Narrative Intelligence

Edited by
Michael Mateas and
Phoebe Sengers

Narrative Intelligence

Advances in Consciousness Research

Advances in Consciousness Research provides a forum for scholars from different scientific disciplines and fields of knowledge who study consciousness in its multifaceted aspects. Thus the Series will include (but not be limited to) the various areas of cognitive science, including cognitive psychology, linguistics, brain science and philosophy. The orientation of the Series is toward developing new interdisciplinary and integrative approaches for the investigation, description and theory of consciousness, as well as the practical consequences of this research for the individual and society.

Series B: Research in progress. Experimental, descriptive and clinical research in consciousness.

Volume 46

Narrative Intelligence
Edited by Michael Mateas and Phoebe Sengers

Narrative Intelligence

Edited by

Michael Mateas
Carnegie Mellon University, Pittsburgh

Phoebe Sengers
Cornell University, Ithaca

John Benjamins Publishing Company
Amsterdam / Philadelphia

∞™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Narrative Intelligence / edited by Michael Mateas and Phoebe Sengers.
p. cm. (Advances in Consciousness Research, ISSN 1381–589X ; v. 46)
Includes bibliographical references and indexes.
1. Narration (Rhetoric)--Psychological aspects. I. Mateas, Michael. II. Sengers, Phoebe. III. Series.

BF456.N37 N37 2002
153-dc21 2002026207
ISBN 90 272 5171 1 (Eur.) / 1 58811 273 X (US) (Hb; alk. paper)
ISBN 90 272 5172 X (Eur.) / 1 58811 274 8 (US) (Pb; alk. paper)

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Table of contents

Chapter 1

Narrative Intelligence

Michael Mateas and Phoebe Sengers
Carnegie Mellon University, Pittsburgh, PA;
Cornell University, Ithaca, NY

Introduction

People are narrative animals. As children, our caretakers immerse us in stories: fairy tales, made-up stories, favorite stories, "Read me a story!" Even when barely verbal, we begin to tell our own proto-stories. As children, narrative frameworks become an important part of the way we learn to approach the world (Nelson 1989).

As adults, we continue to surround ourselves with stories, furnishing our worlds not just with data but with meaning. We say to one another, "Have you heard? Frank and Barb had a fight. She's sick of him letting the dog on the bed. I always told him he'd get in trouble with his permissive ways with that beast." By telling stories we make sense of the world. We order its events and find meaning in them by assimilating them to more-or-less familiar narratives. It is this human ability to organize experience into narrative form that David Blair and Tom Meyer call "Narrative Intelligence" (Blair & Meyer 1997) and around which AI research into narrative coalesces.

A brief history of narrative intelligence

Given the primary importance of narrative in human experience, it is no surprise that story and narrative have long been of interest to AI researchers. In the 1970's and early 80's there was a substantial amount of interest in story understanding and generation in particular. Work in this area was particularly strong in Roger Schank's research group at Yale. Schank and his group explored the issue of what kind of knowledge structures and process a human being must have to understand the meaning of natural language. Since the meaning of a sentence is not determinable in isolation, but requires relating the sentence to sentences around it, to prior experience, and to some larger context, the group's work quickly became focused on understanding narratives. In a series of programs, they developed a theory of the knowledge structures necessary to understand textual narratives. The story-understanding system SAM (Cullingford 1981) used scripts to capture the notion of stereotyped situations or contexts. The scripts captured the typical causal connections holding in a stereotyped situation. The story-understanding system PAM (Wilensky 1981) and the story-generation system Tale-Spin (Meehan 1977) both incorporated a notion of the goals held by characters in a narrative and the various means they have to accomplish these goals. Other work in this group included a model of ideologically-biased understanding (Carbonell 1979), the use of themes to capture aspects of stories more abstract than can be captured just with scripts, plans and goals (Dyer 1983), and a model of narrative memory and reminding (Kolodner 1984).

Work in this area generated an impressive range of systems, particularly given the comparatively primitive hardware technology to which these early researchers were limited. A pleasant discovery for later researchers in re-reading these early reports is a level of charm and wit in system design often unfortunately lacking in contemporary research. Nevertheless, these early narrative systems fell out of favor, suffering from the same fate that befell many 70's AI systems. They were intensely knowledge-based, which meant that they functioned only in very limited domains and could be made more general only by an intensive and probably eventually infeasible knowledge engineering process.

But, perhaps more importantly, as funding for AI dried up during the AI Winter, AI research became more focused on constrained problems with clear, measurable results and immediate practical utility. Researchers tried to make AI more like engineering than like a craft or an art. This required focusing on problems with discrete measurable outcomes in which it is possible to say with certainty that a program achieves or does not achieve the given objective. Yet such a research agenda rules out the ability to work on complex phenomena such as the human use of narratives precisely because the complexity of such a phenomenon rules out the possibility for complete, decisively testable models. Schank makes this clear in his description of the research agenda at Yale (Schank & Reisbeck 1981: 4):

> Thus, for us, theory creation is a process of thought, followed by programming, then by additional thought, with each serving the other. Thus AI really operated under a novel view of science. Normal scientific method holds that first a theory is postulated, and then tested and found to be right or wrong. But in AI our theories are never that complete, because the processes we are theorizing about are so complex. Thus our tests are never completely decisive. We build programs that show us what to concentrate on in building the next program.

Except for occasional exceptions continuing in the Yale tradition, such as Mueller's model of daydreaming (Mueller 1990) and Turner's model of storytelling (Turner 1994), sustained work on narrative disappeared in AI.

Birth of NI

While AI research became refocused, narrative became no less important. Narrative influences simply became felt in other areas of computer science. In these other areas, narrative became an influence as part of a general move towards an interdisciplinary engagement with the humanities. For example, in human-computer interface design, the research focus moved from the hardware interface, through programming language as interface and interactive terminal as interface, to a view of the interface as a computer/human dialog (GUI's are based on this model) and a growing concern with the entire use context (including the social context) as the "interface" (Grudin 1989). This shift in the design focus has been accompanied by a shift in system design methodologies, particularly the adoption of qualitative techniques from the social sciences (e.g. ethnography) and the use of iterative design cycles and rapid prototyping. These new methodologies focus on understanding the use context and managing the inherent incompleteness of any description of that context or the system requirements (Loewgren 1995). As system designers began coping with the rich complexities of designing both for and within a cultural context, they began tapping the long craft tradition of other design fields (e.g. architecture, graphic design, industrial design, etc.) which have been successfully designing artifacts within rich cultural settings for hundreds, if not thousands of years (Winograd 1996). As the field of human-computer interaction became more interdisciplinary (e.g. borrowing anthropological and qualitative sociological techniques), it was just a matter of time before the concept of narrative was examined for interface design principles (Laurel 1991).

Other fields of CS also began tapping humanistic perspectives in general and narrative concepts in particular. For example, in hypertext research, narrative ideas were incorporated, both in the broad sense of narrative as cooperative social construction (Barrett 1989) and in the narrower sense of a narrative as a story written by an author (Bolter and Joyce 1987). Within AI itself, narrative and dramatic concepts reappeared in the form of interactive fiction (Bates 1992, Murray 1998b).

Thus, during the same time period in which AI research abandoned studying complex, culturally grounded phenomena such as meaning in favor of narrowly defined problems with decisive, measurable results, other fields of CS moved in the opposite direction, borrowing and adapting modes of knowledge production from the humanities, design and the arts in order to tackle the complexities of designing computational artifacts for and within cultural contexts. And within this general move towards a humanistic/technical fusion, narrative provides a particularly rich set of ideas for informing such work. It is our contention that this engagement with narrative in other fields of CS has opened up a new opportunity for employing narrative in AI.

Specifically, this work (re)establishes the following conditions within the CS culture:

1. Research methodologies which address rich, complex research questions by employing iterative cycles (e.g. the cycle described in the Schank quote above, where one builds to know what to think and thinks to know what to build) have been revalidated.
2. Interdisciplinary technical work drawing heavily on the humanities, design and the arts has proven useful.
3. Narrative has been recognized as a particularly rich constellation of ideas for informing system design.

The time is ripe for AI to reengage narrative, to explore all the ways in which narrative intersects with intelligence of both the artificial and human varieties. Among the first groups to begin this new exploration was a loose-knit circle of interdisciplinary researchers at the MIT Media Lab (See Chapter 2). They termed this area of work "Narrative Intelligence" (NI). Researchers in the NI group pulled in notions of narrative from other disciplines into a new, creative foment.

The rest of this introduction is structured as follows. The next section, How to Read This Book, provides short descriptions of each of the chapters, organized into topic categories. The following section, Streams of Influence, describes some of the disciplines contributing to NI, and provides short discipline-specific descriptions of how the chapters in this book relate to each of these disciplines. Finally, the section Lay of the Land describes the major themes that have emerged in research in NI and provides short theme-specific descriptions of how the chapters in this book relate to each of these themes.

How to read this book

This book came into being after we organized a symposium on Narrative Intelligence in 1999 (Mateas and Sengers 1999). We were bombarded with submissions, and at the lively and well-attended event itself it was clear that there is a large but latent interdisciplinary community of researchers out there waiting to coalesce around the term Narrative Intelligence. This book is intended to provide these researchers with a focal point: we include historical documents about NI for context (Davis and Travers, Agre, Bruner) and in subsequent chapters attempt to give a feel for the broad range of work, from rhetoric to discourse processing to computer game design, involving systems from natural language processors to interactive autonomous characters to story databases, which is currently happening in and around NI.

Given this plethora of backgrounds, perspectives, approaches, and rhetorical forms, it is important for chapters to be understood in their own contexts. We provide some of that context here. We highly suggest that you continue your tour through the book with Davis and Travers's "A Brief Overview of the Narrative Intelligence Reading Group," which explains the range of influences that form NI and prepares readers for productive engagements with a variety of disciplinary approaches.

Marc Davis and Michael Travers are two of the founders of the NI group at the MIT Media Lab, the major catalyst for much of current NI research. Their "Brief overview of the narrative intelligence reading group" is an intellectual history of the NI group, giving a historical overview of approaches to Narrative Intelligence. Davis and Travers describes how NI research came to exist at the intersection of media theory and artificial intelligence, the difficulties they ran into in trying to synthesize these two approaches, and the work they drew on in a variety of areas to develop what became known as the NI approach.

Human narrative

Marina Umaschi Bers is a researcher at the MIT Media Lab, who has built several "identity construction kits," which support children in thinking through and constructing their identities through the use of story-telling. In "We are what we tell: Designing narrative environments for children," Bers describes how children use narrative in the context of these identity construction environments in order to explore their values. She argues that narrative functions to develop a cohesive sense of self out of diverse and potentially conflicting "subselves."

Jerome Bruner is the founder of Narrative Psychology, an area of research that focuses on how human beings use narrative to understand the world and one another. Narrative psychology insists on the importance of stories in human understanding, contrasting with statistical, logical, and abstract approaches popular in the psychological subfields most often imported into AI. As such, narrative psychology is an important resource for NI researchers who find previously used psychological importations inappropriate for their work. Bruner's classic piece, "The narrative construction of reality," reprinted from *Critical Inquiry*, describes the properties of narrative and how they are used to create understandings of the world, giving researchers an intellectual framework for bringing narrative into their systems.

Kerstin Dautenhahn is a researcher in robotics and agents who focuses on socially intelligent agents and its evolutionary origins in animals, especially primates. Dautenhahn argues that the currently fashionable study of sociality in terms of ant and other 'anonymous' societies is inadequate for understanding what it means to be a social being for humans, primates, and other animals who live in individualized societies with complex forms of social interaction. In "Stories of lemurs and robots: The social origin of story-telling," Dautenhahn relates social intelligence with narrative intelligence, arguing that storytelling has evolved in response to the social structure and social dynamics of primate communities. This means narrative forms an important part of the social glue at least for human societies and perhaps in a prototypical form in primate societies. She describes experiments in agents with a simple ability to use narrative to understand their own and others' behaviors, and work on using social robots to help autistic children to understand increasingly complex social behavior.

Brenda Laurel is one of the pioneers of Narrative Intelligence, having explored the possibility for interactive fiction in her doctoral thesis (Laurel 1986), and subsequently opened up the area of narrative interfaces with Abbe Don and Tim Oren, as described above. More recently, Laurel started a company, Purple Moon, which built software for girls. In "Vital narratives," Laurel discusses the kinds of narrative available in American culture today, arguing for the importance of kinds of narratives that allow for flexible, critical use. She analyses the kinds of relationships which narrative supports, and argues that the best kinds of narratives are inclusive and accessible, stories that can be adapted and made relevant to their listeners' lives.

Story generation

Chris Crawford is a game designer known for several early influential games (Eastern Front, Balance of Power) and for his essays on game design and interactivity (Crawford 1984, 2000). In "Assumptions underlying the Erasmatron storytelling system," Crawford describes the design assumptions underlying the Erasmatron, an interactive story system designed to allow non-programmers (e.g. artists and writers) to build interactive stories.

In "The recombinant history apparatus presents Terminal Time," Steffi Domike, Michael Mateas, and Paul Vanouse describe a novel story generation architecture which generates the spoken narrative, video sequence and sound track for ideologically-biased documentary histories which are generated in response to audience feedback.

In "Story grammars: Return of a theory", R. Raymond Lang argues for the generation of stories using a formalized story grammar. The promise of story grammars is that they provide a formal specification of what is meant by "story structure." Lang describes an implemented grammar-based story generation system called Joseph, and situates this system within the history of story generation research.

In "The Dr. K– Project," artist Brandon Rickman describes a text-based interactive narrative system. Unlike many conventional interactive fiction systems, which simulate virtual environments and then describe them to users, in "Dr. K–" the "world" of the story comes into being and changes based on the history of the interaction. As the audience selects words in the textual narrative, the object or action is brought into focus, causing the system to redescribe the object or action in more detail. Simultaneously, other descriptions may become less detailed, reverting to more generic descriptions. The scene is in continuous flux - the narrative is not unfolded in a linear manner but rather is communicated as a gestalt created by the entire interaction. Rickman contrasts the notion of simulation and fabrication. Where simulations try to provide objective, repeatable, high-fidelity experiences with an emphasis on user control, fabrications provide more indirect user control, a small number of specific viewpoints on the world, and try to expose the representational process. "Dr. K–" can be understood as a fabrication.

In "The Rise and Fall of Black Velvet Flag: An 'Intelligent' System for Youth Culture Documentary" Sheldon Schiffer describes an interactive documentary system that documents the band Black Velvet Flag. A user traverses a custom path through the documentary materials by using a visual query interface. Schiffer is particularly interested in using interactivity to maintain the dynamism of the original source materials. He argues that maintaining this dynamism is particularly appropriate for documenting youth cultural phenomena.

Carol Strohecker, Kevin Brooks, and Larry Friedlander are builders of the interactive fiction system "Tired of Giving In," which tells the story of the US civil rights battle that began with Rosa Parks's refusal to give up her bus seat to a white man. Interaction in this system is based on the notion of the Greek chorus: different characters have different perspectives on unfolding events, and users can ask different characters to tell part of the story. In "Experiments with the theatrical greek chorus as a model for interactions with computational narrative systems," Strohecker, Brooks, and Friedlander describe "Tired of Giving In," and give psychological justification and design sketches for future systems which allow users to shape an interactive fiction and take on different roles through the use of tangible, shared objects.

Agents and narrative

In "Agneta & Frida: Merging web and narrative?," Kristina Höök, Per Persson, and Marie Sjölinder describe the design and evaluation of a concept system for weaving narrative through Web surfing, normally a disjointed series of jumps from page to page. While accompanying users along their Web surf, the characters Agneta and Frida engage in a narrative banter: making (usually sarcastic) comments about what they see on the Web page, commenting on error messages, but also discussing (and living!) their own lives in soap-opera-like vignettes: complaining about the annoying poodle that lives next door or going to the kitchen (off-screen) to make a cup of coffee. The hope is that playing a narrative alongside and connected to the Web browsing experience will help to provide an overall sense of cohesion to the user's experience of Web surfing. Höök, Persson, and Sjölinder designed new user interaction techniques to evaluate such a system, which is not focused on optimizing user functionality, but on providing users with new kinds of experiences.

Katherine Isbister and Patrick Doyle, researchers with roots in Barbara Hayes-Roth's Virtual Theater Project, argue in "Web guide agents: Narative context with charachter" that agents can give human users richer experiences of virtual environments and the Web by telling users stories about the virtual sites they visit together. They analyze the behavior of human tour guides – what sorts of stories they tell, how they decide when and when not to tell stories, how they respond to the cues audiences give them in response. They describe two systems that give tours: one takes humans on a virtual tour of a Japanese castle, the other takes visitors on a tour of a virtual museum.

Sengers employs a cultural-theoretic analysis of the technical assumptions underlying autonomous believable agents to diagnose why the behavior of such agents is often incomprehensible to a human observer. The technical practice of breaking down agents into black-box collections of weakly interacting behaviors results in a lack of behavioral coherence, that is, schizophrenia. The analysis of narrative properties provided by Brunner's narrative psychology (Bruner 1990, 1991) is then used to inform an alternative methodology for the design and implementation of believable agents, a methodology that makes such agents "readable" to a human observer by providing the appropriate cues for inferring a coherent intentional state.

Andrew Stern is a researcher in interactive fiction and believable agents. He is a co-creator of Virtual Petz, one of the first games that allows users to play with (seemingly) intelligent creatures with their own personalities. In "Virtual Babyz: Believable agents with Narrative Intelligence," Stern argues that narratives ("mini-stories") can in fact emerge from the interactions between characters who are modeled as autonomous agents. He describes the engineering and design techniques that were used in order to support the development of narrative in the product Virtual Babyz.

Part IV: Analyzing the stories we tell

Philip Agre's essay "Writing and Representation" was an influential early document within the NI Group at the MIT Media Lab. Agre argues that much work in symbolic representation in AI is influenced by a writing or "text" metaphor which sees representations as effortlessly, without any work on the part of the possessor of the representation, carrying meaning in a context-independent manner. This view of representation has created a series of unsolvable technical impasses within AI. Humanistic critiques (e.g. deconstruction) of the notion of text as a context-independent carrier-of-meaning have revealed that the meaning of a piece of text is a fresh problem in every new context; this meaning is actively constructed by the "user" of the text. Agre explores how this alternative account of writing and text could be used to inform a new approach to representation within AI.

In "Stories and social networks," Sack looks at the interrelationship between stories and social relationships on the internet: which stories get re-told, who cites whom and in what way. He aims for a middle ground between computational linguistics, which he argues generally looks only at the utterances of individuals without concern for their social context, and sociology, in which social networks of storytelling are studied while often ignoring the form and content of the stories involved.

Streams of influence

One of the central aspects of NI work is its inherent interdisciplinarity. If narrative is indeed, as many argue, a fundamental organizing principle of human experience, then it is unsurprising that many different disciplines have an interest in narrative. Work in NI has drawn on conceptions of narrative from many of these sources, including the following. At the end of each discipline description is a list of the papers in this volume which relate to that discipline.

Art

In art, narrative is understood as one, rather powerful, form of representation. Much of contemporary art practice involves self-consciously questioning representational modes, exploring the boundaries, breaking the representation, questioning whose power is being preserved by a representational mode, and hybridizing modes in order to create new ones. Thus, when engaging in narratively-based work, artists rarely tell straightforward narratives employing the standard narrative tropes available within their culture, but rather ironize, layer, and otherwise subvert the standard tropes from a position of extreme cultural self-consciousness. For those studying NI, artistic practice is a useful methodological resource as a way to expose and explore the often unarticulated cultural machinery supporting narrative representation.

In "The recombinant history apparatus presents Terminal Time," Domike, Mateas and Vanouse play with the narrative structure of traditional documentary form by building a system which endlessly replicates this form.

In "The Dr. K– Project," Rickman describes a narrative landscape which, rather than having a mimetic, independent existence, is created in response to audience interaction.

In "The Rise and Fall of Black Velvet Flag: An 'Intelligent' System for Youth Culture Documentary," Schiffer is concerned with maintaining the narrative openness of raw documentary material. He draws on New Media art theory such as Lev Manovich's discussion of database culture and Peter Weibel's discussion of the variable virtual image.

Psychology

In psychology, narrative is thought of as a way in which humans make sense of the world. This notion is particularly advanced in Jerome Bruner's work on narrative psychology (Bruner 1990, 1991). Bruner argues that narrative is fundamental to human understanding of intentional behavior, i.e. that humans make sense of intentional action by assimilating it into narrative structures. This argument is used as a basis for making systems from interfaces to intelligent agents more understandable, by communicating in ways that are easy to assimilate to narrative (Don 1990, Sengers 1999).

In "The narrative construction of reality," Jerome Bruner describes the fundamental properties of human narrative, used as a basis by NI researchers for understanding how to make narrative a part of computational systems.

In "Stories and social networks," Sack explores narrative as a "technology of the self," providing tools that support analysis of identity construction through the telling and re-telling of stories within a social group.

In "Agneta & Frida: Merging web and narrative?," Persson, Höök, and Sjölinder address the human drive to create coherence out of disparate data.

In "Web guide agents: Narrative context with character," Isbister and Doyle analyse the use of narrative by human tour guides to make unfamiliar environments understandable and interesting.

In "Stories of lemurs and robots: The social origin of storytelling," Dautenhahn describes the phylogenetic origins of human narrative intelligence in primate social intelligence. She underlines the importance of telling stories to construct an autobiography, the groundwork for a sense of self. She describes problems in social and narrative intelligence in autistic people, and proposes the use of robots in therapy to develop social and narrative intelligence.

In "We are what we tell: Designing narrative environments for children," Bers motivates and describes the use of narrative in identity construction kits, virtual environments which help children to develop a coherent sense of self and their values.

In "Schizophrenia and narrative in artificial agents," Sengers uses the principles of narrative psychology as derived from Bruner in order to construct artificial agents which are narratively understandable to human users.

Cultural studies

In cultural studies, narrative is studied as a way in which a culture structures and propagates knowledge. Because humans quickly internalize narrative, it is an important form of collective knowledge and can be a basis for ideological manipulation. NI researchers using this concept of narrative are often interested in social or collective forms of narrative and in uncovering hidden narratives. This study of narrative can be reflexively applied to AI research itself, leading to transformations of AI practices. That is, an analysis of the narrative structures and metaphors used to tell the story of progress within AI can illuminate systematic problems caused by these narratives and point the way to new research approaches (Agre 1997; Sack 1992; Sengers 1998; Mateas 2001).

In "Writing and representation," Agre draws on cultural-theoretic analyses of representation to explore unexamined assumptions regarding representation in AI.

In "Vital narratives," Laurel analyses the cultural roles played by different kinds of narratives. She argues that the inflexibility of certain kinds of narrative, such as fundamentalist religious narratives, creates the potential for a great deal of unproductive cultural conflict, while in other cases potentially helpful narratives such as scientific narratives are hindered because they are not made relevant to people's everyday lives. She argues that we need to make ethical decisions about what kinds of narratives, both interactive and noninteractive, we are creating, arguing for flexible, inclusive narratives.

In "The recombinant history apparatus presents Terminal Time," Domike, Mateas and Vanouse explore the role that ideology plays in the construction of history by building a system which caricatures ideologically-biased historical reasoning.

In "The Rise and Fall of Black Velvet Flag: An 'intelligent' system for youth culture documentary," Schiffer explores the role of youth culture in the construction of identity. Narrative constructions of youth culture must be responsive to massive change on the part of the subjects (who in turn are the audience) of this identity formation process.

In "Stories and social networks," Sack uses tools from computational linguistics to support media-studies analysis of social responses to mass media. His work is based on an understanding of the importance of narrative in people's daily lives, and the agency of social groups in retelling and reincorporating cultural narratives. A novel part of his approach from a cultural-studies perspective is that he creates a tool for cultural studies that is also usable by the people being studied.

In "Schizophrenia and narrative in artificial agents," Sengers argues for the similarity of the perspectives of behavior-based artificial agents and institutional psychiatry using the tools of cultural studies. She argues that technical problems in the coordination of behaviors can be traced historically to symptomatic, atomizing approaches to understanding human subjectivity. She builds on the arguments of the anti-psychiatric movement of the 60's to argue that human subjectivity should be represented in artificial agents using a hermeneutic approach which includes narrative.

Literary studies

Literary studies are particularly concerned with analyzing the properties of stories as narrative. These properties can then be used as a basis for story-generation or understanding systems. For example, Vladimir Propp's analysis of the structure of folk tales (Propp 1969) has served as an inspiration for many AI researchers (e.g. Meehan 1977, Turner 1992, Weyhrauch 1997).

More generally, literary studies and literary theory embrace an enormous spectrum of perspectives on story, narrative, and their function in our culture, from Aristotle's theory of poetics to New Criticism to speech act theory to structuralism to Reader Response theory to postmodernism and beyond. Each of these strands involves novel ways of thinking about narrative and its place in human experience that can be tapped for work in NI – the surface has barely been scratched.

In "Story grammars: Return of a theory," Lang provides a brief history of story grammars, a structuralist attempt to formally capture the structure of folktales within a given culture.

In "Vital narratives," Laurel analyses narratives along four different axes: the kinds of relationships they support, their relevance to people's daily lives, the strategies they help fulfill, and their epistemological value.

Drama

Drama is the performance of stories in front of an audience in real-time (i.e. plays and movies). Dramatic stories have different properties from literary stories (i.e. novels); following Laurel (Laurel 1991), dramatic stories have the properties of enactment, intensification, and unity of structure, vs. literary stories which have the properties of description, extensification, and episodic structure. Given the affinity between drama's focus on action and the action-based, real-time, responsive behavior of interactive computer systems, researchers have begun tapping the dramatic tradition, particularly within the areas of interface design and interactive drama (Laurel 1991, Bates 1992, Hayes-Roth, van Gent, and Huber, 1997, Mateas and Stern 2000, Mateas 2000).

In "Agneta & Frida: Merging web and narrative,?" Persson, Höök and Sjölinder use inspiration from film theory to design characters for a narrative interface.

NI is humanistic AI

As you might imagine, this highly divergent list of influences (as well as multiple understandings and definitions of the concept of narrative) has led to a healthy foment in the field. While it is not unusual for AI researchers to draw from a wide variety of other fields for inspiration, it is unusual for those fields to be largely humanistic. We believe this is a special source of interest for NI: it is a field where not only scientific but also humanist notions of experience and humanity fruitfully inform technological conceptions.

The lay of the land

Drawing on a diverse range of influences, researchers have (often independently) explored a wide variety of topics relevant to NI. In the process, several common themes have emerged. At the end of each theme description is a list of papers which relate to that theme.

Narrative interfaces

Several researchers in the field of HCI argue that narrative should be used as a basis for human-computer interfaces (Don 1990, Laurel 1991). If humans often make sense of the world by assimilating it to narrative, then it makes sense to design our systems so as to allow people to use their well-honed narrative skills in interpreting these systems. For example, Don (Don 1990) borrows concepts from the oral storytelling tradition to organize the interface for a multimedia knowledge base. Specifically, she describes three properties of oral storytelling that can guide interface design: storytellers adapt the story to the reactions of the audience, information such as names and lists are embedded within the storyline so that the audience experiences this information as events unfolding in time, and characters with predictable traits are used to prime expectations. Laurel (Laurel 1991) uses the analytic categories of Aristotelian dramatic theory (i.e. spectacle, song, diction, thought, character and plot) to organize interface design.

In "Agneta & Frida: Merging web and narrative?," Persson, Höök and Sjölinder create an interface plug-in which is intended to help people create a narrative understanding of a non-narrative interface, i.e. the Web. In evaluating the system, they develop new HCI methods for evaluating narrativity based on metaphor analysis.

In "Web guide agents: Narrative context with character," Isbister and Doyle construct interface agents as characters who guide users through unfamiliar locations, providing both social context and narrative content.

Narrative agent design

The HCI argument that systems will be more understandable with narrative presentation extends to systems involving artificial agents. Since, as narrative psychologists argue, humans use narrative in particular for understanding intentional behavior, several researchers argue that agents will be more comprehensible if their visible behavior is structured into narrative (Sengers 1999; Lester & Stone 1997). This generally involves the construction of agent architectures that allow agents to make behavioral choices based on the narrative structure of the resulting behavior, often including transition behaviors that knit the agent's various activities into a coherent, narrative whole.

In "Schizophrenia and narrative in artificial agents," Sengers describes such an architecture for narratively understandable agents.

In "Virtual Babyz: Believable agents with narrative intelligence," Stern describes agents that are designed to allow a narrative structure to emerge from their behavior as they act over time.

Agents that use narrative structure

If narrative is one central component of human intelligence, then it should also play an important role in artificial agents which model aspects of human intelligence (Schank 1990; Dautenhahn & Nehaniv 1998). Roger Schank, for example, has developed a model of the interrelationship between stories and memory, describing how stories are understood and how they are recreated from the remembered "gists" of stories. Elsewhere and in "Stories of lemurs and robots: The social origin of storytelling ," Kerstin Dautenhahn argues that human (and possibly animal) experience in the world is shaped by our autobiographies, narratives we tell ourselves about our past and the pasts of other agents (Dautenhahn 1998).

In "Web guide agents: Narrative context with character," Isbister and Doyle describe agents that can use special annotations on web sites to gain access to narrative structure of information and to be able to relay this to human users.

Support for human storytelling

Since stories are an important part of human life, several researchers have begun building systems that support people in telling stories to one another. Some of these systems, such as Kimiko Ryokai's Storymat (Ryokai & Cassell 1999), record and play back stories that people have told. Others, like Marina Umaschi Bers's SAGE Storytellers (Umaschi 1997) and Kevin Brooks' Agent Stories (Brooks 1997), allow people to create their own interactive storytellers and stories.

In "Assumptions underlying the Erasmatron storytelling system," Crawford describes how the Erasmatron is specifically designed to ease the burden for the non-programmer of interactive story design.

In "Agneta & Frida: Merging web and narrative?," Persson, Höök, and Sjölinder aim to support human narrative understanding of the normally non-narrative Web.

In "Stories of lemurs and robots: The social origin of storytelling," Dautenhahn describes a robotic system for helping autistic children to be able to understand human behavior, providing them with training for narrative intelligence.

In "We are what we tell: Designing narrative environments for children," Bers describes several virtual environments based on constructionist learning principles which can be used by children to explore and develop their identities and values in a community, using narrative as an essential element for developing a coherent sense of self.

Story database systems

Some researchers have found it useful to design systems which allow humans to access databases of stories. Presenting information in the form of narratives, they argue, makes it easier and more pleasant for people to process the information. Schank (Schank 1997) has built a training system, Ask Tom, on this principle. It contains a database of stories describing how people have handled commonly occurring problem situations; these stories are triggered by the system when the trainee faces a similar situation. Another example of this kind of work is IBM Research's project on Knowledge Socialization, which looks at

– among other things – ways in which story databases can be used to transfer informal knowledge (Lawrence & Thomas 1999). Cassell and Smith's Victorian Laptop combines a story database with a storytelling support system. As people write their own travel stories, the system retrieves matching stories from a database of Victorian travel narratives, allowing them to compare their experiences with those of travelers from the past (Cassell & Smith 1999).

Story-understanding systems

Story-understanding systems seek to model the processes by which a human "understands" a story. "Understanding" is usually operationalized as the ability to answer questions about a story where the answers are not explicitly given within the story, or as the ability to paraphrase or summarize a story. In order to perform these tasks, story-understanding systems form representations of stories more amenable to manipulation than the surface form, make connections between the stories and some context or background knowledge, and possibly have models of story event importance. Research in story understanding began during AI's classical engagement with narrative (see the section "A Brief History of Narrative Intelligence" above). Even after the shift in AI research agendas following the AI Winter, a small stream of such work continued (e.g. Cox 1996). This body of work plays an important role within NI. By exploring what it means to be the kind of system (either natural or artificial) that understands stories, this work can help inform the design of agents and interfaces that make use of narrative. For example, Sack's work on automatically understanding ideological bias of news stories highlights the importance for narrative of understanding not only the content of what is said, but also the viewpoint that leads it to be told in particular ways (Sack 2001).

In "Stories and social networks," Sack develops a story-understanding system which focuses not on the story itself, but on understanding how people use stories socially and to construct identity.

Story generation systems

Storytelling systems seek to model the knowledge and processes necessary to tell a story. Following Bailey (Bailey 1999), work in storytelling systems can be divided into three major groups: author-centric, story-centric, and character-centric systems (Bailey refers to character-centric systems as world models). Author-centric systems model the thought processes of an author. Character-centric systems model the goals and plans of characters; stories result from characters pursuing their autonomous goals. Story-centric systems model the structural properties of stories themselves (viewing the story as an artifact); the system tells stories by manipulating this structural artifact.

Like story understanding, storytelling work also began during AI's classical engagement with narrative. Interestingly, the three perspectives outlined above all emerged during this classical engagement at roughly the same time. Perhaps the most famous early storytelling system is Tale-Spin (Meehan 1977). Tale-Spin is a character-centric system, modeling the goals and plans of animal characters taken from Aesop's fables. Ani (Kahn 1979), an author-centric system, generates an animation (using a square, triangle and circle to represent characters) telling a simplified version of Snow White. The system is given a high level script describing the authorial goals for the story (what should be conveyed); given this script, it makes all the detailed animation decisions necessary to tell the story. Rumelhart (Rumelhart 1975) takes a story-centric approach, capturing the notion of story as a story grammar. For more detailed descriptions of the history of story-telling systems, see Lang in this volume (Chapter 12).

All three storytelling approaches tend to utilize some form of combinatorial search over a space of primitive story elements. Elliott (Elliot 1998) has explored an alternative approach. His system, while using a fixed script, tells different stories by narrating the stories with different emotional emphases. The emotional behavior of the narration agent is generated by the Affective Reasoner, a cognitive appraisal model of emotion. Elliott's work demonstrates that a storytelling system can leverage the interpretive capabilities of a human observer, in this particular case the ability to understand motivations and emotions.

In "The recombinant history apparatus presents Terminal Time," Domike, Mateas and Vanouse describe a system which generates ideologically-biased histories in response to audience feedback. While Terminal Time is audience interactive, its architecture is influenced more by work in story generation than work in interactive fiction and drama.

In "Story grammars: Return of a theory," Lang describes an implemented story grammar that generates stories in the style of ethnic folktales.

Interactive fiction and drama

The field of interactive fiction and drama seeks to build systems that let the audience experience the story as an interactive participant (this includes, but is not limited to, being a first-person protagonist). System – building work in this area includes approaches that don't specifically make use of AI techniques, such as hypertext fiction and text and graphical adventure games. These approaches have been quite fruitful for exploring the nature of interactivity and the structural possibilities of interactive narrative (Murray 1998b). But for the purposes of this brief overview, we will focus on AI-based approaches to interactive fiction and drama.

Most of the work in interactive drama has approached it from an autonomous-agents perspective. The focus has been on building believable agents that can play roles in stories. The Oz Project built an agent architecture (Loyall & Bates 1991; Loyall 1997) including a model of emotion (Reilly & Bates 1992; Neal Reilly 1996) to support the construction of autonomous characters. Hayes-Roth built agents that improvise activity around a fixed script (Hayes-Roth, van Gent & Huber 1997). Blumberg was originally motivated by the ALife goal of building computational instantiations of ethological models of action selection (Blumberg 1994), but more recently has focused on building architectures to support the construction of characters (Kline & Blumberg 1999). Most of the believable agents architectures make use of some reactive action-selection framework, though there has been some work on using planning techniques to ease the authorial burden (Rizzo et al. 1998).

There has been less work on building systems to support interactive plot. Some work has focused on systems that provide high level plot guidance to believable agents. For example, Weyhrauch (Weyhrauch 1997) built a dramatic guidance system that issues high-level commands to Oz believable agents. Blumberg and Galyean (Blumberg & Galyean 1995) explored the concept of a director giving commands to autonomous characters at multiple levels of abstraction. Other work has focused on tracking the user's progress through a fixed plot, using user actions to trigger the next part of the story. For example, Galyean (Galyean 1995) built a system that uses cinematic techniques adapted to virtual reality to guide a user through a plot. Pinhanez (Pinhanez 1997) built a system that uses a temporal calculus to trigger story events given user actions. Mateas and Stern (Mateas & Stern 2000) are building an interactive drama system which blurs the distinction between strongly autonomous characters and high-level plot control by intermingling believable agent behaviors and plot constructs.

In "Assumptions underlying the Erasmatron storytelling system," Crawford describes the Erasmatron, a system for authoring and playing interactive stories.

In "The Dr. K– Project," Rickman describes a text-based interactive narrative system based on a historical account of two murderers in 1820's Edinburgh.

In "Virtual Babyz: Believable agents with narrative intelligence," Stern describes a character-centric approach to interactive fiction, constructing interactive agents that allow simple plot to emerge from their interactions with one another and the user. He concludes that it is possible to use this to generate loose interactive plots, but that more support for top-down management of plot is needed in order to create tight, well-crafted plots.

Digital interactive video

For digital interactive video, systems automatically construct videos from a database of video clips with interactive guidance from the user. These are closely related to interactive storytelling systems, but face their own range of technical problems because of the use of video material. In Davenport's Autonomist Storyteller System (Davenport & Murtaugh 1997), each video clip is annotated to specify its potential run-time use in a narrative sequence. A search engine assembles the clips into a narrative sequence in real-time. In Synthetic Interviews (Marinelli & Stevens 1998), annotated clips are retrieved in response to utterances processed by a speech recognition engine. This allows a user to have a conversation with video characters; this conversation can be part of a story arc.

In "The recombinant history apparatus presents Terminal Time," Domike, Mateas and Vanouse describe a system which constructs ideologically-biased documentary histories from a database of video and audio material.

In "The Rise and Fall of Black Velvet Flag: An 'intelligent' system for youth culture documentary," Schiffer describes an interactive video system which allows a user to explore their own paths through a database of documentary materials.

Narrative for meta-analysis

AI researchers are people, too. As such, narrative plays an important role in AI research. Some researchers, particularly in cultural studies, study the kinds of narratives AI researchers use in talking about their own work, and how such narratives are woven into choices about what kind of research is worth pursuing (Hayles 1999; Helmreich 1998; Doyle 1997; Sack 1997). A number of AI researchers in turn believe that studying the narratives AI researchers themselves tell can lead to a better self-understanding for AI, and, in turn, yield better AI research (Agre 1997, Sengers 1998, Mateas 2001).

In "Writing and representation," Agre examines the stories AI researchers tell about human representation use, and counters with alternative stories about representation use.

In "Schizophrenia and narrative in artificial agents," Sengers critically analyses the way in which AI researchers talk about their agents, discovering similarities to descriptions of schizophrenic patients. These stories reveal the extent to which AI researchers think of and build their agents as simple mechanisms, although they should appear as complex, living beings. She suggests a compromise approach, in which the agent is thought of as a mechanism which can take advantage of the human propensity to create narrative explanation to create the appearance of living action.

Narrative is many, not one

NI is radically interdisciplinary, drawing on narrative concepts from many humanistic fields of study. Narrative is not a single entity or a single, tightly related set of concepts. As the term is used in humanistic discourse, narrative can mean many things. Narrative can mean a tightly woven story communicated by a strong authorial voice to an audience. Narrative can mean the internal imposition of coherence by which a person makes sense of her life, or the communally constructed group memory by means of which a group organizes past experience. In the broadest sense, narrative can mean an entire worldview (as in "grand" or "master" narrative). And within each of these gross distinctions, there lie yet more distinctions. For example, within the notion of narrative as a tightly woven, author-given story, there lie distinctions such as literary, cinematic and dramatic stories, each of which has its own set of properties and corresponding inspirations and design implications for NI researchers. Thus narrative is a family resemblance concept, a cover term for a rich set of ideas.

The richness of narrative presents some interesting challenges for the emerging field of NI. One challenge is to maintain open lines of communication; with so many different inflections of the concept of narrative, workers will have to make an effort to be clear on the notion of narrative they are using and how it relates (or doesn't relate) to other notions of narrative.

Anoter challenge will be to remain true to the richness of narrative. AI, like the rest of computer science, tends to prefer general and abstract formulations. Applied to narrative, this will result in the attempt to assimilate all narrative phenomena to a single, simplified formulation. In order to build systems, abstraction and simplification are necessary tools. The danger lies in forgetting for what purpose a simplification was made or parhaps that a simplification has even occurred. With a concept as complex and evocative as narrative, there will be particulaly strong pressure to side simplification. If this were to happen, the original richness of narrative, an endless source of inspiration and delight, would be lost.

References

Agre, Philip, E. (1997). *Computation and human experience*. Cambridge: Cambridge University Press.

Bailey, Paul (1999). Searching for storiness: Story-generation from a reader's perspective. In M. Mateas & P. Sengers (Eds.) *Narrative Intelligence: Papers from the 1999 fall symposium* (Technical Report FS-99-01). Menlo Park: AAAI Press.

Barrett, Edward (1989). *The society of text: Hypertext, hypermedia and the social construction of knowledge*. Cambridge, MA: MIT Press.

Bates, Joseph (1992). Virtual Reality, art, and entertainment. *Presence: The journal of teleoperators and virtual environments, 1*, 133–138.

Blair, David & Tom Meyer (1997) Tools for an interactive virtual cinema. In R. Trappl and P. Petta (Eds.), *Creating personalities for synthetic actors: Towards autonomous personality agents*. Berlin: Springer Verlag.

Blumberg, Bruce (1994). Action-selection in Hamsterdam: Lessons from ethology. In *From animals to animats: Proceedings of the third international conference on the simulation of adaptive behavior*. Cambridge, MA: MIT Press.

Blumberg, Bruce & Tinsley Galyean (1995). Multi-level direction of autonomous creatures for real-time virtual environments. In *Proceedings of SIGGRAPH* 95 (pp. 47–54). ACM Press.

Brooks, Kevin (1997). Programming narrative. In *IEEE Symposium on visual languages* (pp. 380–386).

Bolter, Jay David & Michael Joyce (1987). Hypertext and creative writing. In *Hypertext '87 proceedings* (pp. 41–50). Chapel Hill: ACM Press.

Bruner, Jerome (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Bruner, Jerome (1991). The narrative construction of reality. *Critical inquiry, 1*, 1–21.

Carbonell, Jaime (1979). Subjective understanding: Computer models of belief systems. Ph.D. Thesis, Computer Science Department, Yale University.

Cassell, Justine & Jennifer Smith (1999). The Victorian Laptop. In M. Mateas & P. Sengers (Eds.) *Narrative Intelligence: Papers from the 1999 fall symposium* (Technical Report FS-99-01). Menlo Park: AAAI Press.

Cox, Michael (1996). Introspective multistrategy learning: Constructing a learning strategy under reasoning failure. Ph.D. Thesis. Technical Report GIT-CS-96/01. Computer Science Department, Georgia Institute of Technology. Atlanta, Georgia.

Crawford, Chris (1984). *The art of computer game design*. Berkeley: Osborne/McGraw-Hill.

Crawford, Chris (2000). Understanding interactivity. Self-published (available at http://www.erasmatazz.com/book.html).

Cullingford, Richard (1981). SAM. In R. Schank & C. Riesbeck (Eds.), *Inside computer understanding: Five programs plus miniatures* (75-119). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Dautenhahn, Kerstin (1998). Meaning and embodiment in life-like agents. In C. Nehaniv (Ed.), *Plenary working papers in computation for metaphors, analogy and agents* (pp. 24–33). University of Aizu Technical Report 98-1-005.

Dautenhahn, Kerstin & Chrystopher Nehaniv (1998). Artificial Life and natural stories. In *International symposium on Artificial Life and Robotics*: Volume 2 (pp. 435–439). Beppu, Oita, Japan.

Davenport, Glorianna & Michael Murtaugh (1997). Autonomist storyteller systems and the shifting sands of story. *IBM systems journal, 3*, 446–456. Reprint Order No. G321-5652.

Don, Abbe (1990). Narrative and the interface. In Brenda Laurel (Ed.), *The art of human-computer interface design* (pp. 383–391). Reading, MA: Addison-Wesley.

Doyle, Richard (1997). *On beyond living: Rhetorical transformations of the life sciences* (Writing Science). Stanford University Press.

Dyer, Michael (1983). *In depth understanding: A computer model of integrated processing for narrative comprehension*. Cambridge, MA: MIT Press.

Elliott, Clark, Jacek Brzezinski, Sanjay Sheth & Robert Salvatoriello (1998). Story-morphing in the Affective Reasoning paradigm: Generating stories semi-automatically for use with emotionally intelligent multimedia agents. In *Proceedings of the second international conference on autonomous agents* (pp. 181–188). New York: ACM Press.

Galyean, Tinsley (1995). *Narrative guidance of interactivity*. MIT Media Lab Ph.D. Thesis.

Grudin, Jonathan (1989). The computer reaches out: The historical continuity of interface design. In J. Carrasco-Chew & J. Whiteside (Eds.), *Proceedings of the CHI'89 conference on human factors in computer systems* (pp. 141–144). New York: ACM Press.

Hayles, N. Katherine (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. Chicago: University of Chicago Press.

Hayes-Roth, Barbara, Robert van Gent, & Daniel Huber (1997). Acting in character. In R. Trappl and P. Petta (Eds.), *Creating personalities for synthetic actors* (pp. 92–112). Berlin & New York: Springer.

Helmreich, Stefan (1998). *Silicon second nature: Culturing Artificial Life in a digital world*. University of California Press.

Kahn, Ken (1979). Creation of computer animation from story descriptions. Ph.D. Thesis, MIT Artificial Intelligence Lab. AI technical report 540. Boston, MA.

Kline, Christopher & Bruce Blumberg (1999). The art and science of synthetic character design. In *Proceedings of the AISB 1999 symposium on AI and creativity in entertainment and visual art*. Edinburgh, Scotland.

Kolodner, Janet (1984). *Retrieval and organizational strategies in conceptual memory: A computer model*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Laurel, Brenda (1986). *Towards the design of a computer-based interactive fantasy system*. Ph.D. Thesis. The Ohio State University.

Laurel, Brenda (1991). *Computers as theatre*. Reading, MA: Addison-Wesley.

Lawrence, Deborah & John Thomas (1999). Social dynamics of storytelling: Implications for story-base design. In M. Mateas & P. Sengers (Eds.) *Narrative Intelligence: Papers from the 1999 fall symposium* (Technical Report FS-99-01). Menlo Park: AAAI Press.

Lester, James & Brian Stone (1997). Increasing believability in animated pedagogical agents. In W. Lewis Johnson (Ed.), *Proceedings of the first international conference on autonomous agents* (pp. 16–21). ACM Press.

Lieberman, Henry (1995). The visual language of experts in graphic design. In V. Haarslev (Ed.) *Proceedings of the 11th IEEE symposium on visual languages*. Darmstadt, Germany.

Loewgren, Jonas (1995). Perspectives on usability. Technical Report LiTH-IDA-R-95-23. Department of Computer and Information Science, Linkoping University, Linkoping Sweden.

Loyall, A.Bryan & Joseph Bates (1991). Hap: A reactive, adaptive architecture for agents. Technical Report CMU-CS-91-147, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Loyall, A.Bryan (1997). Believable agents. Ph.D. Thesis, Department of Computer Science, Carnegie Mellon University. Technical Report CMU-CS-97-123.

Marinelli, Don & Scott Stevens (1998). Synthetic interviews: The art of creating a "dyad" between humans and machine-based characters. In *Proceedings of the sixth ACM international multimedia conference on technologies for interactive movies* (pp. 11–16). ACM Press.

Mateas, Michael, Steffi Domike & Paul Vanouse (1999). Terminal Time: An ideologically-biased history machine. *AISB quarterly: Special issue on creativity in the arts and sciences, 102*, 36–43.

Mateas, Michael & Phoebe Sengers (Eds.) (1999). *Narrative Intelligence: Papers from the 1999 fall symposium* (Technical Report FS-99-01). AAAI Press.

Mateas, Michael & Andrew Stern (2000). Towards integrating plot and character for interactive drama. In *Working notes of the socially intelligent agents: human in the loop symposium*, 2000 AAAI fall symposium series (pp. 113–118). Menlo Park, CA.: AAAI Press.

Mateas, Michael (2000). A preliminary poetics for interactive drama and games. In *Proceedings of SIGGRAPH 2001, art gallery, art and culture papers* (pp. 51–58).

Mateas, Michael (2001). Expressive AI: A hybrid art and science practice. *Leonardo: Journal of the international society for arts, sciences, and technology, 2*, 147–153.

Meehan, James (1977). *The metanovel: Writing stories by computer*. Ph.D. Thesis. Ann Arbor: University Microfilms International.

Mueller, Erik. (1990). *Daydreaming in humans and machines: A computer model of the stream of thought*. Norwood, New Jersey: Ablex.

Murray, Janet (1998). Building coherent plots in interactive fiction. *IEEE Intelligent Systems*, November/December 1998, 18–21.

Murray, Janet (1998). *Hamlet on the Holodeck*. Cambridge, MA: MIT Press.

Neal Reilly, W. Scott (1996). *Believable social and emotional agents*. Ph.D. Thesis, Department of Computer Science, Carnegie Mellon University. Technical Report CMU-CS-96-138.

Nehaniv, Chrystopher & Kerstin Dautenhahn (1998). Embodiment and memories – Algebras of time and history for autobiographic agents. In R. Trappl (Ed.), *Proceedings of the 14th European meeting on cybernetics and systems research symposium on embodied cognition and Artificial Intelligence*: Vol. 2 (pp. 651–656).

Nelson, Katherine (Ed). (1989). *Narratives from the crib*. Cambridge, MA: Harvard University Press.

Pinhanez, Claudio (1997). Interval scripts: A design paradigm for story-based interactive systems. In *Proceedings of CHI* 97 (pp. 287–294).

Propp, Vladimir (1969). L. Scott (Trans.), L. Wagner (Ed.) *Morphology of the folktale*. 2nd ed. Austin: University of Texas Press.

Reilly, W.Scott & Joseph Bates. (1992). Building emotional agents. Technical Report CMU-CS-92-143, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA.

Rizzo, Paola, Manuela Veloso, Maria Miceli, & Amedeo Cesta (1998). Goal-based personalities and social behaviors in believable agents. *Applied Artificial Intelligence, 13*, 239–271.

Rumelhart, David E. (1975). Notes on a schema for stories. In D.G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211–236). New York: Academic Press, Inc.

Ryokai, Kimiko & Justine Cassell (1999). StoryMat: A play space with narrative memories. In *Proceedings of the 1999 international conference on intelligent user interfaces*. posters/demonstrations (201).

Sack, Warren (2001). Actor-role analysis: Ideology, point of view and the news. In W. Van Peer & S. Chatman (Eds.) *New perspectives on narrative perspective*. New York: SUNY Press.

Sack, Warren (1997). Artificial human nature. *Design issues, 13* (Summer 1997), 55–64.

Sack, Warren (1992). Knowledge compilation and the language design game. In C. Frasson, G. Gauthier, & G. McCalla (Eds.), *Intelligent tutoring systems, Second international conference* (Lecture notes in computer science). Berlin: Springer-Verlag.

Schank, Roger & Reisbeck, Christopher (Eds.) (1981). *Inside computer understanding: Five programs plus miniatures*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Schank, Roger (1990). *Tell me a story: A new look at real and artificial memory*. New York: Scribner.

Schank, Roger (1997). *Virtual Learning: A revolutionary approach to building a highly skilled workforce*. McGraw-Hill.

Sengers, Phoebe (1998). *Anti-boxology: Agent design in cultural context*. Ph.D. Thesis, School of Computer Science, Carnegie Mellon University. Technical Report CMU-CS-98-151.

Sengers, Phoebe (1999). Designing comprehensible agents. In *Sixteenth international joint conference on Artificial Intelligence*: Vol 2 (pp. 1227–1232).

Turner, Scott R. (1994). *The creative process: A computer model of storytelling and creativity*. Mahwah, NJ: Lawrence Erlbaum.

Umaschi, Marina (1997). Soft toys with computer hearts: Building personal storytelling environments. In *CHI'97 proceedings* (pp. 20–21). ACM Press.

Weyhrauch, Peter (1997). Guiding interactive drama. Ph.D. Thesis, School of Computer Science, Carnegie Mellon University. *Technical Report CMU-CS-97-109*. Pittsburgh, PA.

Wilensky, Robert (1981). PAM. In Roger Schank and Christopher Riesbeck (Eds.), *Inside computer understanding: Five programs plus miniatures* (pp. 136–179). Hillsdale, NJ: Lawrence Erlbaum Associates.

Winograd, Terry (Ed.) (1996). *Bringing design to software*. New York, N.Y.: ACM Press.

Chapter 2

A brief overview of the Narrative Intelligence Reading Group

Marc Davis and Michael Travers
University of California at Berkeley; Afferent Systems, Inc. San Francisco

Introduction

In the fall of 1990 at the MIT Media Laboratory, we started a weekly, student-run reading group to explore topics at the intersection of artificial intelligence and literary theory. The group, which we named Narrative Intelligence (NI), quickly took on a life of its own and became a forum in which ideas from philosophy, media theory, and psychology could combine with current research in computational theories of mind and media. Meeting in the basement of the MIT Media Laboratory (so we were both literally and figuratively underground), the group grew to include a local membership of students and faculty from the Media Lab, other MIT departments, Harvard University, and Brown University. The group's mailing list included a large cadre of remote members including luminaries in the fields of human computer interaction, artificial intelligence, and film analysis. Narrative Intelligence became a vital hotbed of interdisciplinary thinking and exploration for its members. It brought together humanists and engineers in the creation of a new cross-disciplinary activity connecting insights from artificial intelligence, media studies, and human computer interaction design. What started out as an attempt to create common ground between two Media Lab students eventually influenced the culture and curriculum of the Media Lab and MIT through the students and faculty that participated in this unofficial, multi-year, interdisciplinary seminar. In this article we describe the formation of the Narrative Intelligence Reading Group, its goals, the core texts and issues it engaged with, reflect on lessons learned, and talk about the future of narrative intelligence.

Formation of the NI Reading Group

In 1990, two graduate students at the MIT Media Lab were trying to talk with each other about topics that seemed of common interest. One, a humanist new to computing (Marc), wanted to build programs that could automatically assemble short movies from archives of video data. The other, a computer scientist with an interest in literary theory (Mike), wanted to program software agents that could understand a simulated world, each other, and themselves. What they found was that while their areas of interest seemed to have common issues (narrative theory and comprehension, knowledge representation, story understanding and generation, and user interface design), the discourse each used to talk about these areas was largely unintelligible to the other. Even such core ideas as "representation," "language," and "communication" meant different things to each of them. Rather than throw up their hands, they wagered that if they could get each other to read the core texts in their respective disciplines, they might be able to construct a common language and a useful discourse. Together with a group of Media Lab students they formed the Narrative Intelligence Reading Group (NI) that met weekly in the basement of the Media Lab building for six years. From 1990–1993, Marc Davis, Mike Travers, and Amy Bruckman actively led and facilitated the group, from 1994–1996, Amy Bruckman and Warren Sack continued the group for another three years. From 1997 to the present, the Narrative Intelligence Reading Group has functioned as a mailing list and resource for its members and others interested in its topics (ni@media.mit.edu).

Motivations for a new interdisciplinary discourse

As early graduate students in the Media Lab, we were faced with trying to synthesize an intellectual framework in which we could situate our work. The desire of the founders and early members of NI to create a common discourse and practice connecting artificial intelligence and literary theory also stemmed from a growing frustration with the limits of our respective disciplines in their ability to inform the analysis, design, and construction of computational media.

In artificial intelligence (AI) we encountered a discipline founded on logicist, formalist, and objectivist conceptions of language, cognition, and computation. Alternative approaches to AI included connectionism and situated action, but both of these schools lacked a coherent theory of representation. Furthermore, existing theories of representation in AI reflected a bias toward understanding representation as based only on textual, logical, or mathematical constructs, and as such, did not offer useful models for thinking about non-textual media. Within the scope of AI and cognitive science, important steps toward a theory of cognition and representation informed by narrative theory had been made by Roger Schank and his students, by George Lakoff and Mark Johnson, and by Phil Agre's work on writing and representation. The analysis and extension of these texts was a key part of our mutual education in NI.

In literary theory, including post-structuralist, semiotic, and reader-response approaches, we found a discipline that had undergone radical transformation and creative growth since the 1960s, but which had been largely oblivious to, and uninfluenced by, both the theoretical roots of computation in the early part of the century and the theoretical implications of the massive changes in communications occurring as a result of the exponential growth of computational technology in the last 20 years. The challenge of literary theory was also that while it provided us with a powerful set of analytical tools (e.g., Saussurean linguistics, semiotics, reception aesthetics, deconstructionism), it offered no guidelines as to how to use them to synthesize and construct computational media informed by that analysis. Beginning with such theorists as Wolfgang Iser and Roland Barthes, we found frameworks we would appropriate for rethinking AI's traditional conceptions of meaning, agency, comprehension, and language.

Crossing disciplinary boundaries: Learning how to talk with each other

Some of the greatest challenges we faced were in making explicit our implicit disciplinary assumptions and practices. This had far reaching and recurring implications for our attempt to read and discuss texts together. On the most basic level there were a host of terms and concepts that were unfamiliar to different members of the group. By reading the core and crossover texts in our disciplines we set out to understand each other's language. But the disciplinary differences were not only terminological: the standards and practices for what constituted acceptable talking, reading, writing, analysis, presentation, and production (texts and artifacts) were all quite different. Table 1 gives a somewhat caricatured, but useful breakdown of the differences.

We encountered these differences time and again, and learned to recognize, understand, and even engage in each other's different practices. The structure of the group helped in this process. The Narrative Intelligence Reading Group was student-initiated and student-run, so it had no curricular or departmental guidelines to adhere to. Each semester its members would meet to establish the

Table 1. Differences in practice between literary theorists and computer scientists

Task	Literary Theorist	Computer Scientist
Read Natural Language Texts	Close analysis of style, rhetorical structure, and implicit meanings to generate an interpretation.	Quick extraction of core concepts and utility of text to generate an instrumentalizable understanding. Style and other explicit rhetorical structures are devalued.
Read Programming Language Texts	No existing practice (this is an enormous oversight of the humanities)	Develop detailed understanding of how the text works so its parts can be appropriated for new text production
Write Natural Language Texts	This is the primary form of text production. Texts are most often analyses of other texts. This writing employs deliberate use of rhetorical devices to persuade the reader. Stylistic innovation is rewarded.	This is a secondary form of text production. Texts are most often analyses and documentation of programs. This writing employs unconscious use of rhetorical devices to persuade the reader. Stylistic innovation is discouraged.
Write Programming Language Texts	No existing practice (programs are texts which can generate texts – they are the realization of post-structuralism's dream of the autonomous self-replicating text, yet literary theorists do not study or write programs)	This is the primary form of text production. This form of writing creates machines, and, in some cases, machines which can make machines.
Discussing and Presenting Work	Speech as a rhetorical art form (the speech is itself an artifact). Innovation, complexity, elegance, and cleverness are valued. Complex texts are read from paper.	Speech as an instrumental means to facilitate the creation of artifacts. Simplicity, lack of ornament, and perspicacity are valued. Present simple slides and a demonstration of the work. Talk is improvised based on slides.

schedule. Within the loose framework of the group's goals, members would suggest texts to read that they would be willing to lead sessions on, then the group would collectively decide on the semester's syllabus and schedule. The group met once a week in the evening, for two hours or longer. Meetings usually had about 15 participants made up of a core group of roughly 12 people who went to most all meetings and a few people drawn from a pool of about 20 other members. Presenters would usually summarize the text, offer questions and frame issues, facilitate the discussion, and then close the discussion with a summary and look ahead to the next week. The schedule and syllabus were flexible and responsive to the needs of the group and the topics it covered. Based on the outcome of a session, we could elect to reorder the schedule, add new texts, or stay with a text for another session. The discussions were lively, multilevel, challenging, and compassionate. No question was too dumb (since most of us were novices outside our core fields), no answer too sacrosanct not to be challenged, questioned, analyzed, and rebuilt by the group. We made it up as we went along; we taught each other our fields and methods, and in the process fashioned a new discourse and practice of our own. It was both challenging and thrilling, and for many of us who participated in NI, it remains a kind of "golden age" of intellectual inquiry, colloquy, and invention.

In addition to the years of weekly meetings, we also went as a group to several events that helped strengthen the social and intellectual fabric of NI: Umberto Eco's talk "On the Quest for a Perfect Language" at Boston University hosted by Marvin Minsky in 1991; the Second International Conference on Cyberspace in Santa Cruz, CA, at which two of our members presented in 1992; talks at MIT by Evelyn Fox Keller, Camille Paglia, and Henry Jenkins; and the Tenth International Conference on Technology in Education in Cambridge, MA, at which some of our local and remote members appeared on a panel together in 1993. We also had some "guest stars" visit NI over the years, including Samuel Delaney (noted science fiction author and literary critic) and Tim Oren (software architect of Apple's Guides project).

As we developed a common discourse based on having read, critiqued, taken apart, and put back together our core texts and theories, we also were able to offer critique and support for our own NI-influenced work. We read each other's papers and offered feedback to each other's conference presentation rehearsals.

After several years of overcoming our disciplinary prejudices and habits, what did eventually emerge was a new type of interdisciplinary methodology for Narrative Intelligence. The primary breakthrough occurred in our developing ways to interleave and cross-pollinate theory (analysis of texts, people, and computational systems) with practice (creating new forms of computational media). By having read, discussed, and critiqued each other's core texts, we were able to develop a common discourse that supported a dialectic between the theoretical frameworks we inherited from artificial intelligence and literary theory and our practical experience of analyzing and building computational media systems.

Core texts and issues

In this partial bibliography, we list the core works that formed the center of our discourse, as well as a few other selections to indicate the diversity of interests in the group. We have divided them somewhat arbitrarily into categories, but in fact almost all of our readings crossed disciplinary boundaries.

Artificial intelligence and cognitive science

At the time we founded NI, mainstream artificial intelligence seemed bogged down in a view of mind based on mathematical logic and objective representation. Cognitive science developed theories of mind with similar assumptions and faced many of the same problems. Dissatisfied with this, we read some critiques from within the field, and identified for ourselves what we thought was useful. From traditional AI, the respective work of Marvin Minsky and Roger Schank was geared to less formal forms of knowledge, including narrative. Schank's group had a longstanding interest in story understanding and generation, and Minsky's Society of Mind theory had integrated some of these ideas into a computational framework.

The situated-action critique of AI (Phil Agre, David Chapman, Rodney Brooks, Gary Drescher, and others) was also influential. Drescher's work drew on constructivist roots that many of us shared. Agre's work was most informed by exposure to literary and social theory, and his paper "Writing & Representation" was one of the "founding documents" of the group. George Lakoff's work on metaphor and critique of objective representation was also influential in our thinking.

Agre, Philip (2001). Writing & representation. In M. Mateas & P. Sengers (Eds.), *Narrative Intelligence*. Amsterdam & Philadelphia: John Benjamins. Originally circulated in 1989 as an unpublished MIT AI Lab report.

Dennett, Daniel & Marcel Kinsbourne (1992). Time and the observer: The where and when of consciousness in the brain *Behavioral and brain sciences*, *15*(2), 183–200.

Drescher, Gary (1991). *Made up minds: A constructivist approach to artificial intelligence.* Cambridge, Massachusetts: MIT Press.

Lakoff, George & Mark Johnson (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Meehan, James (1981). TALE-SPIN. In R. C. Schank & C. K. Riesbeck, (Eds.), *Inside computer understanding: Five programs plus miniatures* (pp. 197–226). Hillsdale, New Jersey: Erlbaum.

Minsky, Marvin (1987). *The society of mind*. New York: Simon and Schuster.

Newell, Allen (1990). *Unified theories of cognition*. Cambridge, Massachusetts: Harvard University Press.

Ortony, Andrew, Gerald L. Clore, & Allan Collins (1998). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.

Schank, Roger C. (1973). Conceptualizations underlying natural language. In R. C. Schank & K. M. Colby, (Eds.), *Computer models of thought and language* (pp. 187–247). San Francisco: W. H. Freeman.

Literary theory

Literary theory provided an important framework for understanding language, communication, and cognition in ways critical of the conceptions underlying artificial intelligence. Roland Barthes' work on semiotics and the "death of the author," Wolfgang Iser's work on reception aesthetics, and Michael Reddy's "toolmakers paradigm" enabled us to go beyond the sender-receiver model of communication underlying most thinking in computer science to one in which meaning is an active and constructive process. Aristotle's rhetorical theory provided an important common toolset for analyzing the structure and style of the texts, artifacts, and theories. Frances Yates' work on memory palaces offered us ancient but highly relevant ways of organizing discourse and memory, and provided fertile metaphors for envisioning new types of computational media systems and interfaces.

Aristotle (1977). *The rhetoric and the poetics of Aristotle*. New York: Random House Modern Library.

Barthes, Roland. (1977). The death of the author. In *Image, music, text* (pp. 142–148). New York: Hill and Wang.

Barthes, Roland. (1977). From work to text. In *Image, music, text* (pp. 155–164). New York: Hill and Wang.

Iser, Wolfgang. (1974). The reading process: A phenomenological approach. In *The implied reader: Patterns of communication in prose fiction from Bunyan to Beckett* (pp. 274–294). Baltimore: The Johns Hopkins University Press.

Iser, Wolfgang. (1989). The play of the text. In *Prospecting: From reader response to literary anthropology* (pp. 249–261). Baltimore: The Johns Hopkins University Press.

Reddy, Michael J. (1993). The conduit metaphor: a case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought. Second edition* (pp. 164–251). Cambridge: Cambridge University Press.
Yates, Frances A. (1966). *The art of memory*. Chicago: University of Chicago Press.

Media studies

From literary theory we broadened our focus to examine scholarly studies of other media technologies, such as film and television. Walter Ong's work deals with the oldest communication technology, writing, and the sharp distinction it generates between oral and literate cultures, and was quite important for us in understanding new hybrid media like MUDs that combine elements of the oral and the written. David Bordwell's book introduces the narrative language developed in the relatively short history of film. Scott McCloud's work is a masterful comic book about comic art that uses techniques of sequential visual representation to explain sequential visual representation. Jenkins debunks the commonly-held image of the passive television viewer, while McLuhan's sweeping and prophetic work practically invented the information age now coming to pass.

Bordwell, David (1985). *Narration in the fiction film*. Madison: University of Wisconsin Press.
Jenkins, Henry (1992). *Textual poachers: Television fans & participatory culture*. New York: Routledge.
McLuhan, Marshall (1964). *Understanding media: The extensions of man*. New York: McGraw-Hill.
McCloud, Scott (1993). *Understanding comics: The invisible art*. Northampton, Massachusetts: Tundra.
Ong, Walter J. (1982). *Orality and literacy: The technologizing of the word*. London: Methuen.

Narrative in psychology and sociology

There have been threads of work in psychology and sociology that center around narrative. Bartlett's view of memory as imaginative reconstruction of events was helpful to us in trying to break away from more static views of mental representation. Nelson's work on children's monologues showed how this reconstruction could take the form of oral self-narratives. These and other works helped sharpen for us the central role of narrative in the construction of the individual and of society.

Applebee, Arthur (1989). *The child's concept of story*. Chicago: University of Chicago Press.
Bartlett, Frederic (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
Freud, Sigmund (1933/1965). Lecture XXXI: The dissection of the psychical personality. In Strachey, James (Trans., Ed.), *New Introductory Lectures on Psychoanalysis* (pp. 51–71). New York: Norton.
Nelson, Katherine (Ed.) (1989). *Narratives from the crib*. Cambridge, Massachusetts: University Press.
Sacks, Harvey (1972). On the analyzability of stories by children. In J. J. Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics: The ethnography of communication* (pp. 325–345). New York: Rinehart & Winston.

User interface theory

All of the participants from the Media Lab (and many of the outsiders) were actively engaged in the research and design of new media and user interfaces, and were applying what we learned in NI to our work. Naturally we were interested in other efforts to apply literary theory or related disciplines to UI design. Brenda Laurel, Abbe Don, and Tim Oren, who were also participants in the group, were the most notable authors in this area. The issue of agents and character-based metaphors was and still is a prominent issue in the UI community. Our focus was on the intimate relationship between character and narrative.

Don, Abbe (1990). Narrative and the interface. In B. Laurel (Ed.), *The art of human computer interface design* (pp. 383–391). Reading, Massachusetts: Addison-Wesley.
Laurel, Brenda (1990). Interface agents: metaphors with character. In B. Laurel (Ed.), *The art of human computer interface design*. Reading, Massachusetts: Addison-Wesley.
Laurel, Brenda (1991). *Computers as theatre*. Reading, Massachusetts: Addison-Wesley.
Oren, Tim, Gitta Salomon, Kristee Kreitman, & Abbe Don (1990). Guides: Characterizing the interface. In B. Laurel (Ed.), *The art of human computer interface design*. Reading, Massachusetts: Addison-Wesley.

Software

We spent some sessions examining software, looking at how narrative was handled, or simply trying to apply some of the intellectual tools we had to criticism and analysis. We tried to apply reception theory and other tools to the analysis of new media forms, including hypertext, The Visual Almanac, an early work from Apple, programs for children, and the then-new genre of screen-savers.

In addition, we presented and critiqued some systems developed by our participants. IDIC applied some of our knowledge about narratives to the task of assembling video sequences. The Programming with Characters project depicted the workings of a program using characters and narrative in order to make it more understandable to a user.

Apple Computer Multimedia Lab (1988). The Visual Almanac. San Francisco: Apple Computer.
Hickman, Craig (1991). Kid Pix 1.0. Novato, California: Broderbund Software.
Nelson, Theodor Holm (1982). Literary Machines. Sausalito, California: Mindful Press.
Rundgren, Todd (1990). Flowfazer ® Music For The Eye. Utopia Grokware.
Sack, Warren & Marc Davis (1994). IDIC: Assembling video sequences from story plans and content annotations. In *IEEE International Conference on Multimedia Computing and Systems* (pp. 30–36). Boston, Massachusetts: IEEE Computer Society Press.
Travers, Michael & Marc Davis (1993). Programming with characters. In *1993 International Workshop on Intelligent User Interfaces* (pp. 269–272). Orlando, Florida: ACM Press.

Social computing

One thread of our discourse led from narrative to character to the real-life presentation of self as modified through computational media. In such media, identities become fluid, the usually implicit rules of social interaction become explicit or otherwise changed, and narrative play takes on new forms.

Borning, Alan & Michael Travers (1991). Two approaches to informal interaction over computer and video networks. In *Proceedings of CHI '91* (pp. 13–19). New Orleans: ACM Press.
Bruckman, Amy (1992). Identity workshop: Emergent social and psychological phenomena in text-based Virtual Reality. Unpublished paper. Online at: http://www.cc.gatech.edu/~asb/old/papers-index-deco1.html#IW.
Grudin, Jonathan (1990). Groupware and cooperative work: Problems and prospects. In B. Laurel (Ed.), *The art of human-computer interface design* (pp. 171–185). Reading, Massachusetts: Addison-Wesley.
Stone, Allucquère Roseanne (1991). Will the real body please stand up? Boundary stories about virtual cultures. In M. Benedikt (Ed.), *Cyberspace: First steps* (pp. 81–118). Cambridge, Massachusetts: MIT Press.

Constructionism in science and learning

The NI group was steeped in the culture of Piagetian constructionism, mostly through the influence of Seymour Papert. Our ideas about narrative were all implicitly of a constructionist bent. Since we were already familiar with this viewpoint in the areas of education and cognition, we decided to broaden our horizon by looking at the constructionist theories of science, which were not so readily accepted at MIT.

Haraway, Donna J. (1991). A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In *Simians, cyborgs, and women: The reinvention of nature* (pp. 149–181). London: Free Association Books.

Keller, Evelyn Fox (1992). Secrets of God, nature, and life. In *Secrets of life, secrets of death: Essays on language, gender and science* (pp. 56–72). New York: Routledge.

Keller, Evelyn Fox (1990). Physics and the emergence of molecular biology: A history of cognitive and political synergy. *Journal of Historical Biology, 23*(3), 389–409.

Papert, Seymour (1980). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books.

Lessons learned and impact

For the members of the Narrative Intelligence Reading Group a new type of intellectual activity became possible: a mutually reinforcing theory and practice of analyzing, designing, and building computational media consciously informed by the humanistic disciplines of literary theory, media studies, psychology, sociology, and philosophy. We also learned that true interdisciplinary work takes perseverance and patience. After 4 years of running a weekly unofficial seminar in the basement of the Media Lab, we found we had begun to have an impact on the institution's curriculum. In 1994, about a third of the doctoral proseminar syllabus for incoming Ph.D. students included "NI" materials. This was largely due to recent MS graduates who were members of NI entering the Ph.D. program. We also found that a number of courses being offered at the Media Lab and around MIT started to reflect the interdisciplinary approach of NI. Over the years of the seminar we had two faculty members who were very active participants (Henry Jenkins and Edith Ackermann) and who advocated NI approaches at MIT. As we became established and well known, we had frequent visits from other faculty. The most telling sign that we had blazed an important intellectual and curricular trail was the appearance in our later years of students who already hybridized literary theory, media studies, and computer science. These were students who were trained both in semiotics and programming languages as undergraduates and expected media technology research to combine them.

The future

While we accomplished much in the 6 years of the Narrative Intelligence Reading Group at the MIT Media Lab, there is so much more to be done. Most humanities departments still look at computation as a mere instrumentality and not as a serious and relevant area of intellectual inquiry. Most computer science programs, and even media technology programs, do not offer courses in which literary and media theory are taught and applied. The challenges facing the humanities and computer science demand not merely an interdisciplinary dialogue, but a redrawing of disciplinary boundaries. Training students and practitioners at this historical moment – as our means of communication are being radically transformed – requires that we develop a theoretically informed praxis that combines the best of our humanistic and computational sciences. The work of the Narrative Intelligence Reading Group was an early step in the important process of redefining what it means to practice (and to teach) a hybridized discipline of computational media studies.

Acknowledgements

Thanks to our local members who blazed the trail in the early years: Edith Ackermann, James Berkeley, Kathy Biddick, Amy Bruckman, Janet Cahn, Anil Chakravarthy, Abbe Don, Michelle Evard, Michelle Fineblum, Lenny Foner, Larry Friedlander, Ricki Goldman-Seagal, Jennifer Gonzalez, Gregory Gargarian, Nira Granott, Joel Henderson, Ian Horswill, Gilberte Houbard, Alex Jacobson, Henry Jenkins, Michael Johnson, Yasmin Kafai, Greg Kimberly, Golan Levin, Henry Lieberman, Pattie Maes, Kevin McGee, Teri Mendelsohn, Margaret Minsky, Shawn O'Donnell, Martin Roberts, Alan Ruttenberg, Warren Sack, Josh Smith, Carol Sperry, Carol Strohecker, Jeremy Wertheimer, Alan Wexelblatt, and Uri Wilensky. Thanks also to our active remote members: Joe Bates, Brenda Laurel, Stuart Moulthrop, Tim Oren, Sandy Stone, and Terry Winograd. For any of the NIers that we have inadvertently left out, our apologies, and please let us know.

Part I

Human Narrative

Chapter 3

The narrative construction of reality

Jerome Bruner
New York University, NY

1. Surely since the Enlightenment, if not before, the study of mind has centered principally upon how man achieves a "true" knowledge of the world. Emphasis in this pursuit has varied, of course: empiricists have concentrated upon the mind's interplay with an external world of nature, hoping to find the key in the association of sensations and ideas, while rationalists have looked inward to the powers of mind itself for the principles of right reason. The objective, in either case, has been to discover how we achieve "reality," that is to say, how we get a reliable "fix" on the world, a world that is, as it were, assumed to be immutable and, as it were, "there to be observed."

This quest has, of course, had a profound effect on the development of psychology, and the empiricist and rationalist traditions have dominated our conceptions of how the mind grows and how it gets its grasp on the "real world." Indeed, at mid-century Gestalt theory represented the rationalist wing of this enterprise and American learning theory the empiricist. Both gave accounts of mental development as proceeding in some more or less linear and uniform fashion from an initial incompetence in the grasp of reality to a final competence, in one case attributing it to the working out of internal processes or mental organization, and in the other to some unspecified principle of reflection by which – whether through reinforcement, association, or conditioning – we came to respond to the world "as it is." There have always been dissidents who challenged these views, but conjectures about human mental development have been influenced far more by majoritarian rationalism and empiricism than by these dissident voices.

In more recent times, it was Piaget who became the spokesman for the classic rationalist tradition by arguing the universality of a series of invariant developmental stages, each with its own set of inherent logical operations that successively and inexorably led the child to construct a mental representation of the real world akin to that of the detached, dispassionate scientist. While he did not quite drive the empiricist learning theorists from the field (they have begun the take new life again through the formulation of "connectionist" computer-simulations of learning), his views surely dominated the three decades following the Second World War.

Now there is mounting criticism of his views. The growth of knowledge of "reality" or of the mental powers that enable this growth to occur, the critics argue, is neither unilinear, strictly derivational in a logical sense, nor is it, as it were, "across the board." Mastery of one task does not assure mastery of other tasks that, in a formal sense, are governed by the same principles. Knowledge and skill, rather, are domain specific and, consequently, uneven in their accretion. Principles and procedures learned in one domain do not automatically transfer to other domains. Such findings were not simply a "failure to confirm" Piaget or the rational premise generally (see Segal et al. (1985)). Rather, if the acquisition of knowledge and of mental powers is indeed domain specific and not automatically transferable, this surely implies that a domain, so called, is a set of principles and procedures, rather like a prosthetic device, that permits intelligence to be used in certain ways, but not in others. Each particular way of using intelligence develops an integrity of its own – a kind of knowledge-plus-skill-plus-tool integrity – that fits it to a particular range of applicability. It is a little "reality" of its own that is constituted by the principles and procedures that we use within it.

These domains, looked at in another way, constitute something like a culture's treasury of toolkits. Few people ever master the whole range of toolkits: we grow clever in certain spheres, and remain incompetent in others in which, as it were, we do not become "hitched" to the relevant toolkit. Indeed, one can go even further and argue, as some have, that such cultural toolkits (if I may so designate the principles and procedures involved in domain specific growth) may in fact have exerted selection pressures on the evolution of human capacities. It may be, for example, that the several forms of intelligence proposed by Howard Gardner (which he attempts to validate by the joint evidence of neuropathology, genius, and cultural specialization) may be outcomes of such evolutionary selection (Gardner 1983). The attraction of this view is, of course, that it links man and his knowledge-gaining and knowledge-using capabilities to the culture of which he and his ancestors were active members. But it brings deeply into question not only the universality of knowledge from one domain to another, but the universal translatability of knowledge from one culture to another. For in this dispensation, knowledge is never "point-of-viewless."

This is a view that is very compatible with another trend that has arisen in the analysis of human intelligence and of "reality construction." It is not a new view, but it has taken new life in a new guise. Originally introduced by Vygotsky, and championed by his widening circle of admirers, the new position is that cultural products, like language and other symbolic systems, mediate thought and place their stamp on our representations of reality (Shore 1996; Vygotsky 1978; Vygotsky 1962; Stigler et al. 1990). In its latest version, it takes the name, after Seeley-Brown and Collins, of *distributed intelligence* (Brown et al. 1988). An individual's working intelligence is never "solo." It cannot be understood without taking into account his or her reference books, notes, computer programs and data bases, or most important of all, the network of friends, colleagues, or mentors on whom one leans for help and advice. Your chance of winning a Nobel Prize increases immeasurably if you have worked in the laboratory of somebody who has already won one, not because of pull but because of access to the ideas and criticisms of those who know better.[1]

2. Once one takes such views as seriously as they deserve, there are some interesting and non-obvious consequences. The first is that there are probably a fair number of important domains supported by cultural toolkits and distributional networks. A second is that the domains are probably differentially integrated in different cultures, as anthropologists have been insisting for some years now (Gladwin 1979; Rosaldo 1989; Geertz 1983; Bruner 1990). And a third is that many domains are not organized by logical principles or associative connections, particularly those that have to do with man's knowledge of himself, his social world, his culture. Indeed, most of our knowledge about human knowledge-getting and reality-constructing is drawn from studies of how people come to know the natural or physical world rather than the human or symbolic world. For many historical reasons, including the practical power inherent in the use of logic, mathematics, and empirical science, we have concentrated upon the child's growth as "little scientist," "little logician," "little mathematician." These are typically Enlightenment-inspired studies. It is curious how little effort has gone into discovering how man comes to construct the social world and the things that transpire therein. Surely, such challenging fine works as E. E. Jones's magisterial *Interpersonal Perception* make it clear that we do not achieve our mastery of social reality by growing up as "little scientists," "little logicians," or "little mathematicians" (Jones 1990). So while we have learned a very great deal indeed about how we come eventually to construct and "explain" a world of nature in terms of causes, probabilities, space-time manifolds, etc., we know altogether too little about how we go about constructing and representing the rich and messy domain of human interaction.

It is with just this domain that I want now to concern myself. Like the domains of logical-scientific reality construction, it is well buttressed by principles and procedures. It has an available cultural toolkit or tradition on which its procedures are modelled, and its distributional reach is as wide and as active as gossip itself. Its form is so familiar and ubiquitous that it is likely to be overlooked, in much the same way as we suppose that the fish will be the last to discover water. As I have argued extensively elsewhere, we organize our experience and our memory of human happenings mainly in the form of narrative – stories, excuses, myths, reasons for doing and not doing, and so on. Narrative is a conventional form, transmitted culturally and constrained by each individual's level of mastery and by his conglomerate of prosthetic devices, colleagues and mentors. Unlike the constructions generated by logical and scientific procedures which can be weeded out by falsification, narrative constructions can only achieve "verisimilitude." Narratives, then, are a version of reality whose acceptability is governed by convention and "narrative necessity" rather than by empirical verification and logical requiredness, although ironically, we have no compunction about calling stories true or false (For a fuller, more discursive account of the nature and products of narrative thought, see (Bruner 1986, 1990) and my more recent (Bruner 2002). See also (Sarbin 1986)).

I propose now to sketch out ten features of narrative, rather in the spirit of constructing an armature upon which a more systematic account might be constructed. As with all accounts of forms of representation of the world, I shall have a great difficulty in distinguishing what may be called the narrative mode of *thought* from the forms of narrative *discourse*. As with all prosthetic devices, each enables and gives form to the other, just as the structure of language and the structure of thought eventually become inextricable. Eventually it becomes a vain enterprise to say which is the more basic – the mental processes or the discourse form that expresses it –for, just as our experience of the natural world tends to imitate the categories of familiar science, so our experience of human affairs comes to take the form of the narratives we use in telling about them.

Much of what I have to say will not be at all new to those who have been working in the vineyards of narratology or who have concerned themselves with critical studies of narrative forms. Indeed, the ancestry of many of the ideas that will concern me can be traced back directly to the debates that have been going on among literary theorists over the last decade or two. My comments are echoes of those debates now reverberating in the human sciences – not only in psychology, anthropology, and linguistics, but also in the philosophy of language. For once the "Cognitive Revolution" in the human sciences brought to the fore the issue of how "reality" is represented in the act of knowing, it became apparent that it did not suffice to equate representations with images, with propositions, with lexical networks, or even with such more temporally extended vehicles as sentences. It was perhaps a decade ago that psychologists became alive to the possibility of narrative as a form not only of representing but of constituting reality, a matter of which I shall have more to say presently. It was at that point that cognitively inclined psychologists and anthropologists began to discover that their colleagues in literary theory and historiography were deeply immersed in asking comparable questions about textually situated narrative. I think one can even date the "paradigm shift" to the appearance of a collection of essays narrative inquires in 1981 – *On Narrative* (Mitchell 1981).

If some of what I have to say about the features of narrative, then, seem "old hat" to the literary theorist, let him or her bear in mind that the object is different. The central concern is not how narrative as text is constructed, but rather how it operates as an instrument of mind in the construction of reality. And now to the ten features of narrative.

3.

1. Narrative diachronicity. A narrative is an account of events occurring over time. It is irreducibly durative. It may be characterizable in seemingly non-temporal terms (as a "tragedy" or a "farce") but such terms only summarize what are quintessentially patterns of events occurring over time. The time involved, moreover, as Ricoeur has noted, is "human time" rather than abstract or "clock" time (Ricoeur 1984). It is time whose significance is given by the meaning assigned to events within its compass. William Labov, one the greatest students of narrative, also regards temporal sequence as essential to narrative but he locates this temporality in the meaning-preserving sequence of clauses in narrative *discourse* itself (Labov 1967, 1981). While this is a useful aid to linguistic analysis, it nonetheless obscures an important aspect of narrative representation. For there are many conventionalized ways of expressing the sequenced durativity of narrative even in discourse, like flashbacks and flashforwards, temporal synechdoche, and so on. As Nelson Goodman warns, narrative comprises an ensemble of ways of constructing and representing the sequential, diachronic order of human events, of which the sequencing of clauses in spoken or written "stories" is only one device (Goodman 1981). Even non-verbal media have conventions of narrative diachronicity, as in the "left-to-right" and "up-to-down" conventions of cartoon strips and cathedral windows. What underlies all conventionalized forms for representing narrative is a "mental model" that has its unique pattern of events over time that gives it its defining property. And to that we shall come presently.

2. Particularity. Narratives take as their ostensive reference particular happenings. But this is, as it were, their vehicle rather than their destination. For stories plainly fall into more general types; they are about boy-woos-girl, bully-gets-his-comeuppance, etc. In this sense the particulars of narratives are tokens of broader types. Where the boy-woos-girl script calls for the giving of a gift, for example, the gift can equally well be flowers, perfume, or even an endless golden thread. Any of these may serve as an appropriate token or emblem of a gift. Particularity achieves its emblematic status by its embeddedness in a story that is in some sense generic. And, indeed, it is by virtue of this embeddedness in genre, to look ahead, that narrative particulars can be "filled in" when they are missing from an account. The "suggestiveness" of a story lies, then, in the emblematic nature of its particulars, its relevance to a more inclusive narrative type. But for all that, a narrative cannot be realized save through particular embodiment.

3. Intentional state entailment. Narratives are about people acting in a setting, and the happenings that befall them must be relevant to their intentional states while so engaged – to their beliefs, desires, theories, values, etc. When animals or non-agentive objects are cast as narrative protagonists, they must be endowed with intentional states for the purpose, like the Little Red Engine in the children's story. Physical events play a role in stories chiefly by affecting the intentional states of their protagonists. As Baudelaire put it, "The first business of an artist is to substitute man for nature."

But intentional states in narrative never fully determine the course of events, since a character with a particular intentional state might end up *doing* practically anything. For some measure of agency is always present in narrative, and agency presupposes choice – some element of "freedom." If people can predict anything from a character's intentional states, it is only how he will feel or how he will have perceived the situation. The loose link between intentional states and subsequent action is the reason why narrative accounts cannot provide causal explanations. What they supply instead is the basis for *interpreting* why a character acted as he or she did. Interpretation is concerned with "reasons" for things happening, rather than strictly with their "causes," a matter to which we turn next.

4. Hermeneutic composability. A preliminary word of explanation is needed here. The word *hermeneutic* implies that there is a text or a text analogue *through* which somebody has been trying to express a meaning and *from* which somebody is trying to extract a meaning. This in turn implies that there is a difference between what is *expressed* in the text and what the text might *mean*, and furthermore that there is no unique solution to the task of determining *the* meaning for *this* expression. Such hermeneutic interpretation is required when there is neither a *rational* method of assuring the "truth" of a meaning assigned to the text as a whole, nor an *empirical* method for determining the verifiability of the constituent elements that make up the text. In effect, the best hope of hermeneutic analysis is to provide an intuitively convincing account of the meaning of the text as a whole in the light of the constituent parts that make it up. This leads to the dilemma of the so-called "hermeneutic circle" – in which we try to justify the "rightness" of one reading of a text in terms of other readings rather than by, say, rational deduction or empirical proof. The most concrete way of explicating this dilemma or "circle" is by reference to the relations between the meanings assigned the whole of a text (say a story) and its constituent parts. As Charles Taylor puts it, "we are trying to establish a reading for the whole text, and for this we appeal to readings of its partial expressions; and yet because we are dealing with meaning, with making sense, where expressions only make sense or not in relation to others, the readings of partial expressions depend on those of others, and ultimately of the whole" (Taylor 1979: 28).

This is probably nowhere better illustrated than in narrative. The accounts of protagonists and events that constitute a narrative are selected and shaped in terms of a putative story or plot that then "contains" them. At the same time, the "whole" (the mentally represented putative story) is dependent for its formation upon a supply of constituent candidate parts. In this sense, as we have already noted, parts and wholes in a narrative rely upon each other for their viability (Ricoeur 1984). In Vladimir Propp's terms, the parts of a narrative serve as "functions" of the narrative structure as a whole (Propp 1968; Propp 1984). But that whole cannot be constructed without reference to such appropriate parts. This puzzling part-whole textual interdependence in narrative is, of course, an illustration of the defining property of what is called the "hermeneutic circle." For a story can only be "realized" when its parts and whole can, as it were, be made to live together.

This hermeneutic property marks narrative both in its construction and in its comprehension. For narratives do not exist, as it were, in some real world, waiting there patiently and eternally to be veridically mirrored in a text. The act of constructing a narrative, moreover, is considerably more than "selecting" events either from real life, from memory, or from fantasy and then placing them in an appropriate order. The events themselves need to be *constituted* in the light of the overall narrative – in Propp's terms, to be made "functions" of the story. This is a matter to which we will return later.

Now let me return to "hermeneutic composability." The telling of a story and its comprehension *as* a story depend upon the human capacity to process knowledge in this interpretive way. It is a way of processing that has, in the main, been grossly neglected by students of mind raised either in the rationalist or in the empiricist traditions. The former have been concerned with mind as an instrument of right reasoning, with the means we employ for establishing the necessary truth inherent in a set of connected propositions. Piaget was a striking example of this rational tradition. Empiricists, for their part, rested their claims upon a mind capable of verifying the constituent "atomic propositions" that comprised a text. But neither of these procedures, right reason or verification, suffice for explicating how a narrative is either put together by a speaker or interpreted by a hearer. This is the more surprising since there is compelling evidence to indicate that narrative comprehension is among the earliest powers of mind to appear in the young child and among the most widely used forms of organizing human experience (see, for example, Nelson (1989) and Bruner (1990)).

Many literary theorists and philosophers of mind have argued that the act of interpreting in this way is forced upon us only when a text of the world to which it presumes to refer is in some way "confused, incomplete, cloudy..." (Taylor 1985: 15). Doubtless we are more aware of our interpretive efforts when faced with textual or referential ambiguities. But I would take strong exception to the general claim that interpretation is forced upon us only by a surfeit of ambiguity. The illusion created by skilful narrative that this is not the case, that a story "is as it is" and needs no interpretation, is produced by two quite different processes. The first should probably be called "narrative seduction." Great story tellers have the artifices of narrative reality construction so well mastered that their telling preempts momentarily the possibility of any but a single interpretation – however bizarre it may be. The famous episode of a Martian invasion in the broadcast of Orson Welles's *War of the Worlds* provides a striking example (Cantril 1940). Its brilliant exploitation of the devices of text, context, and *mis en scene* predisposed its hearers to one and only one interpretation, however bizarre it seemed to them in retrospect. It created "narrative necessity," a matter we understand much less well than its logical counterpart, logical necessity. The other route to making a story seem self-evident and not in need of interpretation is via "narrative banalization." It is when we take a narrative as so socially conventional, so well-known, so in keeping with the canon, that we can assign it to some well-rehearsed and virtually automatic interpretive routine. These constitute what Roland Barthes called "readerly" texts in contrast to "writerly" ones that challenge the listener or reader into unrehearsed interpretive activity (Barthes 1985).

In a word, then, it is not textual or referential ambiguity that compels interpretive activity in narrative comprehension, but narrative itself. Narrative seduction or narrative banalization may produce restricted or routine interpretive activity, but this does not alter the point. "Readerly" story interpretation or "hack" story constructions can be altered by surprisingly little instruction (See, for example, (Elbow 1986)). And the moment a hearer is made suspicious of the "facts" of a story or the ulterior motives of a narrator, he or she immediately becomes hermeneutically alert. If I may use an outrageous metaphor, automatized interpretations of narratives are comparable to the "default settings" of a computer: an economical, time- and effort-saving way of dealing with knowledge – or, as it has been called, a form of "mindlessness" (Langer 1989).

Interpretation has a long history in biblical exegesis and in jurisprudence. It is studded with problems that will become more familiar shortly, problems that have to do more with context than text, with the conditions on telling rather than with what is told. Let me tag two of them better to identify them for subsequent discussion. The first is the issue of *intention*: "why" the story is told how and when it is, and interpreted as it is by interlocutors caught in different intentional stances themselves. Narratives are not, to use Roy Harris's felicitous phrase, "unsponsored texts" to be taken as existing unintentionally as if cast by fate upon a printed page (Harris 1989). Even when the reader takes them in the most "readerly" way, he usually attributes them (following convention) as emanating from an omniscient narrator. But this condition is itself not to be overlooked as uninteresting. It probably derives from a set of social conditions that give special status to the written word in a society where literacy is a minoritarian prerogative.

A second contextual issue is the question of *background knowledge* – of both the story teller and the listener, and how each interprets the background knowledge of the other. The philosopher Hilary Putnam, in a quite different context, proposes two principles: the first is a Principle of Benefit of Doubt, the second a Principle of Reasonable Ignorance: the first "forbids us to assume that . . . experts are factually omniscient," the second that "any speakers are philosophically omniscient (even unconsciously)" (Putnam 1975: 278). We judge their accounts accordingly. At the other extreme, we are charitable toward ignorance and forgive children and neophytes their incomplete knowledge, "filling in" for them as necessary. Or Sperber and Wilson, in their well known discussion of "relevance," argue that in dialogue we typically presuppose that what an interlocutor says in replying to us is topic-relevant and that we most often assign an interpretation to it accordingly in order to make it so, thereby easing our task in understanding Other Minds (Sperber & Wilson 1986). We also take for granted, indeed we institutionalize situations in which it is taken for granted that the "knowledge register" in which a story is told is different from the one in which it is taken up, as when the client tells the lawyer his story in "life talk" and is listened to in "law talk" so that the lawyer can advise about litigation (rather than life). The analyst and the analysand in therapy are comparable to the lawyer and client in legal consultation.[2]

Both these contextual domains, intention attribution and background knowledge, provide not only bases for interpretation but, of course, important grounds for negotiating how a story shall be taken – or, indeed, how it should be told, a matter better reserved for later.

5. Canonicality and breach. To begin with, not every sequence of events recounted constitutes a narrative, even when it is diachronic, particular, and organized around intentional states. Some happenings do not warrant telling about and accounts of them are said to be "pointless" rather than story-like. A Schank-Abelson script is one such case: it is a prescription for canonical behavior in a culturally defined situation – how to behave in a restaurant, say (Schank & Abelson 1977). Narratives require such scripts as necessary background, but they do not constitute narrativity itself. For to be worth telling, a tale must be about how an implicit canonical script has been breached, violated or deviated from in a manner to do violence to what Hayden White calls the "legitimacy" of the canonical script (White 1981). This usually involves what Labov calls a "precipitating event," a concept that Barbara Herrnstein-Smith puts to good use in her exploration of literary narrative (Labov 1967, 1981; Herrnstein-Smith 1978).

Breaches of the canonical, like the scripts breached, are often highly conventional and are strongly influenced by narrative traditions. Such breaches are readily recognizable as familiar human plights – the betrayed wife, the cuckolded husband, the fleeced innocent, etc. Again, they are conventional plights of "readerly" narratives. But both scripts and their breaches also provide rich grounds for innovation – as witness the contemporary literary-journalistic invention of the "yuppy" script or the formulation of the white-collar criminal's breach. And this is, perhaps, what makes the innovative story teller such a powerful figure in a culture. He may go beyond the conventional scripts, leading people to see human happenings in a fresh way, indeed, in a way they had never before "noticed" or even dreamed. The shift from Hesiod to Homer, the advent of "inner adventure" in Lawrence Sterne's *Tristram Shandy*, the advent of Flaubert's perspectivalism, or Joyce's epiphanizing of banalities – these are all innovations that probably shaped our narrative versions of everyday reality as well as changing the course of literary history, the two perhaps being not that different.

It is to William Labov's great credit to have recognized and provided a linguistic account of narrative structure in terms of two components – what happened and why it is worth telling (Labov 1967, 1981). It was for the first of these that he proposed his notion of irreducible clausal sequences. The second captures the element of breach in canonicality, and involves the use of what he calls *evaluation* for warranting a story's tellability as evidencing something unusual. From initial orientation to final coda, the language of evaluation is made to contrast with the language of clausal sequence – in tense, aspect, or other marking. It has even been remarked that in sign languages, the signing of sequence and of evaluation are done in different places in the course of telling a story, the former at the center of the body, the latter off to the side.

The "breach" component of a narrative can be created by linguistic means as well as by the use of a putatively delegitimizing precipitating event in the plot. Let me explain. The Russian Formalists distinguished between the "plot" of a narrative, its *fabula*, and its mode of telling, what they called its *sjuzet*. Just as there are linearization problems in converting a thought into a sentence, so there are problems in, so to speak, representing a *fabula* in its enabling *sjuzet* (for a discussion of uses of this distinction by the Russian Formalists, see (Bruner 1986)). The literary linguist, Tzvetvan Todorov, whose ideas we shall visit again later, argues that the function of inventive narrative is not so much to "fabulate" new plots as to render previously familiar ones uncertain or problematical, challenging a reader into fresh interpretive activity – echoing Roman Jakobson's famous definition of the writer's task, "to make the ordinary strange" (Todorov 1977; for a good statement of Roman Jakobson's view, see (Jakobson 1960)).

6. *Referentiality.* The acceptability of a narrative obviously cannot depend upon its correctly referring to reality, else there would be no fiction. Realism in fiction must then indeed be a literary convention rather than a matter of correct reference. Narrative "truth" is judged by its verisimilitude rather than its verifiability. There seems indeed to be some sense in which narrative, rather

than referring to "reality," may in fact create or constitute it, as when "fiction" creates a "world" of its own – Joyce's "Dublin" where places like St.Stephen's Green or Grafton Street, for all that they bear familiar labels, are no less real or imaginary than the characters he invents to inhabit them. In a perhaps deeper sense, indeed, it may be that the plights and the intentional states depicted in "successful" fiction sensitize us to experience our own lives in ways to match. Which suggests, of course, that the distinction between narrative fiction and narrative truth is nowhere nearly as obvious as common sense and usage would have us believe. *Why* comon sense insists practically upon such a sharp distinction being drawn is quite another problem, perhaps related to the requirement of "bearing witness." But that lies beyond the scope of this essay.

What does concern us, rather, is why the distinction is intrinsically difficult to make and sustain. Surely one reason lies in what I earlier called the hermeneutic composability of narrative itself. For such composability creates problems for the conventional distinction between "sense" and "reference." That is, the "sense" of a story as a whole may alter the reference and even the referentiality of its component parts. For a story's components, insofar as they become its "functions" or captives, lose their status as singular and definite referring expressions. St. Stephen's Green becomes, as it were, a type rather than a token, a class of locales including the locus so named in Dublin. It is an invented referent not entirely free of the meanings imparted by the real place, just as a story that requires a "betrayal" as one of its constituent functions, can convert an ordinarily mundane event recounted into something that seems compellingly like a betrayal. And this, of course, is what makes circumstantial evidence so deadly and so often inadmissible in courts of law. Given hermeneutic composability, referring expressions within narrative are always problematic, never free of the narrative as a whole. What is meant by the "narrative as a whole"? This leads us to the so-called "law of genres," to which we turn next.

7. Genericness. We all know that there are recognizable "kinds" of narrative: farce, black comedy, tragedy, the *Bildungsroman*, romance, satire, travel saga, etc. But as Alastair Fowler so nicely puts it, "genre is much less a pigeonhole than a pigeon" (Fowler 1982: 37). That is to say, we can speak of genre both as a property of a text or as a way of comprehending narrative. Mary McCarthy wrote short stories in several literary genres. She later gathered some of them together in an order of the increasing age of the chief female protagonist, added some interstitial "evaluation" sections, and published the lot as an autobiography entitled, *Memories of a Catholic Girlhood*. Thereafter (and doubtless to her dismay) readers interpreted her new stories as further installments of autobiography. Genres seem to provide both writer and reader with commodious and conventional "models" for limiting the hermeneutic task of making sense of human happenings – ones we narrate to ourselves as well as ones we hear others tell.

What are genres, viewed psychologically? Merely conventionalized representations of human plights? There are surely such plights in all human cultures: conflicts of family loyalty, the vagaries of human trust, the vicissitudes of romance, etc. And it might even seem that they are universal, given that the classics can be done in "modern dress" and the tales of exotic peoples be locally translated. But I think that emphasis upon plights and upon their putative universality may obscure a deeper issue. For plight is only the plot form of a genre, its *fabula*. But genre is also a form of telling, its *sjuzet*. Even if genres specialize in conventionalized human plights, they achieve their effects by using language in a particular way. And to translate the "way of telling" of a genre into another language or culture where it does not exist requires a fresh literary-linguistic invention (see Brower (1959)). It contains critical essays on the task of translating fiction and non-fiction into English by some of the great practitioners of the art.). The invention may, of course, be culturally out of reach. Language, after all, is contained within its uses. It is not just a syntax and a lexicon. The so-called "inward turn of narrative" in Western literature, for example, may have depended upon the rise of silent reading, which is a rather recent invention. If the reflectiveness produced by silent reading was then intensified by the creation of new genres – the so-called modern and post-modern novels – we might well expect that such genres would not be easily accessible to the Western non-reader and even less so to a member of a non-literate culture.

While genres, thus, may indeed be loose but conventional ways of representing human plights, they are also ways of telling that predispose us to use our minds and sensibilities in particular ways. In a word, while they may be representations of social ontology, they are also invitations to a particular style of epistemology. As such, they may have quite as powerful an influence in shaping our modes of thought as they have in creating the realities that their plots depict (Heath 1983; Ochs & Schieffelin 1983; Ochs et al. 1989; Feldman 1989). So, for example, we celebrate innovations in genre as changing not only the content of imagination but its modus operandi: Flaubert for introducing a perspectival relativism that dethroned both the omniscient narrator and the singular "true" story, Joyce for slyly substituting free association to break the constraints of semantic and even syntactic conventionalism, Beckett for shredding the narrative continuities we had come to take for granted in story telling, Calvino for converting post-modern anti-foundationalism into classic mythic forms, and so on.

Narrative genre, in this dispensation, can be thought of not only as a way of constructing human plights, but as providing a guide for using mind, insofar as the use of mind is guided by the use of an enabling language.

8. Normativeness. Because its "tellability" as a form of discourse rests upon a breach of conventional expectation, narrative is necessarily normative. A breach presupposes a norm. It is this founding condition of narrative that has led students of the subject, from Hayden White and Victor Turner to Paul Ricoeur, to propose that narrative is centrally concerned with cultural legitimacy (see especially (White 1978; Turner 1982)). A new generation of legal scholars, not surprisingly, has even begun to explore the implicit norms inherent in legal testimony, which, of course, is principally narrative in form (*Michigan Law Review 1989*, see also *Amsterdam and Bruner 2000).*

While everybody from Aristotle to the so-called narrative grammarians, all agree that a story pivots on a breach in legitimacy, the differences in how the notion of breach is conceived are themselves revealing of differing cultural emphases. Take Kenneth Burke's celebrated account of the dramatic "pentad." The pentad consists of an Agent, an Action, a Scene, a Goal, and an Instrument, the appropriate balance between these elements being defined as a "ratio" determined by cultural convention. When this "ratio" becomes unbalanced, when conventional expectation is breached, Trouble ensues. And it is Trouble that provides the engine of drama, Trouble as an imbalance between any and all of the five elements of the pentad: Nora in *A Doll's House*, for example, is a rebellious Agent in an inappropriately bourgeois Scene, etc. Precipitating events are, as it were, emblems of the imbalance. Burke's principal emphasis is on plight, *fabula*. It is, as it were, concerned ontologically with the cultural world and its arrangements, with norms as they "exist."

In the second half of our century, as the apparatus of skepticism comes to be applied not only to doubting the legitimacy of received social realities but also to questioning the very ways in which we come to know or construct reality, the normative program of narrative (both literary and popular) changes with it. "Trouble" becomes epistemic: Julian Barnes writes a stunning narrative on the *episteme* of Flaubert's perspectivalism, *Flaubert's Parrot*; or Italo Calvino produces a novel, *If on a Winter's Night a Traveler*, in which the issue is what is text and what context; and theories of poetics change accordingly. They too take an "epistemic turn." And so the linguist Tzvetvan Todorov sees the poetics of narrative as inhering in its very language, in a reliance on the use of linguistic transformations that render any and all accounts of human action more subjunctive, less certain, and subject withal to doubt about their construal. It is not simply that "text" becomes dominant but that the world to which it putatively refers is, as it were, the creature of the text (see, for example, (Suleiman & Crosman 1980)).

The normativeness of narrative, in a word, is not historically or culturally "once for all." Its form changes with the preoccupations of the age and the circumstances surrounding its production. Nor is it required of narrative, by the way, that the Trouble with which it deals be resolved. Narrative, I believe, is designed to contain uncanniness rather than to resolve it. It does not have to come out on the "right side." What Frank Kermode calls the "consolation of narrative" is not the comfort of a happy ending, but the comprehension of plight that, by being made understandable, becomes bearable (Kermode 1981).

9. Context sensitivity and negotiability. This is a topic whose complexities we have already visited in an earlier discussion of "hermeneutic composability" and the interpretability of narrative. In considering context, the familiar issues of narrative intention and of background knowledge arise again. With respect to the first of these, much of literary theory has abandoned Coleridge's dictum that the reader should suspend disbelief and stand, as it were, naked before the text. Today we have "reader response" theory and books entitled *The Reader in the Text* (Iser 1989; Suleiman & Crosman 1980). Indeed, the prevailing view is that the notion of totally suspending disbelief is at best an idealization of the reader and, at worst, a distortion of what the process of narrative comprehension involves. Inevitably, we assimilate narrative on our own terms, however much (in Wolfgang Iser's account) we treat the occasion of a narrative recital as a specialized speech act (Iser 1974). We inevitably take the teller's intentions into account, and do so in terms of our background knowledge (and, indeed, in the light of our presuppositions about the teller's background knowledge).

I have a strong hunch, which may at first seem counter-intuitive, that it is this very context sensitivity that makes narrative discourse in everyday life such a viable instrument for cultural negotiation. You tell your version, I tell mine, and we rarely need legal confrontation to settle the difference. Principles of charity and presumptions of relevance are balanced against principles of sufficient ignorance and sufficient doubt to a degree one would not expect where criteria of consistency and verification prevailed. We seem to be able to take competing story versions with a perspectival grain of salt, much more so than in the case of arguments or proofs. Judy Dunn's remarkable book on the beginning of social understanding in children makes it plain that this type of negotiation of different narrative versions starts early and is deeply imbedded in such practical social actions as the offering of excuses, not merely in story telling per se (Dunn 1988). I think it is precisely this interplay of perspectives in arriving at "narrative truth" that has led philosophers like Richard Rorty to abandon univocally verificationist views of truth in favor of pragmatic ones ((Rorty 1979); see also (Taylor 1989)). Nor is it surprising that anthropologists have increasingly turned away from positivist descriptions of cultures toward an interpretive one in which not objective categories but "meanings" are sought for, not meanings imposed *ex hypothesi* by an outsider, the anthropologist, but ones arrived at by indigenous participants immersed in the culture's own processes for negotiating meaning (see particularly (Geertz 1983); see also (Rabinow & Sullivan 1979) and (Stigler et al. 1990)).

On this view, it is the very context dependence of narrative accounts that permits cultural negotiation which, when successful, makes possible such coherence and interdependence as a culture can achieve.

10. Narrative accrual. How do we cobble stories together to make them into a whole of some sort? Sciences achieve their accrual by derivation from general principles, by relating particular findings to central paradigms, by couching empirical findings in a form that makes them subsumable under altering paradigms, and by countless other procedures for making science, as the saying goes, "cumulative." This is vastly aided, of course, by procedures for assuring verification though, as we know, verificationist criteria have limited applicability where human intentional states are concerned, which leaves psychology rather on the fringe.

Narrative accrual is not foundational in the scientist's sense. Yet narratives do accrue and, as anthropologists insist, the accruals eventually create something variously called a "culture" or a "history" or, more loosely, a "tradition." Even our own homely accounts of happenings in our own lives are eventually converted into more or less coherent autobiographies centered around a Self acting more or less purposefully in a social world (see, for example, Chapter 4 in (Bruner 1990)). Families similarly create a corpus of connected and shared tales and Elinor Ochs's studies in progress on family dinner-table talk begin to shed light on how this is accomplished.[3] Institutions too, as we know from the innovative work of Eric Hobsbawm, "invent" traditions out of previously ordinary happenings and then endow them with privileged status (Hobsbawm & Ranger 1983). And there are principles of jurisprudence, like *stare decisis*, that guarantee a tradition by assuring that once a "case" has been interpreted in one way, future cases that are "similar" shall be interpreted and decided equivalently. Insofar as the law insists upon such accrual of cases as "precedents," and insofar as "cases" are narratives, the legal system imposes an orderly process of narrative accrual.

There has been surprisingly little work done on this fascinating subject, although there are stirrings among anthropologists (influenced principally by Clifford Geertz) and among historiographers (prodded by Michel Foucault's ground breaking *The Archeology of Knowledge*) (Geertz 1988; Clifford 1988; Foucault 1972). What kinds of strategies might guide the accrual of narratives into larger scale cultures or traditions or "world versions"? Surely one of them must be through the imposition of bogus *historical-causal entailment*: e.g., the assassination of Archduke Ferdinand is seen as "causing" the outbreak of the First World War, or Pope Leo III's coronation of Charlemagne as Holy Roman Emperor on Christmas Day in 800 is offered as "a first step on the way toward" or as a precursor of the enactment of the European Community in 1992. There is a vast literature of caution against such simplicities by both philosophers and historians, but it has not in the least diminished our passion for converting *post hoc* into *propter hoc*.

Another strategy might be called, for lack of a better expression, *coherence by contemporaneity*: the belief that things happening at the same time must be connected. I made the wry discovery, writing my own intellectual autobiography several years ago, that once I had discovered in the *New York Times* Index what else had been happening at the time of some personal event, I could scarcely resist connecting the lot into one coherent whole – connecting, not subsuming, not creating historical-causal entailments, but winding it into story. My first scientific paper (on maturing sexual receptivity in the female rat), for example, was published about the time Neville Chamberlain had been duped by Hitler at Munich. My original story before consulting the *Times* Index was vaguely about a nineteen-year-old's first discovery, rather like a Bildungsroman. The post-Index story, with Munich now included, was an exercise in irony: young Nero fiddling with rats while Rome burned! And by the same compelling process, we invent the Dark Ages, making everything all of a piece until, finally, the diversity becomes too great and then we invent the Renaissance.

Once shared culturally – distributed in the sense discussed earlier – narrative accruals achieve, like Emile Durkheim's collective representation, "exteriority" and the power of constraint (Durkheim 1965).[4] The Dark Ages come to exist, and we come to cluck with wonder at the "exceptionality" of any non-traditional philosopher or deviant theologian who lived in its shadows. I am told that the ex-President and Nancy Reagan sent a letter of sympathy to a nationally known soap opera character who had just gone blind – not the actor, but the character. But that is not unusual: culture always reconstitutes itself by swallowing its own narrative tail – Dutch boys with fingers in the dike, Columbus Christianizing Indians, the Queen's honors list, the Europhilia that dates from Charlemagne.

What creates a culture, surely, must be a "local" capacity for accruing stories of happenings of the past into some sort of diachronic structure that permits a continuity into the present – in short, to construct a history, a tradition, a legal system, instruments assuring historical continuity if not legitimacy. I want to end my list of narrative properties on this rather "obvious" point for a particular reason. The perpetual construction and reconstruction of the past provide precisely the forms of canonicality that permit us to recognize when a breach has occurred and how it might be interpreted. The philosopher, W. T. Stace, proposed two philosophical generations ago that the only recourse we have against solipsism (the unassailable view that argues that we cannot prove the existence of a real world, since all we can know is our own experience) is that human minds are alike and, more important, that they "work in common" (Edwards 1967). One of the principal ways in which we work "mentally" in common, I would want to argue, is by the process of joint narrative accrual. Even our individual autobiographies, as I have argued elsewhere, depend upon being placed within a continuity provided by a constructed and shared social history in which we locate our Selves and our individual continuities (Bruner 1990: Chapter 4). It is a sense of belonging to this canonical past that permits us to form our own narratives of deviation while maintaining complicity with the canon. Perhaps Stace was too concerned with metaphysics when he invoked this process as a defense against solipsism. We would more likely say today that it must surely be a major prophylactic against alienation.

4. Let me return now to the original premise – that there are specific domains of human knowledge and skill and that they are supported and organized by cultural tool kits. If we accept this view, a first conclusion would be that in understanding the nature and growth of mind in any setting, we cannot take as our unit of analysis the isolated individual operating "inside her own skin" in a cultural vacuum. Rather, we must accept the view that the human mind cannot express its nascent powers without the enablement of the symbolic systems of culture. While many of these systems are relatively autonomous in a given culture – the skills of shamanism, of specialized trades, and the like – some relate to domains of skill that must be shared by virtually all members of a culture if the culture is to be effective. The division of labor within a society goes only so far. Everybody within a culture must in some measure, for example, be able to enter into the exchange of the linguistic community, even granted that this community may be divided by idiolects and registers. Another domain that must be widely (though roughly) shared for a culture to operate with requisite effectiveness is the domain of social beliefs and procedures – what we think people are like and how they must get on with each other, what elsewhere I have called "folk psychology" and what Harold Garfinkel has called ethnosociology (Garfinkel 1967). These are domains that are, in the main, organized narratively.

What I have tried to do in this paper is to describe some of the properties of a world of "reality" constructed according to narrative principles. In doing so, I have gone back and forth between describing narrative mental "powers" and the symbolic systems of narrative discourse that make the expression of these powers possible. It is only a beginning. My objective has been merely to lay out the ground plan of narrative realities. The daunting task that remains now is to show in detail how, in particular instances, narrative organizes the structure of human experience – how, in a word, "life" comes to imitate "art" and vice versa.

Notes

1. Zuckerman, Harriet, personal communication. She can be reached for further information at the Department of Sociology, Columbia University, New York.

2. See Spence (1982). An unwillingness on the part of a patient to accept the psychoanalyst's version or interpretation of a narrative is likely to lead to an examination and reformulation by the latter of the former's story as having to do with the patient's "resistance." The patient's version is made to conform to the psychiatrist's version as a price for the therapy's continuation. While lawyers, typically, in translating the client's personal "story" into a legal narrative, offer the client options in how the "facts of case" shall be legally framed – whether things "add up" to a narrative about contracts, torts, or rights to due process, say – the final legal story is, nonetheless, forced into a "canonical" narrative that conforms to prevailing biases in the society while also corresponding to some precedent in the law. So, for example, in recent American jurisprudence, the "facts of the case" of *Bowers vs. Hardwick* is interpreted as a violation of sodomy statutes of the State of Georgia rather than as an instance of the exercise of the individual's rights to privacy as guaranteed by the Fourth Amendment to the United States Constitution. The "fact" that a homosexual act is, in this case, between consulting adults is thereby ruled by the Court as "irrelevant" to the legal story. For a discussion of the effects of imposing "official" jurisprudential story forms on everyday narratives, see (Lane Scheppele 1989).

3. I am greatly indebted to Professor Ochs for letting a group of us in an informal seminar at UCLA during the Winter Term of l990 view her tapes of these sessions and share her views on the processes involved.

4. For a more psychological account of this process, referred to by the author as "ontic dumping," see (Feldman 1987).

References

Amsterdam, Anthony & Bruner, Jerome (2000). *Minding the Low*. Cambridge: Harvard University Press.

Barthes, Roland (1985). *The responsibility of forms: Critical essays on music, art, and representation*. New York: Hill and Wang.

Brower, Reuben (Ed.) (1959). *On translation*. Cambridge, MA: Harvard University Press.

Bruner, Jerome (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Bruner, Jerome (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Bruner, Jerome (2002). *Making Stories: Law, Literature, Life*. New York: Farrar, Straus & Giroux.

Cantril, Hadley (1940). *The invasion from Mars*. Princeton: Princeton University Press.

Clifford, James (1988). *The predicament of culture*. Cambridge, MA: Harvard University Press.

Dunn, Judy (1988). *Beginnings of social understanding*. Cambridge, MA: Harvard University Press.

Durkheim, Emile (1965). *The elementary forms of the religious life*. New York: Free Press.

Edwards, Paul (Ed.) (1967). Entry for "W. T. Stace." *Encyclopedia of philosophy*. New York: Macmillan and Free Press.

Elbow, Peter (1986). *Embracing contraries: Explorations in learning and teaching*. New York: Oxford University Press.

Feldman, Carol (1987). Thought from language: The linguistic construction of cognitive representations. In Jerome Bruner & Helen Haste (Eds.), *Making sense: The child's construction of the world*. New York: Methuen.

Feldman, Carol (1989). Monologue as problem-solving narrative. In K. Nelson (Ed.), *Narratives from the Crib*. Cambridge, MA: Harvard University Press.

Foucault, Michel (1972). *The archeology of knowledge*. New York: Pantheon.

Fowler, Alastair (1982). *Kinds of literature*, Cambridge, MA: Harvard University Press.

Gardner, Howard (1983). *Frames of mind*. New York: Basic Books.

Garfinkel, Harold (1967). *Studies in ethnomethodology*. Englewood NJ: Prentice Hall.

Geertz, Clifford (1983). *Local knowledge*. New York: Basic Books.

Geertz, Clifford (1983). Thick interpretation. In *Local knowledge*. New York: Basic Books.

Geertz, Clifford (1988). *Works and lives: The anthropologist as author*. Stanford: Stanford University Press.

Gladwin, Thomas (1970). *East is a big bird*. Cambridge, MA: Harvard University Press.

Goodman, Nelson (1981). Twisted tales: Or story, study, or symphony. In W. J. Thomas Mitchell (Ed.), *On narrative* (pp. 99–115). Chicago: University of Chicago Press.

Harris, Roy (1989). How does writing restructure thought. *Language and communication, 9*, 99–106.

Heath, Shirley Brice (1983). *Ways with words*. Cambridge: Cambridge University Press.

Herrnstein-Smith, Barbara (1978). *On the margins of discourse: The relation of literature to language*. Chicago: University of Chicago Press.

Hobsbawm, Eric & Terrence Ranger (Eds.) (1983). *The invention of tradition*. Cambridge: Cambridge University Press.

Iser, Wolfgang (1974). *The implied reader*. Baltimore: Johns Hopkins University Press.

Iser, Wolfgang (1989). *Prospecting: From reader response to literary anthropology*. Baltimore: Johns Hopkins University Press.

Jakobson, Roman (1960). Linguistics and poetics. In T. Sebeok (Ed.), *Style in language* (pp. 350–377). Cambridge, MA: MIT Press.

Jones, E. E. (1990). *Interpersonal perception*. San Francisco: Freeman.

Kermode, Frank (1981). Secrets and narrative sequence. In W. J. Thomas Mitchell (Ed.), *On Narrative* (pp. 79–99). Chicago: University of Chicago Press.

Labov, William & Joshua Waletzky (1967). Narrative analysis. In *Essays on the verbal and visual arts* (pp. 12–44). Seattle: University of Washington Press.

Labov, William (1981). Speech actions and reactions in personal narrative. *Georgetown University roundtable on languages and linguistics*, 219–247.

Lane Scheppele, Kim (1989). Foreword to a special issue on "Legal storytelling." *Michigan law review*, *87* (8), 2073–2098.

Langer, Ellen (1989). *Mindfulness*. Reading MA: Addison Wesley.

Michigan law review (1989). Special issue on "Legal storytelling." *87* (8).

Mitchell, W. J. Thomas (Ed.) (1981). *On narrative*. Chicago: University of Chicago Press.

Nelson, Katherine (Ed.) (1989). *Narratives from the crib*. Cambridge, MA: Harvard University Press.

Ochs, Elinor & Bambi Schieffelin (1983). *Acquiring conversational competence*. London: Routledge.

Ochs, Elinor, Carolyn Taylor, Dina Rudolph, & Ruth Smith (1989). *Narrative activity as a medium for theory-building*. Los Angeles, Department of Linguistics, University of Southern California.

Propp, Vladimir (1968). *Morphology of the folktale*. Austin: University of Texas Press.

Propp, Vladimir (1984). *Theory and history of folklore*. Minneapolis: University of Minnesota Press.

Putnam, Hilary (1975). *Mind, language, and reality*. Cambridge: Cambridge University Press.

Rabinow, Paul & William Sullivan (1979). *Interpretive social science: A reader*. Berkeley: University of California Press.

Ricoeur, Paul (1984). *Time and narrative*, Vol. I, Chicago: University of Chicago Press.

Rorty, Richard (1979). *Philosophy and the mirror of nature*. Princeton: Princeton University Press.

Rosaldo, Renato. (1989). *Culture and truth: The remaking of social analysis*. Boston: Beacon Press.

Sarbin, Theodore. (1986). *Narrative psychology: The storied nature of human conduct*. New York: Praeger.
Schank, Roger & Robert Abelson (1977). *Scripts, plans, goals, and understanding*. Hillsdale NJ: Erlbaum.
Seeley Brown, John, Allan Collins, & Paul Duguid (1988). Situated cognition and the culture of learning, *Educational researcher*, *18* (1), 32–42.
Segal, Judith, Susan Chipman, & Robert Glaser (1985). *Thinking and learning skills*. Hillsdale NJ: Erlbaum.
Shore, Bradd (1996). *Culture in Mind: cognition, culture, and the problem of meaning*, New York: Oxford University Press.
Spence, Donald (1982). *Narrative truth and historical truth*. New York: Norton.
Sperber, Dan & Dierdre Wilson (1986). *Relevance: Cognition and communication*. Cambridge, MA: Harvard University Press.
Stigler, James, Richard A. Shweder, & Gilbert Herdt (Eds.). (1990). *Cultural psychology*. Chicago: University of Chicago Press.
Suleiman, Susan & Inge Crosman (Eds.) (1980). *The reader in the text: essays on audience and interpretation*. Princeton: Princeton University Press.
Taylor, Charles (1979). Interpretation and the sciences of man. In Paul Rabinow and William M. Sullivan (Eds.), *Interpretive social science: A reader* (pp. 25–71). Berkeley: University of California Press.
Taylor, Charles (1985). Interpretation and the sciences of man. Chapt. 1, *Philosophy and the Human Sciences* (pp. 15–57). Cambridge: Cambridge University Press.
Taylor, Charles (1989). *Sources of the self*. Cambridge, MA: Harvard University Press.
Turner, Victor (1982). *From ritual to theater: The human seriousness of play*. New York: Performing Arts Journal Publications.
Todorov, Tzvetan (1977). *The poetics of prose*. Ithaca: Cornell University Press.
Vygotsky, Lev (1962). *Thought and language*. Cambridge: MIT Press.
Vygotsky, Lev (1978). *Mind in society*. Cambridge: Harvard University Press.
White, Hayden (1981). The value of narrativity in the representation of reality. In W. J. Thomas Mitchell (Ed.), *On narrative* (pp. 1–24). Chicago: University of Chicago Press.
White, Hayden (1978). *Tropics of discourse: Essays in cultural criticism*. Baltimore: Johns Hopkins University Press.

Chapter 4

Stories of lemurs and robots

The social origin of story-telling

Kerstin Dautenhahn
University of Hertfordshire, Hatfield

Introduction

This chapter[1] discusses narrative intelligence in the context of the origins of primate (social) intelligence. The relationship between social intelligence and narrative intelligence is outlined, with a particular emphasis on 1) the phylogenetic origins of primate (narrative) intelligence, and 2) the ontogenetic origin of autobiographical stories. The chapter is based on the assumption that in order to fully understand the importance and role of narrative in human intelligence one needs to draw attention to 'where stories come from', i.e. addressing whether story-telling can be linked to communication mechanisms that are evolutionary older but served a similar function, under which conditions and constraints story-telling capacities might have evolved, to what extent narrative intelligence is linked to social intelligence, etc, see Read and Miller (1995). This chapter will address some of these questions on the *origin of narrative* in *primates* and hopes to complement research that focuses on the particular structure and role of narrative in *humans*, e.g. (Turner 1996). Since the ontogenetic (developmental) aspects of story-telling are discussed in more detail elsewhere, e.g. in (Engel 1995/1999; Nelson 1986), this chapter focuses on selected research in primatology and autobiographical memory. The *Narrative Intelligence Hypothesis* (NIH) is explained, according to which the evolutionary origin of stories and narrativity was correlated with increasing social dynamics in primate societies, in particular the need to communicate about third-party relationships. Human narrative intelligence might have evolved because the structure of narrative is particularly suited to communicate about the social world, although in present human societies narrative and social-telling is used in a variety of contexts where social matters and communication might not necessarily be central (cf. narrative in arts, advertisement, entertainment etc.). After an introduction and discussion of the NIH, the possible implications of narrative intelligence research for understanding autism, and autism therapy are discussed. It is argued that *narrative technology* can potentially meet the social and cognitive needs of young primate story-tellers. The chapter concludes by outlining requirements for artificial *socially intelligent story-tellers*.

Primate intelligence: Getting to know each other

Primate societies belong to *individualized societies*. Here we find complex recognition mechanisms of kin and group members. This gives rise to complex kinds of social interaction and the development of various forms of social relationships and networks. On the behavioral level, long-lasting social bonding, attachment, alliances, dynamic (not genetically determined) hierarchies, social learning, development of traditions, etc., are visible signs of individualized societies. In humans, the evolution of language, culture and an elaborate cognitive system of mindreading and empathy are characteristics of human social intelligence in individualized societies (Dautenhahn 1997). As a consequence of the latter, humans not only pay attention to other agents and their interactions individually (interactions between distinct personalities), but also use their mental capacities to reason about other agents and social interactions.

In primate societies an individual is not only socially situated (being part of and surrounded by a social environment) but also socially embedded (Edmonds & Dautenhahn 1998), which means that the agent needs to pay attention to other agents and their interactions individually. Particularly, human primates are specialized in predicting, manipulating and dealing with highly complex social dynamics (involving direct relationships as well as third-party relationships); as we discuss below in more detail, they possess language as an effective means of preserving group coherence, *social bonding* (Dunbar 1993), and communicate about themselves and others in terms of stories. Humans not only deal with very complex relationships but seem to have *mental models* of themselves, others and the social world (cf. Baron-Cohen et al. (1985), Whiten (1991), Baron-Cohen (1995)). Humans live in individualized societies (as do some other species of birds and mammals). An increasingly complex social field and an increasing need to communicate effectively with each other were likely to have been among the important constraints in the evolution of human minds.

The use of the term 'minds' in this chapter is based on research into *theory of mind* and *mindreading* where people discuss whether and to what extent humans or other animals are able to reflect on their own mental states (e.g. desires, intentions, beliefs) and those of others. Researchers have studied whether human intelligence is particularly specialized in mindreading (Premack & Woodruff 1978; Povinelli & Preuss 1995). Minds are certainly attributed to members of *Homo sapiens* (and, as some evidence suggests, several other hominid species might have existed with *minds)*, but other candidates exist among mammals (e.g. non-human apes, dolphins, elephants) and birds (e.g. parrots and members of the crow family). Interestingly, species which we describe as possessing a 'mind' are all highly social. Even the solitary life style of *Pongo pygmaeus* or orangutans, (who nevertheless seem to be highly social in their ability to recognize and interact with each other) is rather a secondary adaptation to a particular environment which demands a spatially distributed social organization.

The *Social Intelligence Hypothesis* (SIH), sometimes also called *Machiavellian Intelligence Hypothesis* or *Social Brain Hypothesis*, suggests that the primate brain and primate intelligence evolved in adaptation to the need to operate in large groups where structure and cohesion of the group required a detailed understanding of group members. For important contributions to the SIH see, e.g., Chance and Mead (1953); Jolly (1966); Humphrey (1976/1988); Brothers (1990); and chapters in Byrne and Whiten (1988); Whiten and Byrne (1997). This hypothesis does not exclude possibly important ecological variables that might have provided initial demands and might have supported primate evolution. However, it is assumed that social complexity that demanded the evolution of social skills (which allow the interpretation, prediction, and manipulation of conspecifics) has been a prominent selective factor accelerating primate brain evolution, given that maintaining a large brain is very costly. To give an example, an adult human brain weighs about 2 % of the total body, but consumes 20 % of total energy intake, (Aiello & Wheeler 1995). Identifying friends and allies, predicting others' behavior, knowing how to form alliances, manipulating group members, making war, love and peace, are important ingredients of primate politics (de Waal 1982). In contrast to strepsirhine primates (lemurs and lorises), monkeys and apes show a variety of sophisticated social behavior: using alliances and cooperation in competition for resources, relying on support by others when acquiring dominance ranks, putting considerable effort into building, reconciliation and maintenance of long-lasting and intensive inter-personal social relationships, knowing personal characteristics and affiliations of group members and using techniques of social manipulation (Byrne

1999). Note that complex social behavior is also shown by other social non-primate mammalian species. Thus, there are two interesting aspects to human sociality: it served as an evolutionary constraint which led to an increase of brain size in primates, which in turn led to an increased capacity to further develop social complexity.

Dunbar and his collaborators found evidence (cf. Dunbar (1992); Dunbar (1993); Barton & Dunbar (1997); Dunbar (1998)) that the size of a cohesive social group in primates is a function of relative neocortical volume (volume of neocortex divided by rest of the brain). Such a correlation has not been found for ecological variables (hypothesizing e.g. that dietary considerations or the size of home ranges caused an increase in brain size). It is therefore suggested that social complexity played a causal role in primate brain evolution, namely that in order to manage larger groups, bigger brains are needed to provide the required 'information processing capacity'. The neocortex, which accounts for 50-80 % of total brain volume in primates, is generally associated with cognitive processes such as reasoning, mental manipulations and consciousness. Compared with more primitive parts of the brain, the neocortex size substantially increases from insectivores to prosimians, anthropoids, and humans. Indeed, it has been shown that primate species with relatively larger neocortices exhibit more complex social strategies than species with smaller neocortices (Pawlowski et al. 1998). It is also suggested that the relationship between encephalization (relationship between brain size and body size) and social complexity is not unique to primates. For example, findings reported in (Marino 1996) suggest that *cetaceans* (whales, dolphins, and porpoises) and primates show similar relationships between relative brain size and group size (a measure of social complexity). Similarly, Dunbar and Bever (1998) show that neocortex size predicts group size in carnivores and some insectivores. Also, bats that have stable social groups have a larger neocortex than bats that do not live in stable social groups (Barton & Dunbar 1997).

According to the SIH, primates are good primatologists, namely they are experts on social matters in a *laser-beam* form of intelligence. According to the SIH, during the evolution of human intelligence a transfer took place from social to non-social intelligence[2] so that hominid primates could transfer their expertise from the social to the non-social domain. An interesting aspect of this kind of transfer is discussed in (Mithen 1996). He explains the evolution of anthropomorphic thinking with an accessibility between the domains of social intelligence and natural history intelligence so that "people could be thought of as animals, and animals could be thought of as people" (Mithen 1996: 224). Furthermore, the accessibility between the domains of social and technical intelligence led to the possibility to think about people in terms of objects to be manipulated, in a similar way as physical objects can be manipulated. Although it is still unknown why hominids needed or chose to live in social groups, this *feedback principle* soon led to the development of highly sophisticated levels of organization and control in human societies (cf. Russell (1993)).

In non-human primate societies cohesion is maintained through time by *social grooming*. Social grooming patterns generally reflect social relationships:

> "The vervets clearly differentiated between the animals they groomed regularly and those they didn't. A grooming partner is something special, someone who deserves particular attention, who should be supported in moments of need, on whose behalf the taking of risks is warranted." (Dunbar 1996: 22)

Given the neocortical size of modern humans, we can extrapolate from the non-human primate regression and predict a group size of 150 for human societies. This number limits the number of relationships that an individual human can monitor simultaneously, it is the upper group size limit which still allows social contacts that can be regularly maintained, supporting effective coordination of tasks and information-flow via direct person-to-person contacts. Such relationships are *personal relationships*, they have sufficient depth to be relied on, they provide the basis of mutual support and coalition formation that are necessary in cases of attack or the need to access resources. The number 150 is supported by evidence from analyzing contemporary and historical human societies. Dunbar suggests that 1) there is a cognitive limit to the number of individuals with whom any one person can maintain stable relationships (depending on personal knowledge, face-to-face interactions), 2) that this limit (which he terms *cognitive group size*) is a direct function of relative neocortex size, and 3) that this in turn limits group size. But how do humans preserve cohesion in groups of 150 individuals, a function that (physical) social grooming serves in non-human primate societies? In terms of survival needs (resting, feeding etc.) primates can only afford to spend around 20 % of their time on social interactions and social grooming. However, a group size of 150 predicted for humans would require that about 42 % of the total time budget of a human primate are devoted to social grooming. It was therefore suggested by Dunbar (1993) that in order to preserve stability and coherence in human societies, human *language* has evolved as an efficient mechanism of *social bonding*, replacing social grooming mechanisms in non-human primate societies with direct physical contact (allowing only much smaller groups). Following this argument, language allowed an increase in group size while still preserving stability and cohesion within the group.

In the context of the evolution of human intelligence, Richard Byrne pointed out (Byrne 1997) that the Social Intelligence Hypothesis might account for the evolution of primate intelligence, but offers little explanation for the evolution of specific ape and human kinds of intelligence (e.g. involving mental representations): clear evidence for a systematic monkey-ape difference in neocortex ratio is lacking. Great apes do not form systematically larger groups than monkeys do, which draws attention to physical rather than social factors (e.g. tool use, processing plant food etc.) that drove the evolution of mental representations in apes and humans. Why have in particular human apes evolved sophisticated representational and mental skills, are there any candidate factors that could have accelerated the evolution of human intelligence? Again, it seems most reasonable to start looking for factors in the social field of humans, given the fundamental social nature of human minds and how minds and human behavior develop, e.g. (Brothers 1990; Aronson 1994). Narrative psychology and studies on the development of autobiographic memory, e.g. (Nelson 1993; Conway 1996), and a *self* point towards an important factor, namely that *stories* are the most efficient and natural human way to communicate, in particular to communicate about others (Bruner 1987, 1990, 1991). According to Read and Miller (1995: 139), "Stories are so functional because social interaction is central to human beings, and stories are fundamentally about social interaction…stories are central to the human cognitive system because they capture the essence of social interaction, the structure of human action". Following this line of argument, the *Narrative Intelligence Hypothesis*, (Dautenhahn 1999c) proposes that the evolutionary origin of communicating in stories was correlated with increasing social dynamics among our human ancestors, in particular the necessity to communicate about *third-party relationships* (which in humans seems to reach the highest degree of sophistication among all apes, cf. gossip and manipulation, (Sinderman 1982)). As will be explained in more detail below, according to this hypothesis, human narrative intelligence might have evolved because the structure of narrative is particularly suited to communicate about the social world.

An evolutionary trend seens to exist from physical contact (non-human primates) to language (hominids) to communicating in stories (modern, highly 'enculturated' humans living in complex societies) correlated with an increase in complexity and sophistication of social interaction and mindreading. This trend demonstrates the evolution of increasingly efficient mechanisms for *time-sharing* the processes of social bonding. While physical grooming is generally a dyadic activity, language can be used in a variety of ways extending the dyadic use in dialogues to e.g. one-to many communication as it is today used extensively in the mass media (television, books, email etc.). It has been estimated (Dunbar 1993) that the human bonding mechanism of language is about 2.8 times as efficient as social grooming (the non-human primate bonding mechanism). Indeed, evidence suggests that conversational groups usually consist of one speaker plus 2 or three listeners. Of course larger groups can be formed easily, but in terms of actively participating and following different arguments within the group 1+2(3) seem to be the upper limit for avoiding information processing overload in the primate social brain. Also, language because of its representational nature affords documentation, preservation in storage media and transmission of (social) knowledge to the next generation, as well as communication between geographically separated locations (cf. (Donald 1993) for a discussion of language and external symbols in human cultural evolution).

Discussions in the social domain (e.g. on social relationships and feelings of group members) are fundamentally about *personal meaning*, different from e.g. discussions in the technical domain (e.g. about how to operate a tool or where to find food). We suggest that narrative might be the 'natural' format for encoding and re-constructing meaningful, socially relevant information (e.g. emotions and intentions of group members). According to Dunbar (1993) people spend about 60 % of conversations on gossiping about relationships and personal experiences. Humans use language to learn about other people and third-party relationships, to manipulate people, to bond with people, to break up or reinforce relationships.

Thus, a primary role of language might have been to communicate about social issues, to get to know other group members, to synchronize group behavior, to preserve group cohesion. Language is based on representations and the possibility to combine them in arbitrary ways. Representations need not be 'symbols', they can be spatial or visual in nature, and can be verbal or non-verbal. Apes can be trained to a subset of American Sign Language in order to communicate with humans, see e.g. studies with the chimpanzees Washoe and Nim (Gardner & Gardner 1969; Terrace et al. 1979), the gorilla Koko (Patterson & Linden 1981), or the orangutan Chantek (Miles 1990). Alternatively, icon-based keyboards (lexigrams) have been used in human-ape communication, e.g. Savage-Rumbaugh's studies with chimpanzees and bonobos such as Kanzi (Savage-Rumbaugh et al. 1986).

However, as of today, there is no convincing evidence that apes are using a symbolic, representational system in the wild on a level of complexity that can be compared to human language. Non-human apes do communicate extensively with each other, using gestures, vocalizations, eye-contact, and a range of 'body language'. While these means allow efficiently to communicate about the "here and now", they do not support a broadening of the *temporal horizon* (i.e. communicating about the past, future (Nehaniv 1999)), and events and group members that are absent. Obviously, there has not been a strong selective advantage for non-human apes in developing elaborate symbolic representational systems, although primate politics shows that non-primates do take into consideration the past and the future when deciding on how to behave socially, e.g. when predicting the behavior of conspecifics. Thus, non-human apes seem to possess mental representations, but it is unclear whether these representations are symbolic. Therefore, in terms of mental and communication skills humans and other apes have a lot in common, they possess mental representations and communication systems, but only humans possess an elaborate *symbolic/linguistic representational system* that is necessary for communicating via *human language* (cf. (Cheney & Seyfarth 1990) for discussions on communication systems in vervet monkeys). Interestingly, as Oliphant (Oliphant 1999) points out, a representational system which can learn word-meaning associations need not be computationally very expensive. Therefore, the information processing capacity of the brain can not be responsible for the fact that humans use language and chimpanzees in the wild do not. However, it is important to note that the form of human languages as such is meaningless. Words and sentences become meaningful only as a result of a cognitive effort that *creates* meaning and puts messages in *context*. The ability to construct and give meaning to representations is a 'computationally' expensive process, e.g. it requires identification and interpretation of the context of the communicative event, such as the personality/character of the sender (is he trustworthy?), the relationship between 'sender' and 'recipient' of a message (potential mate? competitor?), important third-party relationships, positions in the group hierarchy etc. Thus, one and the same 'message' can have potentially many different interpretations and 'meanings', depending on the complexity of the primate social field (discussed below), the number of different roles an individual can have, and the potential to create new roles and relationships.

Although humans use gestures, facial expressions, body language and other non-verbal means to convey (social) meaning, human communication is dominated by verbal communication, which is serial in nature (although in face-to-face interaction accompanied by non-verbal cues). Thus, given the serial communication channel of human language, what is the best means to communicate social issues, namely learning about the who, what, and why? Physical social grooming, the main group cohesion mechanism in non-human primates is 'holistic', parallel, spatial, sensual, meaningful. How can a stream of symbols that are in themselves meaningless convey meaning such as bodily grooming does? I argue that narrative structure seems to be particularly suited: usually a narrative gives a certain introduction of the characters (making contact between individuals, actors, listener and speaker), develops a plot, namely a sequence of actions that convey meaning (value, pleasurable, unpleasurable), usually with a high point and a resolution (reinforcement or break-up of relationships), and focuses on *unusual* events rather than stereotypical events. In this way, stories seem to give language a structure which resembles (and goes beyond) physical grooming, namely replacing physical presence and actions by the creation of a mental picture of physical actions, providing the stage, actors, intentions and a storyline. Story-telling also gives more flexibility than social grooming as to the actors and content of the stories: stories can include people that are part of the current audience, as well as absent persons, historical characters, fictional characters, etc. Stories that are told by a skilled story-teller (e.g. using appropriate body language, exploiting prosody, and possessing a rich repertoire of verbal expressions) can give very good examples of the power of words. The format of a story can provide sensual, emotional, and meaningful aspects to otherwise 'factual' information, e.g. poetry gives numerous examples of stories that can elicit emotional responses and influence people.

Thus, both story-telling in humans and social grooming in non-human primates are efficient social bonding mechanisms.

To summarize this section, *narrativity*, the capacity to communicate in terms of stories, is regarded as an efficient means to communicate social matters, and the origin of narratives might therefore have been a crucial milestone in the evolution of primate social intelligence (Read & Miller. 1995: 150) "It is because of the social, and the need to effectively manage social interactions, that we developed stories – stories made for the cognitively complex humans that we are. It is our stories that make us human". According to the *Narrative Intelligence Hypothesis* (NIH), the evolutionary origin of stories and narrativity was correlated with increasing social dynamics in human primate societies, in particular the need to communicate about third-party relationships. The evolution of the human story-telling mind was possibly correlated with the evolution of complex mechanisms of social understanding and a complex social field. This suggests that if we intend to develop a socially intelligent agent (Dautenhahn 1998) which can truly understand and respond to stories in human-agent interaction, then we need to model at least to a certain extent social relationships and primate social life. In the following sections we analyze the primate social field, and in more depth social understanding and the role of narrative in autobiography.

The primate social field

The primate family tree split up about forty million years ago into prosimians, which might resemble early arboreal primates (e.g. lemurs), and anthropoids (monkeys, apes, incl. humans). The problems of social life are especially complex for species whose cognitive skills create a complex *social field* which is based on several fundamental components:

1. Individuals specifically recognize other conspecifics in their groups as individuals and as kin. Primate societies are *individualized societies*. The social world of primates is primarily vision-dominated, recognition of friends and relatives and their behavior is therefore strongly based on visual cues, e.g. faces need to be recognized and memorized.

Two separate mechanisms have been proposed for kin-recognition: early familiarity (i.e. previous experience with the individuals in question) and phenotypic matching (using visual or non-visual cues). Generally, it is assumed that kin recognition in primates depends on previous experience. However, chimpanzees have been shown to be able to match related but unknown individuals by visual cues, in the same way as humans can match persons in a family album. In the wild, chimpanzees form loosely organized fission-fusion communities where even closely related individuals spend considerable time apart. Under such conditions phenotypic kin recognition could be greatly advantageous. As Parr and de Waal showed (Parr et al. 1999), chimpanzees can perceive similarities in the faces of related but unfamilar individuals, indicating visual kin recognition at a purely phenotypic level. Their results show that chimpanzees can match very well faces of mothers and their sons, but not mother-daughter pairs. This preference might be due to the particular ecological and social conditions of chimpanzee life.

How individual recognition substantially increases social complexity is shown by the following example described in (Philips & Austad 1996: 265):

> "...imagine a social group composed of six individuals, two unrelated sets of three full siblings. Consider an individual within that group seeking to join two other individuals for the purposes of cooperative hunting. With recognition only of group members versus nongroup members, there is only one recognizable hunting group – himself plus two other group members. If kinship were also recognized, then this individual could discriminate between three kind of groups (two fellow sibs, two nonsibs, one sib and one nonsib). If all group members were individually recognizable, our focal individual could potentially join twenty unique groups."

Thus, the more individuals can be recognized, the greater the number of social contexts recognized which can potentially lead to different responses and interpretations of communicated signals. If an animal can recognize group members individually, then it opens up a large set of choices, choices of who to join with, collaborate with, make friends with etc. The animal's situation is then much different from that of an animal that perceives itself as a member of a large anonymous group of (almost) identical group members. Living in an individualized group poses great cognitive challenges and can enhance the richness and diversity of social life in a group.

2. Individuals can understand and predict at least part of the behavior of other animals. Emotional information needs to be processed, in particular they need to recognize and act on cues to other animals' emotional states.

A variety of behavioral and contextual clues are used to predict another animal's behavior. The human ape is possibly the most social animal of all primates, and shows highly complex social structures and organizations. Elaborate mechanisms of social understanding, including sympathy and empathy (discussed below), a rich body language and facial expressions which are used to express internal states, moods etc. facilitate communication. Humans from a certain age on also attribute mental states to others, they possess a *theory-of-mind* (cf. (Leslie 1987), (Baron-Cohen 1995)) and can reason about beliefs, desires, wishes and goals of others. The abilities of humans to get along with each other, despite frequent violent encounters, is remarkable. Imagine one hundred chimpanzees, unfamiliar with each other, crowded in a metro coach. Very soon injuries, even deaths of animals are almost certain to occur. However, millions of (human) commuters survive exactly the same scenario day after day. Surviving in large 'anonymous' groups of people is controlled in human society by a number of norms and regulations. Thus, humans can not only understand and predict individuals, they can apply the same mechanisms to a crowd (as a kind of meta-organism).

3. Individuals remember aspects of previous interactions with group members and so form *dyadic, direct relationships* with them.

This involves remembering rank and past affiliations of group members. Even personal histories (e.g. who helped or received help) might be remembered. Cognitive processes of learning and memory make this possible.

4. Individuals need to remember dyadic relationships in the whole group, i.e. interactions other group members have with each other. This allows them

to understand the social relationships of others, i.e. their *third-party relationships*. Such relationships need to be recognized and memorized. Individuals need to be able to manipulate information about a set of relationships, e.g. for the purpose of forming alliances or tactical deception.

Kinship (based on certain patterns of association rather than on genetics), friendship (based on relatively recent aggressive or affiliative encounters) and dominance rank are all involved in the most important kinds of relationships recognized by primates. Many other avian and mammalian species are able to recognize individual group mates, remember past interactions with them, and predict their behavior, but it is not clearly established whether and/or to what extent they understand third-party relationships, i.e. relationships that group members have with one another. Enculturated animals (e.g. chimpanzees that grow up in a human family, but also pet animals such as dogs and cats) often show quite human-like social tactics (e.g. deception, cf. (Byrne 1997)), and they can even show cognitive skills different from their mother-reared cousins (e.g. improvement of imitative skills in enculturated chimpanzees, cf. (Tomasello et al. 1993)). It is at present therefore difficult to compare primate social intelligence with social intelligence in non-primate mammals. For more information on the primate social field see (Tomasello & Call 1997), (Dunbar 1998).

In terms of social complexity (and cognitive processes needed to deal with it), the world of an animal which takes into account third party actions is more complex than the world of an animal which only interacts dyadically. The social problems are still greater if an animal takes into account the probable thoughts as well as actions of its partners in interaction (Byrne & Whiten 1997).

The social life of Lemur catta

Here is an example of the social life of a non-human primate. The primate Center at Duke University gives the following information on *Lemur catta*, see Figure 1, a prosimian primate unique to Madagascar:

> "Ring-tailed lemurs are found in social groups of 3–25 individuals. Females remain in the group to which they were born for their entire lives, while males may change groups when they reach sexual maturity. Ringtail groups range over a considerable area each day in search of food. All group members use this common home range, and groups are often aggressive towards other groups at the borders of these areas. Females are usually dominant to males, which gives them preferential access to food and the choice of whom to mate with.

> (Female dominance in primates is unique to prosimians.) Social bonds within the group are established and reinforced by grooming. Prosimians groom in a rather unique way, all prosimians (ringtail lemurs included) have six lower teeth that stick straight out from their jaw, forming a comb that the animals use to groom their fur and the fur of other members of their social group." (http://www.duke.edu/web/primate/).

Lemur catta is very popular with many people because these creatures are seen as very gentle and 'friendly' primates. According to Jolly (1966), the fact that social lemurs show the usual primate type of society and social learning without the capacity to manipulate objects as monkeys do, might indicate the primacy of social intelligence in the evolution of primate intelligence. Although it is likely that lemurs can interpret a variety of social cues and use body language and social grooming as social cohesion mechanisms, they are not known to be elaborate story-tellers. According to Nelson (Nelson 1993: 12), when human primates are growing up "an important development takes place when the process of sharing memories with others through language becomes available as a means of reinstating memory...Language opens up possibilities for sharing and retaining memories in a culturally shared format for both personal and social functions. Sharing memory narratives is important to establish the new social function of autobiographical memory, as well as to make reinstatement through language possible." Thus, autobiographical memory as we know it, i.e. human-style autobiographical memory, seems to go hand in hand with the development of language. Lemurs are not likely to be able to communicate with us by telling stories about themselves and others, although their non-verbal communication system might be rich (and, as one can speculate, possibly even have narrative structure (Dautenhahn 2001)). However, humans interpret the lives of these gentle and beautiful lemurs in the most natural way, namely as stories and tales, and we cannot do otherwise.

For investigations into animal minds we cannot hand out questionnaires or conduct interviews, information can only be gained via observing natural behavior in the wild and/or conducting laboratory experiments under controlled conditions. Due to the difficult nature of gaining results that can withstand scientific/methodological scrutiny, many issues regarding animal minds (e.g. imitation, empathy, mindreading) are still highly controversial. We cannot directly look into a lemur's mind, neither do we know what kind of stories elephants or *cetaceans* are telling, and what a story could mean to their lives in the first place. However, imagine that young dolphins grow up while being taught the structure of narratives through story-telling, with their parents, peers and relatives, then the structure of these stories can be expected to be well adapted to life and

Figure 1. Foto of lemur catta. http://www.scz.org/animals/l/rtlemur2.html.

living as a dolphin, and adapted to the structure of the dolphin's mind, and it might turn out not to be compatible to the human mind. The way humans tell stories might only be one instantiation in a huge space of possible story-telling minds, natural and artificial.

Stories, social understanding, and autobiographic agents

Previously we suggested that two mechanisms are important to human social understanding: 1) empathic resonance, the ability to 'open' oneself towards another self, and to re-experience part of the other person's experiences, and 2) biographical reconstruction, the interpretation of another person's behavior and appearance based on the situatedness of another's mind in time and space (Dautenhahn 1997). The behavior and appearance of any biological agent can only be understood with reference to its history, considering its context, past, present and future situations. This is particularly important for life-long learning human agents who are continuously learning about themselves and their environment and are able to modify their goals and motivations. Autobiographical memory develops during the lifetime of a human being, and the capacity to fully develop an autobiography is not innate. In Nelson's discussion of the social origins of autobiographical memory in children she supports the *social interaction hypothesis*, namely that children gradually learn the forms of how to talk about memory with others, and thereby learn how to formulate their own memories as narratives (Nelson 1993).

Humans are constantly telling and re-telling stories about themselves and others. Humans are autobiographic agents, agents which are embodied and situated in a particular environment (including other agents), and which dynamically reconstruct their individual 'history' (autobiography) during their lifetimes (Dautenhahn 1996). The biologist Steven Rose uses the term lifelines in order to refer to a living organism's trajectory through time and space which make each organism an *individual*: "...it is in the nature of living systems to be radically indeterminate, to continually construct their – our – own futures, albeit in circumstances not of our own choosing" (Rose 1997: 7).

Telling (part of) a plausible autobiographical story to others is more than relating a plausible sequence of episodic events; it includes the construction of a plausible story based on one's goals, intentions and motivations. If we listen to a story originating from a completely different cultural background, the main problem of understanding is usually not to figure out what the actors *do*, but *why* they are doing it, i.e. understanding their goals and intentions. Once we understand the underlying motivations for their behavior, it helps us to make the link to similar situations which we, the listeners, experienced ourselves. We then might recall events which are from their appearance completely different, but with a similar meaning for us, which allows an understanding on a level of similarity which addresses the experiential, rather than cognitive, aspects of story understanding.

This creative aspect of story-telling, i.e. to tell autobiographic stories about oneself and create biographic re-constructions about other persons, is linked to the empathic, experiential way of relating other persons to oneself. Story-telling is a central mechanism in human social understanding.

Relationship between social and narrative intelligence: The case of autism

I argued above that in human evolution narrative capacities evolved from the need to effectively manage social dynamics, socially bond with others, exchanging information on third-party relationships etc. In this section I discuss that an impairment of narrative skills might contribute to difficulties people with autism have with social relationships.

People with autism have generally great difficulty in social interactions and developing relationships with other people. They are impaired in reading social cues and facial expressions, which makes the human social world around

Figure 2. Two children with autism simultaneously interacting with the Labo-1 robot used in the AURORA project.

them frightening and unpredictable (Baron-Cohen 1995), (Trevarthen et al. 1996/98). A variety of therapy approaches are available, and the author is involved in the AURORA project (Autonomous robotic platform as a remedial tool for children with autism, http://www.aurora-project.com/) that develops a mobile robot as an interactive and therapeutic *toy,* (Dautenhahn 1999a, b), (Werry & Dautenhahn 1999), (Dautenhahn & Werry 2000). Figure 2 shows two children with autism simultaneously playing with the robot, part of a series of trials where we investigated the role of the robot as a social mediator (Werry et al. 2001).

Previously, (Dautenhahn 1997) I suggested that an impairment of the processes of empathic understanding and biographical reconstruction might contribute to the symptoms which people with autism show, who are generally not able to build up 'normal' social relationships, nor can they show 'adequate' behavior in social interactions (Howlin et al. 1999). People with autism definitely possess strong emotions, but they seem to lack the ability to view other persons as *mental agents*, as opposed to *physical objects*, which is a crucial prerequisite for empathy and attribution of emotions and mental states to other people. Moreover, children with autism generally do not show pretend-play with dolls or stuffed animals.

A set of standardized experiments are usually used to identify autistic symptoms in children, among them experiments in which a particular story is presented and the child has to answer questions about the actors' current beliefs (false belief test). The *Sally-Anne test* (Baron-Cohen et al. 1985) is about two dolls. (1) Sally and Anne are together in a room, (2) Sally puts a marble in a basket and leaves the room, (3) Anne takes the marble out of the basket and puts it into a box, (4) Sally returns. The child is then asked where Sally will look for the marble. This short story can be presented to the children in a variety of formats, e.g. told verbally with/without objects and cartoons, enacted with puppets or human beings etc. Normal children until the age of four and most autistic children (of all ages) give 'Anne's box' as the answer, i.e. they cannot attribute to Sally a different belief then they have themselves (and they know that the marble is now in Anne's box). Tests like the Sally-Anne test require you to be able to distinguish yourself and your beliefs and perceptions from those of others: what I {know, believe, perceive, feel} is not necessarily identical with what you {know, believe, perceive, feel}. This ability is not innate; children develop this ability during their first years of life. By the age of 3-4 years a child's *theory-of-mind* is usually well developed, while most children with autism will not succeed at this. The term 'theory-of-mind' has recently been replaced by the term *mind-reading*, in order to express that the skill to understand the social world is not necessarily *theory-based* (e.g. based on a set of axioms and logical rules). Moreover, interpersonal processes of joint attention and/or empathy are alternative approaches to understanding autism, see discussion e.g. in (Dautenhahn 1997). Failure of children with autism to pass the Sally-Anne test has usually been interpreted as a failure in the development of theory-of-mind or mindreading skills, cf. (Baron-Cohen 1995). However, as I will discuss in this section, an alternative explanation for such a failure, although related to mindreading skills, could lie in a failure to properly interpret, re-construct and understand *stories*, thus indicating an impairment of narrative capacities in children with autism, as suggested by psychologists such as Jerome Bruner and Carol Feldman (Bruner & Feldman 1993).

According to the developmental psychologist Katherine Nelson (Nelson 1986) children experience their day as a series of scripts (as suggested in (Schank & Abelson 1977)) and routines which help them to structure their world of experiences and language. Scripts help them to understand what is going to happen and who is going to do what. Nelson's evidence demonstrates the primacy of scripts as an organizing tool for children. However, as Bruner points out (Bruner 1991), narratives require scripts as necessary background (the *skeleton*), but they do not constitute narrativity itself. Scripts are not *worth telling* unless they include the *unusual*, breaches, violations to the script which make a story interesting. Thus, children only become true story-tellers once they can create and remember stories about the unusual, the specific, events and experiences that contribute to their unique and individual autobiography.

Interestingly, some people with autism show animal empathy (i.e. they can 'understand' the behavior and feelings of animals, (Grandin 1995)), so a mechanism of empathic resonance (with animals) seems to exist. Moreover, some high-functioning people with autism can learn and train themselves in social behavior to some extent, by learning and applying generic rules of human interaction, although they usually fail to recognize idiosyncratic social cues (i.e. they fail to construct another person's individual biographic history). Thus, we can expect that when people with autism are confronted with a complex 'social story' (enacted by actors in movies or comics, or by normal people in real life), that the more 'human-like' the actors in a story are, the more sophisticated their behavior is, i.e. the more biographical reconstruction of the story is required, the more difficulty people with autism will have in understanding the story.

Children with autism need structure in their lives, they prefer to stick to a fixed daily routine, and they have difficulty to remember and describe what *actually* happened to them, in contrast to what *usually* happens to them. These attributes are reminiscent of Nelson's evidence that the memory of preschool children is structured around the usual, routine episodes, until children become skilled story-tellers. This indicates an impairment of narrative skills in children with autism, in particular those narratives which are special and individual and which contribute to autobiographical memory. One reason for the difficulties people with autism have in relating, understanding and communicating with other people might therefore lie in an impairment of narrative, story-telling skills, i.e. an impairment of the ability to represent the characteristic narrative shape of human action and interaction (Bruner & Feldman 1993). From early childhood on, through transactions with others, e.g. in mutual imitation games (Nadel et al. 1999), children learn the 'narrative format' of human interaction, an important milestone in the development of a child's understanding of other minds (cf. discussion in Jordan (1999)). Humans are not only *mental agents*, they are agents with a history, autobiographic agents, interlinked with the histories of other agents in the social field. Social understanding requires an autobiographic agent which is able to re-construct its own and other people's experiences, an agent with a history, an agent which has a body as the point of reference which gives a unique perspective on the (social) world, which allows one to generalize from experiences and to reconstruct specific, individual experiences.

Interestingly, Howlin et al. (1999) who developed a cartoon-based practical guide to teaching children with autism to mindread, pointed towards the importance of *social context and history* in teaching social understanding to children with autism: "Understanding – and reacting appropriately to – people's emotions, involves more than the ability to recognize a few clear and relatively simple emotions from pictures and cartoons. Whether a situation is construed as being happy, sad or frightening will depend, not only on the current context but on the past history of the individual(s) involved. Moreover, facial expression alone may not always be a true representation of how someone is feeling – a smile, for example maybe used in a brave attempt to disguise sadness or pain. And, being able to recognize certain unambiguous emotions in other people, may not necessarily help children with autism fully understand or cope with their own emotional responses, especially if these differ from those of others."

How might one help children in general, and children with autism in particular to become skilled story-tellers? Usually, for children growing up in a social context, surrounded and encouraged by story-telling adults and other children, and exposed to a variety of stories that are written, told or performed, story-telling skills are part of normal development (Engel 1999), without necessarily being explicitly taught. For children with autism, who were not able to follow that 'normal' path that leads to becoming a skilled story-teller, story-telling skills would have to be taught, explicitly, and in this sense 'artificially', i.e. making things explicit that are normally 'picked up' in a social context. For example, parents and peers do normally not tell a child explicitly 'remember this', 'don't remember this'. The autism researcher Powell (Stuart Powell 1999, pers. comm.) recommends that in teaching people with autism pointers have to be given explicitly about what is important and useful (to remember) and what is not, in this way helping them to structure their memory in order to create autobiographical stories that they can tell.

In this section I discussed the importance of story-telling and autobiography in the social and cognitive development of children. Systems that support children's story-telling (e.g. as investigated in many projects part of the EU initiative Experimental School Environments (ESE), cf. (Machado et al. 1999), (Bobick et al. 1999), (Benford et al. 2000)) might play an important role in a story-oriented education for pre-school and older children. Such new *narrative technology* can potentially meet the social and cognitive needs of young primate story-tellers. This section also discussed how research in narrative intelligence could potentially be applied to autism therapy.[3] Narrative technology of this kind needs to make narrative skills explicit, make the implicit visible, highlight the underlying structure, point out (and possibly explain) what is important to remember and what is not.

A new generation of humanoid robots might even be used in autism therapy in order to test and teach social skills. Humphrey (1976:1988) argues for the necessity of developing a laboratory test of 'social skill' for primates. His suggestion is:

> "The essential feature of such a test would be that it places the subject in a transactional situation where he can achieve a desired goal only by adapting his strategy to conditions which are continually changing as a consequence partly, but not wholly of his own behavior. The 'social partner' in the test need not be animate (though my guess is that the subject would regard it in an 'animistic' way); possibly it could be a kind of 'social robot', a mechanical device which is programmed on-line from a computer to behave in a pseudo-social way."

Thus, for Humphrey a test of social intelligence does not measure social 'reasoning', but addresses a social interaction situation. Nowadays we do have humanoid social robots (e.g. Breazeal et al. 1999, 2000) which, if they are accepted by human and non-human primates, could take the role of the interaction partner in such a social intelligence test. Generally, interactions between animate and inanimate social agents can indicate what kind of social knowledge is necessary in order to achieve a certain social behavior, e.g. how much 'theory' a social (and autobiographic) agent requires in order to be able to read others' minds. Systematic experimental tests e.g. with a social robot might also shed light on the role of narrative in social intelligence as discussed in this chapter.

Requirements for narrative agents

The evolution and development of natural social intelligence and story-telling is based on the primate social field. This chapter explained that (1) research in primatology points to the importance of social intelligence for the evolution of primate intelligence, and (2) autism shows how fundamentally an impairment of social skills, and possibly narrative skills, can influence the life of people, even if they show good non-social skills of intelligence. Thus, it seems that in order to make artificial (robotic or software) agents story-tellers, they need to be primarily *socially intelligent agents*[4] (Dautenhahn 1998). Based on our previous analysis of the primate social field the following list of necessary requirements for a story-telling agent is suggested (this is not supposed to be an exhaustive list):

1. Individualized societies: The capacity to identify and recognize individual group members.
2. Social Networks: the capacity to establish, maintain, remember and utilize social networks. Ability to *predict* the behavior of others and outcomes of interaction. Agents need enough 'experience' and background knowledge

in order to predict the future, and make the link to the past and present. Three basic elements in the primate social field are the following:

2.1 Ability to remember and learn interactions with others and to build *direct relationships*: As discussed above the upper limit of the group size was estimated for humans as 150, representing a cognitive limit on the number of individuals with whom one person can maintain stable relationships, as a function of brain size. The 'brain' of a software or robotic agent (at least in terms of storage capacity) can be huge. Thus, agents can have many friends.

2.2 Identifying third-party relationships (relationships among other group members), ability to remember and learn interaction between others. Since human communication is dominated by gossiping about other people, artificial agents talking about other agents seems to be suggested.

2.3 Ability to *understand* others, most elaborated in humans which show complex mechanisms of empathy, biographical reconstruction, and an individual autobiography. Agents need social skills, ways to figure out what other agents are doing and the ability to communicate with them.

2.4 Recognition of conspecifics as members in a group hierarchy/social structure (e.g. structures of kinship, allies, dominance hierarchies, etc.)

3. Efficient mechanisms of social bonding, either via physical grooming (in non-human primate societies) or via language and communication in narratives as efficient ways of *social bonding*, important for maintaining the coherence of social groups at different levels of social organization.
4. Social learning: the capacity to use others as *social tools* (Dautenhahn 1995), via social learning mechanisms with varying degrees of what the animals learn from each other (cf. social facilitation versus imitation)

We hope that in future work these requirements can be sufficiently addressed in the construction of socially intelligent *narrative agents*, e.g. socially intelligent robots, cf. (Dautenhahn & Nehaniv 1998), (Dautenhahn 1999), (Dautenhahn & Billard 1999), (Dautenhahn & Coles 2000).

Conclusion

Narrative agents as we know them, e.g. humans and other primates, are social agents, grow up in a society, learning about other agents and how to predict their behavior. Also, narrative might be at the center of who (we think) we are.

> "Our fundamental tactic of self-protection, self-control, and self-definition is not building dams or spinning webs, but telling stories – and more particularly

> concocting and controlling the story we tell others – and ourselves – about who we are. These strings or streams of narrative issue forth as if from a single source – not just in the obvious physical sense of flowing from just one mouth, or one pencil or pen, but in a more subtle sense: their effect on any audience or readers is to encourage them to (try to) posit a unified agent whose words they are, about whom they are: in short, to posit what I call a *center of narrative gravity*." (Dennett 1989/91)

In this chapter I discussed the issue of narrative and story-telling from the perspective of primate social behavior and primate evolution, hoping that knowledge of who we are (as a species and as an individual primate) helps us understand the broader context and significance of narrative in human life. A more detailed discussion and analysis of the transactional format of narratives in human and other animals is given in (Dautenhahn 2001).

Currently, a number of research project are devoted to building narrative software, virtual or physical environments e.g. (Glos & Cassell 1997), (Machado & Paiva 1999), (Umaschi-Bers & Cassell 1999), (Bobick et al. 1999), (Montemayor et al. 2000), (Benford et al. 2000). Supporting, and possibly expanding, human narrative intelligence is expected to impact human minds and our notions of sociality and what we call our *selves*. In parallel, investigations into autonomous story-telling agents might result in agents (robotic or software) with genuine narrative minds, able to tell us interesting stories, listen to and understand our stories, and make us laugh. As I argued in this chapter, the kind of stories these agents will tell us will be shaped by the social field and the cultural environment of human societies in which these agents *grow up*. Thus, it is up to us whether the stories of the future will be nightmares, fairy-tales, comedies or adventures.

Acknowledgements

The AURORA project is supported by EPSRC (GR/M62648), *Applied AI Systems Inc.*, and the National Autistic Society (NAS). Thanks to the editors Phoebe Sengers and Michael Mateas whose comments helped improving a previous version of this chapter.

Notes

1. This chapter is partially based on (Dautenhahn 1999c).

2. See (Gigerenzer 1997: 265) for an analysis of the SIH as a "collection of loosely related assertions about the special role of the social (i.e. intraspecific) in the intellectual life of humans and other primates".

3. See other work (not necessarily involving narrative) that uses computer and robot technology in autism therapy, e.g. (Weir & Emanuel 1979), (Strickland 1996), (Blocher 1999), (Dautenhahn 1999), (Werry & Dautenhahn 1999), (Dautenhahn & Werry 2000), (Dautenhahn 2000a), (Charitos et al. 2000), (Parsons et al. 2000), (Michaud et al. 2000), Autism & Computing (http://www.shifth.mistral.co.uk/autism/NAS/).

4. Selected literature on Socially Intelligent Agents: (Dautenhahn & Numaoka 1998; Dautenhahn 2000b; Edmonds & Dautenhahn 1999; Dautenhahn 2000c, 2000d).

References

Aiello, L. C. & P. Wheeler (1995). The expensive tissue hypothesis. *Current anthropology*, *16* (36), 184–193.

Aronson, E. (1994). *The social animal.* New York: W.H. Freeman.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* Cambridge, MA & London: The MIT Press.

Baron-Cohen, S., A. M. Leslie, & U. Frith (1985). Does the autistic child have a "theory of mind". *Cognition, 21,* 37–46.

Barton R. A. & R. I.M. Dunbar (1997). Evolution of the social brain. In (Whiten & Byrne, 1997: 240–263).

Benford, S., B. B. Bederson, K.-P. Akesson, V. Bayon, A. Druin, P. Hansson, J. P. Hourcade, R. Ingram, H. Neale, C. O'Malley, K. T. Simsarian, D. Stanton, Y. Sundbald, & G. Taxen (2000). Designing storytelling technologies to encourage collaboration between young children. *Proc. CHI 2000*, April 1–6 2000, The Hague, The Netherlands.

Blocher, K. H. (1999). *Affective Social Quest (ASQ): Teaching emotion recognition with interactive media and wireless expressive toys*. MSC Thesis, MIT, USA.

Bobick, A. F., S. S. Intille, J. W. Davis, F. Baird, C. S. Pinhanez, L. W. Campbell, Y. A. Ivanov, A.Schütte, & A. Wilson (1999). The Kidsroom: A perceptually-based interactive and immersive story environment. *Presence, 8* (4), 369–393.

Breazeal, C. & B. Scassellati (1999). How to build robots that make friends and influence people. *Proc. IROS99*. Kyonjiu, Korea.

Breazeal, C. & B. Scassellati (2000). Infant-like social interactions between a robot and a human caretaker. *Adaptive behavior, 8* (1).

Brothers, L. (1990). The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts in neurosciences*, *1* (1), 27–51.

Byrne, R. W. & A. Whiten, (Eds.), (1988). *Machiavellian intelligence*. Clarendon Press.

Byrne, R. W. (1997). Machiavellian intelligence. *Evolutionary anthropology, 5* (5), 172–180.

Byrne, R. W. (1999). Human cognitive evolution. In M. C. Corballis & S. E. G. Lea (Eds.), *The descent of mind: Psychological perspectives on hominid evolution* Chapter 4 (pp. 71–87). Oxford University Press.

Bruner, J. (1987). *Actual minds, possible worlds*. Cambridge: Harvard University Press.

Bruner, J. (1990). *Acts of meaning*. Cambridge: Harvard University Press.
Bruner, J. (1991). The narrative construction of reality. *Critical inquiry, 18* (1), 1–21. Also appears in this volume.
Bruner, J. & C. Feldman (1993). Theories of mind and the problem of autism. In S. Baron-Cohen et al. (Eds.), *Understanding other minds: Perspectives from autism*. Oxford: Oxford University Press.
Chance, M. R. A. & A. P. Mead (1953). Social behaviour and primate evolution. *Symposia of the society for experimental biology VII (evolution)*, 395–439.
Charitos, D., G. Karadanos, E. Sereti, S. Triantafillou, S. Koukouvinou, & D. Martakos (2000). Employing virtual reality for aiding the organisation of autistic children behaviour everyday tasks. In P. Sharkey, A. Cesarani, L. Pugnetti, & A. Rizzo (Eds.), *Proc. 3rd international conference on disability, Virtual Reality & associated technologies* (pp. 147–152). University of Reading.
Cheney, D. L. & R. M. Seyfarth (1990). *How monkeys see the world*. University of Chicago Press.
Conway, M. A. (1996). Autobiographical knowledge and autobiographical memories. In D. C. Rubin (Ed.), *Remembering our past. Studies in autobiographical memory* (pp. 67–93). Cambridge University Press.
Dautenhahn, K. (1995). Getting to know each other: Artificial social intelligence for autonomous robots. *Robotics and autonomous systems, 16*, 333–356.
Dautenhahn, K. (1996). Embodiment in animals and artifacts. In *Embodied cognition and action* (pp. 27–32). AAAI Press, Technical report FS-96-02, 1996.
Dautenhahn, K. (1997). I could be you: The phenomenological dimension of social understanding. *Cybernetics and systems journal, 28* (5), 417–453.
Dautenhahn, K. (1998). The art of designing socially intelligent agents: Science, fiction, and the human in the loop. *Applied Artificial Intelligence, 12* (7–8), 573–617.
Dautenhahn, K. (1999a). Robots as social actors: AURORA and the case of autism. *Proc. CT99, The third international cognitive technology conference*, August, San Francisco, 359–374, URL: http://www.cogtech.org/CT99/Dautenhahn.htm (last accessed 19th July 2002).
Dautenhahn, K. (1999b). Embodiment and interaction in socially intelligent life-like agents. In C. L. Nehaniv, (Ed.), *Computation for metaphors, analogy and agents* (pp. 102–142). Springer Lecture Notes in Artificial Intelligence, Volume 1562.
Dautenhahn, K. (1999c). The lemur's tale – Story-telling in primates and other socially intelligent agents. In M. Mateas & P. Sengers, (Eds.), *Proc. Narrative Intelligence*, AAAI Fall Symposium 1999, AAAI Press, Technical Report FS-99-01 (pp. 59–66).
Dautenhahn, K. (2000a). Design issues on interactive environments for children with autism. In P. Sharkey, A. Cesarani, L. Pugnetti, & A. Rizzo (Eds.), *Proc. 3rd international conference on disability, Virtual Reality & associated technologies* (pp. 153–161). University of Reading.
Dautenhahn, K. (Ed.) (2000b). *Human cognition and social agent technology*. John Benjamins Publishing Company.
Dautenhahn, K. (Ed.) (2000c). Simulation Models of Social Agents, special issue of *Adaptive behaviour*, *7* (3–4).

Dautenhahn, K. (Ed.) (2000d). *Socially intelligent agents – The human in the loop*, Technical Report FS-00-04, http://www.aaai.org/Press/Reports/Symposia/Fall/fs-00-04.html, AAAI Press.

Dautenhahn, K. (2001). The Narrative Intelligence hypothesis: In search for the transactional format of narratives in humans and other social animals. *Proc. CT2001, The fourth international conference on cognitive technology*, Lecture Notes in Computer Science, Springer Verlag, 248–266.

Dautenhahn, K. & C. L. Nehaniv (1998). Artificial life and natural stories. In *Proc. third international symposium on Artificial Life and robotics* (AROB III'98 – January 19–21, 1998, Beppu, Japan), volume 2, 435–439.

Dautenhahn, K. & C. Numaoka (Eds.) (1998 and 1999). *Socially intelligent agents*, Special Issues of *Applied Artificial Intelligence*, *12* (7–8) *13* (3).

Dautenhahn, K. & Aude Billard (1999). Studying robot social cognition within a developmental psychology Framework. *Proc. Eurobot99, Third European workshop on advanced mobile robots* (pp. 187–194), September 1999, Switzerland, 1999.

Dautenhahn, K. & I. Werry (2000). Issues of robot-human interaction dynamics in the rehabilitation of children with autism. *Proc. From animals to animats, The sixth international conference on the simulation of adaptive behavior (SAB2000)* (pp. 519–528), 11–15 September 2000, Paris, France, 2000.

Dautenhahn, K. & Steven Coles (2001). Narrative Intelligence from the bottom up: A computational framework for the study of story-telling in autonomous agents. In *Journal of artificial societies and social simulation (JASSS)*, Special issue on Starting from Society: The application of social analogies to computational systems, *4* (1).

Dennett, D. C. (1989/91). The origins of selves. *Cogito,* 3, 163-73, Autumn 1989. Reprinted in Daniel Kolak and R. Martin, eds., (1991), *Self & identity: Contemporary philosophical issues*, Macmillan.

de Waal, F. (1982). *Chimpanzee politics: Power and sex among apes*. Jonathan Cape, London.

Donald, M. (1993). Precis of *Origins of the modern mind: Three stages in the evolution of culture and cognition*. *Behavioral and brain sciences, 16,* 737–791.

Dunbar, R. I. M. (1992). Neocortex size as a constraint on group size in primates. *Journal of human evolution*, *20,* 469–493.

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and brain sciences*, *16*, 681–735.

Dunbar, R. I. M. (1996). *Grooming, gossip and the evolution of language*. Faber and Faber Limited.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary anthropology*, *6,* 178–190.

Dunbar, R. I. M. & J. Bever (1998). Neocortex size predicts group size in carnivores and some insectivores. *Ethology*, *104,* 695–708.

Edmonds, B. & K. Dautenhahn (1998). The contribution of society to the construction of individual intelligence. Technical Report CPM-98-42, Centre for Policy Modelling, Manchester, Metropolitan University, UK, 1998.

Edmonds, B. & K. Dautenhahn (Eds.) (1999). Social Intelligence, special issue of *Computational and mathematical organisation theory, 5* (3).

Engel, S. (1995/99). *The stories children tell: Making sense of the narratives of childhood*. W. H. Freeman.

Gardner, R. A. & B. T. Gardner (1969). Teaching sign language to a chimpanzee, *Science, 165,* 664–672.

Glos, J. W. & Justine Cassell (1997). Rosebud: A place for interaction between memory, story, and self. In. J.-P. Marsh, C. L. Nehaniv, & B. Gorayska, (Eds.), *Proceedings of the second international conference on cognitive technology* (pp. 88–97). IEEE Computer Society Press, 1997.

Gigerenzer, G. (1997). The modularity of social intelligence. In A. Whiten and R. W. Byrne, (Eds.), *Machiavellian intelligence II: Extensions and evaluations*, Chapter 10 (pp. 264–288). Cambridge: Cambridge University Press.

Grandin, T. (1995). *Thinking in pictures.* Doubleday Publisher.

Howlin, P., S. Baron-Cohen & J. Hadwin (1999). *Teaching children with autism to mind-read*. John Wiley.

Humphrey, N. K., (1976/1988). The social function of intellect. In R. W. Byrne & A. Whiten, (Eds.), *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans* (pp. 13–26). Oxford: Clarendon Press.

Jordan, R. (1999). *Autistic spectrum disorders – An introductory handbook*. London: David Fulton Publishers.

Jolly, A. (1966). Lemur social behavior and primate intelligence. *Science, 153,* 501–506.

Leslie, A. M. (1987). Pretense and representation: The origins of 'theory of mind'. *Psychological review, 94* (4), 412–426.

Machado, I., C. Martinho, & A. Paiva (1999). Once upon a time. In M. Mateas & P. Sengers (Eds.), *Proc. Narrative Intelligence*, AAAI Fall Symposium 1999, AAAI Press, Technical Report FS-99-01 (pp. 115–119).

Marino, L. (1996). What can dolphins tell us about primate evolution? *Evolutionary anthropology*, *5* (3), 81–86.

Michaud, F., P. Lepage, J.-D. Leroux, M. Clarke, F. Bélanger, Y. Brosseau, & D. Neveu (2000). Mobile robotic toys for autistic children. *Proc. international symposium on robotics*, Montréal, May 2000.

Miles, H. L. W. (1990) The cognitive foundations for reference in a signing orangutan. In S. T. Parker & K. R. Gibson (Eds.), *"Language" and intelligence in monkeys and apes* (pp. 511–539). Cambridge: Cambridge University Press.

Mithen, S. (1996). *The prehistory of the mind*. London: Thames and Hudson.

Montemayor, J., H. Alborzi, A. Druin, J. Hendler, D. Pollack, J. Porteous, L. Sherman, A. Afework, J. Best, J. Hammer, A. Kriskal, A. Lal, T. Plaisant Schwenn, L. Sumida, & R. Wagner (2000). From PETS to Storykit: Creating new technology with an intergenerational design team. *Proc. 2000 workshop on interactive robotics and entertainment (WIRE-2000)* April 30–May 1, 2000, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.

Nadel, J., C. Guerini, A. Peze, & C. Rivet (1999). The evolving nature of imitation as a format of communication. In J. Nadel & G. Butterworth (Eds.), *Imitation in infancy* (pp. 209–234). Cambridge: Cambridge University Press.

Nehaniv, C. L. (1999). Narrative for artifacts: Transcending context and self. In M. Mateas & P. Sengers, (Eds.), *Proc. Narrative Intelligence* (pp. 101–104). AAAI Fall Symposium 1999, AAAI Press, Technical Report FS-99-01.

Nelson, K. (1986). *Event knowledge: Structure and function in development*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Nelson, K. (1993). The psychological and social origins of autobiographical memory, *Psychological science, 4* (1), 7–14.

Oliphant, M. (1999). Cultural transmission of communication systems: Comparing observational and reinforcement learning models. In K. Dautenhahn & C. L. Nehaniv, (Eds.), *Proc. AISB'99 symposium on imitation in animals and artifacts* (pp. 47–55). Society for the Study of Artificial Intelligence and Simulation of Behaviour, 1999.

Philips, M. & S. N. Austad (1996). Animal communication and social evolution. In M. Bekoff & D. Jamieson, (Eds.), *Readings in animal cognition* (pp. 257–267). MIT Press.

Parr, L. A. & F. . M. de Waal (1999). Visual kin recognition in chimpanzees. *Nature, 399,* 647–648.

Parsons, S., L. Beardon, H. R. Neale, G. Reynard, R. Eastgate, J. R. Wilson, S. V. G. Cobb, S. D. Benford, P. Mitchell, & E. Hopkins (2000). Development of social skills amongst adults with Asperger's Syndrome using virtual environments: the 'AS Interactive' project. In P. Sharkey, A. Cesarani, L. Pugnetti, & A. Rizzo (Eds.), *Proc. 3rd international conference on disability, Virtual Reality & associated technologies* (pp. 163–170). University of Reading.

Patterson, F. & E. Linden (1981). *The education of Koko*. New York: Owl books.

Pawlowski, B., C. B. Lowen, & R. I. M. Dunbar (1998). Neocortex size, social skills and mating success in primates. *Behaviour*, *135,* 357–368.

Povinelli, D. J. & T. M. Preuss (1995). Theory of mind: Evolutionary history of a cognitive specialization. *Trends in cognitive neurosciences, 18*, 418–424.

Premack, D. & G. Woodruff (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences, 4,* 515–526.

Read, S. J. & L. C. Miller (1995). Stories are fundamental to meaning and memory: For social creatures, could it be otherwise? In R. S. Wyer, (Ed.), *Knowledge and memory: The real story*, Chapter 7 (pp. 139–152). Hillsdale, NJ: Lawrence Erlbaum.

Rose, S. (1997). *Lifelines. Biology, freedom, determinism*. Penguin Books.

Russell, R. J. (1993). *The lemurs' legacy. The evolution of power, sex, and love*. NY: G.P. Putnam's Sons.

Savage-Rumbaugh, E. S., K. McDonald, R. A. Sevcik, W. D. Hopkins, & E. Rubert (1986). Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (*Pan paniscus*). *Journal of experimental psychology: General, 115*, 211–235.

Schank, R. C. & R. P. Abelson (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.

Sindermann, C. J., (1982). *Winning the games scientists play*. New York & London: Plenum Press.

Strickland, D. (1996). A virtual reality application with autistic children. *Presence: Teleoperators and Virtual Environments, 5* (3), 319–329.

Terrace, H. S., L. A. Petitto, R. J. Sanders, & T. G. Bever (1979). Can an ape create a sentence? *Science, 206,* 891–902.

Tomasello, M. & J. Call (1997). *Primate cognition*. Oxford: Oxford University Press.

Tomasello, M., A.C. Kruger, & H.H. Ratner (1993). Cultural learning. *Behavioral and brain sciences, 16* (3), 495–552.

Trevarthen, C., K. Aitken, D. Papoudi, & J. Robarts (1996/98). *Children with autism*. London & Philadelphia: Jessica Kingsley Publishers.

Turner, M. (1996). *The literary mind*. Oxford University Press.

Umaschi Bers, M. & J. Cassell (1999). Children as designers of interactive storytellers: "Let me tell you a story about myself". In K. Dautenhahn, (Ed.), *Human cognition and social agent technology*, Chapter 16 (pp. 61–83). Amsterdam: John Benjamins Publishing Company.

Weir, S. & R. Emanuel (1976). Using LOGO to catalyse communication in an autistic child. Technical report, DAI Research Report No. 15. University of Edinburgh, 1976.

Werry, I. & K. Dautenhahn (1999). Applying robot technology to the rehabilitation of autistic children. *Proc. SIRS99*, 7th international symposium on intelligent robotic systems '99 (pp. 265–272).

Werry, I., Dautenhahn, K., Ogden, B. & W. Harwin (2001). The role of a robotic mediator in autism therapy. Proc. *The fourth international conference on cognitive technology: Instruments on mind (CT2001),* Monday 6th–Thursday 9th August, 2001 at University of Warwick, United Kingdom, Springer Verlag, Lecture Notes in Computer Science, subseries Lecture Notes in Artificial Intelligence, 57–74.

Whiten, A., (Ed.), (1991). *Natural theories of mind: Evolution, development and simulation of everyday mindreading*. Basil Blackwell.

Whiten, A., & R. W. Byrne, (Eds.), (1997). *Machiavellian intelligence II: Extensions and evaluations.* Cambridge: Cambridge University Press.

Chapter 5

Vital narratives

Brenda Laurel
Los Gatos, CA

American culture has a hole in its heart. Vital narrative forms that nurture and define us are vanishing. Our culture also has a hole in its head. We are unclear about the epistemological roots of the various forms of narrative we are exposed to, and, more importantly, we are befuddled about how to judge what sorts of actions might appropriately flow from these narratives.

This is an opinion piece. The paper is fundamentally about different kinds of narratives that I think we need to have and some that I think we need to change in the world today. It is about a certain kind of narrative literacy that can help people to form more complex and appropriate narratives for their lives. I am focused on kids and teens as an audience because they are the future, but I think that the analysis applies to adults as well.

Any of the types of narratives I discuss here can be represented as interactive fiction, games, or web-based content. In the case of a story that is written or told, the widely accepted reader-response theory says that the reader plays an active role in constructing the story. Readers find ways to make good stories personally relevant. Empathy, for example, plays a key role in the construction of relevance.

In interactive media, the reader's role in construction is more pronounced. It is important that the form not get in the way of this act of construction. Elements of personal agency – that is, the ability to do something with the material in the real or virtual world – is a key to engaging kids in interactive narrative. That said, I leave the formal questions to others (for a change) and look in this paper at the content and intent of various kinds of stories as they provide scaffolding for constructing meaning and deriving actionable goals.

Storytelling as relationship

Story is an object; *storytelling* is a relationship. The statement that we are lacking certain genres of vital stories really means that we are missing certain kinds of vital relationships.

A few years ago, one of the questions I asked in the context of a larger research project on kids was, do your parents tell you stories? Many kids said that their parents read them stories, but very few said their parents told them stories. Children's literature can be seen as a set of tales tailored in a general way to the needs and questions of kids. When a parent reads a child a story (or belongs to a parent-child book club, for example) then the relationship is present indirectly in the reading. Parents *should* read to their children.

But children also need personal narratives to connect them to their relatives, friends, their culture, and other world-views. It emerged in my research that the family storyteller was often the grandmother or other relative in a family who could tell "remember when" stories. But as families are more often separated geographically, the elder storyteller may not be able to form storytelling relationships. This is one way in which the technologies of telecommunications might serve us well. Certainly, the disappearance of rural life contributes to the disappearance of storytelling. Since the beginning of the 20th century, the number of Americans engaged in farming has shrunk from 50% to 2%. Along with them go the rural cultures that fostered storytelling.

Doubtless, many parents tell their kids stories about their own youth, as cautionary tales, for instance. But how do we talk to our children about our lives now? How do we speak – figuratively or literally – about our own beliefs and ethics? One kind of storytelling relationship that children often lack is the stories of parents' lives and work, the choices that must be made, the difficulties that must be faced, the joys of doing a job well. Such stories help children understand what it is to be an adult, and to expose children to the constructions that an adult must have to navigate the world.

Another blow to the culture of American storytelling has been struck by media, beginning with radio and silent film. As every parent of a preschooler knows, the temptation is great to allow television (the blessed kind, like "Children's Television Workshop") to act as a babysitter and source of education. That is all well and good, except for the relationship part. Kids have pretend relationships with characters, but that is not the same thing as a relationship with a person. This, too, can be addressed in some ways by technology; for example, on personal storytelling websites like Bubbe's Back Porch [www.bubbe.com] and The Fray [www.fray.com]. There are even websites for kids to tell and hear stories about various social problems like bullying [see, for example, www.bullying.org.].

In cultures where storytelling holds a significant place, storytellers choose tales to address the needs or context of their listeners. Native lore of all sorts falls in this category. The performance is live and responsive. Stories told by films or TV shows or even books do not possess the same quality of responsiveness or personal connection. Again, movies, TV, and books provide many excellent stories for children (and many damaging ones). But the stories and media of popular culture cannot be substituted for the storytelling relationships that children also need. One way I propose to analyze our genres of stories, then, is along the axis of relationships.

Other criteria for evaluating stories

During our research for Purple Moon, a company that I co-founded in 1996 to create interactive CD-ROMs and web materials for girls, we learned that personal relevance was an important key to creating stories that engage young people (Laurel 2001). As a mother of three teen girls, I have heard over and over again that they find no personal relevance in history, science, or math. Nowhere in the curriculum is it specified that a teacher has the responsibility of helping young people see why the subject matter is important or how it connects with their lives. This is the axis of personal relevance.

Many kinds of stories are told for strategic reasons. The desired outcomes are changes of opinion or belief, or action on the part of the hearer. Ancient and contemporary generals tell stories to motivate troops. Mothers Against Drunk Driving have stories that galvanize opposition to drunk driving. Political commercials tell stories that attempt to persuade voters to elect particular candidates. This is the axis of strategy and outcomes.

A fourth axis may be thought of as epistemological – that is, the truth value of a story and the way that we determine that value. I would like to look at several genres of stories present in our culture and to evaluate them along these four axes. Through this exercise we may think about the effects of existing kinds of stories and identify some kinds of stories that seem to be missing. Please remember that my evaluations are subjective, based primarily on my own observations as a student of culture, and not on any body of research. I would propose that such research might be very valuable in our ongoing efforts to tune our narrative universe.

Religious narratives

Starting with perhaps the most complex kind of story, I want to look at religious narratives. They are of great value to believers. They provide guidance and the comfort of faith. Religious narratives may also bring believers to face very difficult decisions as dictated by the ethics and morality of their faith. Religious stories are often deployed by priests, leaders, or believers in strategic and relational ways. Storytelling relationships are formed when people read scripture to address particular life issues for the hearer, deploying sacred text as a *story*, to be connected by the hearer to his or her life through narrative intelligence. But a demonstration of the trouble that can be caused by religious narratives seems always to be at hand, and none is more immediate than the current conflict in the Middle East.

By religious narratives, I mean canonical holy books or stories. In the big three patriarchal monotheistic religions – Islam, Judaism, and Christianity – the canonical narratives are easily identifiable, although each has been through centuries of reworking, intentional or otherwise. There are two ways that the faithful look at these texts. Those who look at them through the exercise of narrative intelligence will find story, history, metaphor, and other devices employed to make certain ethical and moral points clear. Those who look at them literally - fundamentalist believers – see them as having a truth value that is higher than all other stories. So, for example, the religious guarantees made by canonical narratives regarding the "ownership" of land claimed by both Palestinians and Jews have a higher truth value for the faithful than human narratives of suffering, historical narratives of occupation, or scientific narratives of genetic identity. Young people need to be challenged to question the literal truth of religious narrative and to explore how they might be applied in nuanced and open-ended ways.

How religious narratives rank on the four axes we've identified depends almost entirely on how the reader approaches the text. For the fundamentalist, we get one set of evaluations; we get a very different set for those who are believers but do not take the scriptures literally. Indeed, because of fundamentalism, religious stories tend to displace or forbid other forms of narrative from coming into play. Interestingly, the Catholic church stashes its non-canonical narratives in the lives of saints. This and its rituals hint at the pagan context in which the Church was born and established itself in medieval European and other pagan cultures.

In terms of relationship, both fundamentalists and the more open-minded faithful find high relationship value in religious stories – relationship with God, and relationship with the community of believers. Those fundamentalists with contrasting beliefs (e.g., Jew vs. Muslim) are placed in extremely negative relationships with one another. The non-fundamentalist faithful are more likely to allow other kinds of knowledge to provide a context for their exercise of faith. The relationships that religious narratives call for with those outside the faith depends on the particular narrative. For example, compassion plays a large role in most Christian faiths, but this aspect in action has historically been thwarted by the requirement to evangelize, bringing disease and despair to many native peoples.

Religious narratives have high personal relevance for both fundamentalist and non-fundamentalist believers. They also have high (negative) personal relevance to combatants in religious wars. For those outside the faith, personal relevance is a function of exposure, context, and pre-existing relationship.

Religious narratives are often invoked in calls to action. For example, Christians and Jews are exhorted to obey the ten commandments. Mormons and Jehovah's Witnesses are required to actively evangelize. In many faiths at different times and places, religious narratives are invoked to require war or persecution. Examples are the Crusades, the Inquisition, and the ongoing conflict between Jews and Palestinians in Israel.

Many young people in areas of conflict around the world are exposed only to religious narratives and calls to battle at home and at school. This sort of brainwashing was present at least as early as the Children's Crusades, and probably long, long before. The political or strategic narratives that are drummed into these children grow out of religious narratives. These narratives are likely to inherit the authority of true belief and preclude alternative ways of thinking about conflict or personal agency. An extremely difficult but enormously valuable goal would be to introduce these single-minded young people to other sorts of narratives that engage their narrative intelligence, critical thinking skills, and compassion or empathy.

Some religious institutions are more accepting of other kinds of stories to inform spiritual practice and, in the case of strategic narratives, broaden the search space for solutions to problems. The Episcopal and Unitarian churches in the United States are good examples. In these institutions, the fundamentalism of true belief is tempered by a view of religious texts as guidelines for ethics and morality rather than literal truths.

Religious narratives can also give birth to cultural narratives that provide structure and guidance for everyday life. They are derivative of religious belief without the constraints of fundamentalism. The cultural narratives of Judaism, for example, provide rituals and shape communities that are also able to admit of other sorts of narrative in thinking about the world. Strategically, concentrating on these derivative narratives – political, cultural, and ethical – tends to relax the stranglehold of true belief.

Folklore, spiritual and pagan narratives

Spiritual and pagan narratives and folklore stand in contrast to religious narratives in several ways. In such narratives, stories are understood more symbolically and metaphorically, giving broader participation to the reader. They are coded to refer to various aspects of life without necessarily being taken literally. Some examples are some of the spiritual narratives of Buddhism, Hinduism, and first-peoples' stories, and even some of the folklore that is now invoked around Christian holidays, such as the stories about Santa Claus, the Easter Bunny, or Halloween. Characteristically, such stories are identified by religious fundamentalists as "pagan." The term has been applied in a derogatory tone by monotheistic religion to groups that are polytheistic (e.g. Hindu), poetically theistic (e.g. Wicca), or non-theistic (e.g. "folk").

"Pagan" continues to be a troubling word for many people. In Germany and France, for example, the term today translates into something like "skinhead." I use the term "pagan" here in its original context, meaning literally, "of the land." Paganism characteristically involves highly articulated relationships with nature and tends to be quite specific in the value of local landscape. Celtic pagans associated spirits and stories with natural locations like wells, springs, and groves. Antique pagans (Greek and Roman) tended to associate their deities with qualities rather than locations, but sacred locations like Delphi and what is now Bath were also honored with local deities. One of the ways that the Romans pacified conquered peoples was to place statues of their local deities in Rome along with their own pantheon. In Shinto practice, still widely popular in Japan, deities and spiritual beings are most often tied to places. Confucian stories express a strictly ordered set of "right relations" and responsibilities among people, families, clans, and the state.

On a tour of Chaco Canyon years ago, I learned from my guide (an anthropologist) that the creation stories of the Anasazi and their progeny were probably told so that the landmarks in the story were visible from where the teller stood. The point, he mused, was probably that a child should be able to see where the world began and to think of where she stood as a place where spiritual power was strong. Here in the San Francisco Bay Area, many local native people place the creation at Mount Diablo or Mount Tamalpias, clearly visible before the days of pollution from almost anywhere in the region. Thousands of such examples exist all over the world.

Contemporary pagan practices in the United States come in many flavors (Adler 1979). Some, based primarily on Alistair Crowley and the Golden Dawn movement, have strict liturgies that call upon pagan gods of Celtic or Roman origin and stand close to "true belief." Others are more syncretic, drawing upon lore and ritual from a variety of sources, including Native American stories. Some invent stories and rituals related to their own locations and communities.

From Aesop's Fables (written by a Roman slave) to stories of Coyote, Bear, Anansi and other tricksters to stories of faeries and trolls, pagan spiritual stories and folklore often rely on non-human characters to stand in for aspects of human behavior. Other characters, like the Navajo Rainbow Woman, the Corn Maidens, or Grandmother Spider stand in for aspects of cosmology. In Celtic paganism, characters like Cernunnos and the triple goddess stand in for the mysteries and cycles of nature and work as symbolic cosmologies.

Many religious narratives, such as the injunction against cutting down an olive grove, probably have their roots in lore that addressed the exigencies of life long before they were incorporated by the present religion. Likewise, many Catholic saints such as Brigid and Sophia were likely appropriated from pre-Christian folklore.

Pagan stories attempt to capture aspects of humanity's relationship to the natural world, the solar cycles and the cycles of agriculture, and husbandry of the land and its creatures. They exercise our narrative intelligence to combine knowledge, intuition, and ethics with respect for the Mysteries of consciousness, creation, and purpose. Many spiritual tales, folklore, and pagan stories exhibit pro-environmental or Gaian qualities. Typically, such stories represent the world as a complete living being, including animals, humans, plants, and landscape into the same system. The tendency is to seek for right relationships with each and all.

Pagan practitioners of all sorts (including native people) have, of course, a high positive relationship with their communities and typically, with the Earth or a Great Spirit as well as with the characters in their narratives. Believers in patriarchal religion typically have a negative reaction to paganism because it comes into conflict with their faith, and because some groups – for example, fundamentalist Christians - may associate paganism with Satan through ignorance of actual pagan practice. Non-practitioners may dip into folklore to address the particular needs of a child through a storytelling relationship (e.g., telling the story of Little Red Riding Hood to a girl who feels powerless).

As with religion, for the practitioner of pagan spirituality personal relevance pervades everyday life. Non-fundamentalist believers and non-religious people tend to be more tolerant, capable of gleaning wisdom from pagan stories and folklore (as well as from religious texts) through narrative intelligence.

For the practitioners, pagan narratives and their implicit ethics inform action in the world and contemplation of one's relationship to nature. Some non-pagans may view strategy and outcomes as relatively more positive because the spiritual narratives and folklore often come to conclusions that are coherent with non-believers' political and ethical stances. For most, the strategy and outcomes of pagan practice are not well understood and, as such, the response tends to be negative.

In regards to truth value, it should be remembered that some native people and pagans treat their narratives as religion while others see them as guidelines for spiritual practice. Religious pagans may privilege their narratives in the same way that other religious fundamentalists do. By contrast, spiritual practitioners can typically also accept scientific and historical narratives as having equal truth value. For believers in other religions, pagan narratives are often seen as negative or even dangerous. Even so, the pagan revival or neo-pagan movement in the United States is strengthened by its relationship with nature. Many contemporary environmentalists and "folk" sense a measure of truth in the pagan attitude toward the natural world.

Spiritual, folk, and pagan stories are some of the vital narratives that we are in danger of losing. The trend toward globalization, the disappearance of diverse cultures and their stories, and the overriding narratives of religion, defense, development and economic "growth" put these forms at risk.

Scientific narratives

Like many spiritual narratives and pagan stories, scientific narratives can be characterized as dialogues with nature. The tools of science are based on rationality rather than story and ritual. Both scientists and pagans honor the natural world, and both understand the existence of Great Mysteries. The writings of such scientists as Newton, Einstein, and Feynman are aglow with wonder and joy.

In 1972, Joseph Campbell described the change in consciousness that resulted from our first view of the earth from space (Campbell 1986). Seeing our blue planet alone in the starry blackness, Campbell says, we suddenly understood that rather than coming *into* this world, we come *out* of it, or as Alan

Watts put it, "as a vine grapes, so the Earth peoples" (Watts 1999). The scientific understanding of Gaia as an organism was first explored Vladimir Vernadsky in 1944–1945 (Vernadsky 1997) and later articulated by James Lovelock (Lovelock 1979; Joseph 1990). Lovelock strove to construct a scientific narrative that demonstrated the Earth to be a whole living being, of which we are co-dependent parts. Taken together with Campbell and Watts' narrative inversions, Lovelock's work, extended by works of such scientists as Lyn Margulis and Dorion Sagan (Margulis & Sagan 1987), enhanced the impact of the Gaia hypothesis, both as a new line of scientific thinking and a new story for our relationship to our planet.

Scientific narratives always appear in situated contexts, often involving struggles to assert scientific findings in ways that do not offend religious institutions. Galileo, Copernicus, Darwin, and countless others have engaged in such struggles. In contrast, Newton thought that "Nature" was 'God's book" – and that by reading Nature's "laws", he could come closer to the mind – and purpose – of the Christian God (Fauvel 1988). Relatively few modern scientists hold this view; most simply assume the existence of an external world, eschewing solipsism, and want to know How It Works.

The theory of evolution provides an excellent example. From the Victorian era to the present day, the tendency of evolution to move toward greater complexity has been popularly misunderstood to mean that evolution moves toward greater perfection. Many believe evolution to be the unfolding according to God's Divine Plan, which culminates in Man.

But contemporary scientists like Lyn Margulis and Stephen Jay Gould (Gould 1989, 2002) see things differently. Their research suggests that evolution is neither the unfolding of a divine plan nor the inevitable march of sentience toward more and more spectacular manifestations. Devolution – or a movement toward less complexity – is also part of the process. And for humanity, like countless species that have disappeared over the millennia, extinction is a real possibility. Evolution is a process, not an outcome.

The particular sort of self-reflective consciousness that humans have seems unique on Earth. But as Margulis and Sagan argue, the most ancient microorganisms "invented" us through the process of evolution. Our consciousness is a manifestation, not exclusively of human brains, but of the entire biosphere, including the microcosmic life that surrounds and inhabits our own bodies. The biblical text in which God places Man in dominion over the natural world creates a attempts to separate humans from the rest of Nature. Gaians, environmentalists, and most scientists disagree.

Scientific narratives encounter resistance from many sources on religious grounds. Others see science as the tool of economic imperialism, as in, for example, the ongoing arguments about the patenting of seeds and the genetic engineering of foods. Resistance also comes from fear of change and the unknown. When science explains something that was heretofore mysterious and for which "folk" or religious explanations had been employed, many resist what they see as an assault on their beliefs. And yet, as science proceeds, the realm of Mystery is not reduced but redefined.

Many forget that the history of science shows that scientific findings and theories change as more is known and better tools are invented. An interesting difference between scientific and other kinds of narratives, as Karl Popper pointed out (Popper 1992), is that scientific narratives are disprovable. While the narrative of an almighty God or the existence of Faeries underground cannot be disproved, scientific findings can be – and often are. Disproving an antiquated scientific narrative (like geocentricism or the idea that the continents have always been where they are now) might be a very good way to engage students in the process and epistemology of science.

Scientists are explorers. Some go down 5,000 feet in bathyspheres to see things that have heretofore been unseen, and they come back and report to us. Others walk on the moon. My reason for placing scientific stories among the other forms I treat in this chapter is that I believe that they are powerful tools, and they present fundamentally different ways of knowing than other kinds of stories.

Perhaps the greatest obstacle to acceptance of scientific narratives is the language in which they are cast. One must understand scientific language to some degree to interpret the raw narratives of science – that is, the research papers, books, and specialized periodicals in which scientific narratives appear. Both scientific journalism and scientific educational materials attempt to make science accessible to the general public. Sometimes this works. But generally, one might say that journalistic reports often err on the side of sensationalism, while science textbooks fail to establish personal relevance. In some ways, interactive simulation as a form of scientific narrative can reduce or eliminate these pitfalls.

The response of many K-12 students to science teaching is, "why should I care?" Every teacher of science must be prepared to tell stories that answer this fundamental question. In educational and popular contexts, scientific information needs to be supported by particular attention to personal relevance. For example, in 2000–2001, my first-year graduate students in Media Design at Art Center College of Design devoted themselves to helping teens understand the human genome (see http://mdp.arcenter.edu/code23). Our first interviews with high school students indicated a fairly low level of knowledge and interest in the topic. But when we came to those same students with questions regarding policy, applications, and ethical questions that were associated with the science, we uncovered strong interest and a willingness of students to construct their own personal relevance. Our response was to create a transmedia system that incorporated such "hooks" for personal relevance into the scientific material.

Scientific narratives are a kind of vital narrative that faces many challenges in contemporary cultures. Its characteristic inaccessibility leads many to defer to business and government to determine the proper uses of science and the policies governing the directions and outcomes of scientific exploration. Religious communities and others based on a shared sense of oppression (as in some flavors of feminist separatism) see the assertion of truth value in science as an attack or an extension of the patriarchy. Increasingly, however, the practice of good citizenship relies upon scientific literacy. Science may be rated differently on our axes if we make certain changes in how it is represented (narrated) and taught.

In terms of students and the general public, the axis of relationship can be boosted by more scientists making narratives directed at this audience. Such scientists as Carl Sagan, Stephen Jay Gould and Stephen Hawking have created a "scientist as hero" image through their accessible, popular works, although they have had to endure derision for "pandering to the public" from the professional community. Certainly, more women scientists who achieve notoriety or take up the challenge of writing directly for a general audience from time to time would strengthen the appeal of science to girls and women.

Science writers like Margaret Wertheim and Matt Ridley also help to make scientific narrative accessible. Strategy and outcomes for science are high in visible areas like medicine, space exploration, and robotics. Personal relevance can be spread throughout science by beginning with highly relevant topics like, AIDS prevention, genetic engineering, or weather and climate. Popular culture can also play a part to address relevance through films like *Gattica* and good speculative fiction by authors like Greg Bear (e.g., *Darwin's Radio*), as well as popular science literature and news coverage.

Personal encounters with science, in environment such as the Exploratorium in San Francisco or through in-class experimentation and field work, may be the most powerful way to engage kids in science. Los Gatos High School uses this approach in their curriculum. Freshman science classes that are intended for the non-honors students are focused on environmental and space science –

down and dirty, up and out. These provide "hooks" of personal relevance for the reluctant science student.

At the John Woolman School, a Quaker institution, students do a service project related to science each year. These projects range from beach clean-ups to counting tortoises in the Mojave desert and working on ways to prevent dirt bikes from killing tortoises and disturbing their habitats. Science melds with action and activism in such projects and can prove quite rewarding and enlightening to students. The aspects of wonder, joy, and service can be incorporated into a "field curriculum" for science. The tools of handheld technology may greatly enhance the ability of students to do scientific work in the field.

Historical narratives

On May 10, 2002, *The New York Times* reported that "[at] a time when ancient cultures and conflicts are increasing American involvement around the world, American students show a poor command of history" (Schemo 2002). In my experience as a researcher and parent, it seems that history is taught primarily as "chronicle" – that is, as an exercise in remembering dates and names. It seems that causality, ethics, politics, personal stories and oral history are not adequately explored in many classrooms. The antithesis to the typical classroom experience is something like the Holocaust Museum, which contains many different kinds of narratives and artifacts to help visitors to see the whole picture.

One of the reasons that history often takes the form of chronicle is the desire to avoid highly colored points of view. History attempts to be "objective." Yet the actual events and experiences that history tries to represent are most often enormous conflicts in values, political or nationalist narratives, and points of view. Erasing these "subjective" differences removes the heart of history. Of course, personal stories, journals, and oral history are primary materials that can give emotional and political views into the complexities of history.

In the *Guides* project at Apple in 1990, Tim Oren, Abbe Don and I worked on a system for presenting different points of view on history in a computerized database (Laurel et al. 1990). We worked with the concept of "Guides" – characters who represented different points of view – as ways of navigating the information. In addition to encyclopedic content, the database was peppered with personal stories, and the "Guides" had stories of their own to tell. Our subject was Westward Movement, and the "Guides" consisted of a Native American, a pioneer woman, and a trader/trapper, each performed by real people. A particular Guide would suggest articles or stories that supported his or her point of view.

Oral histories and personal stories abound on the Web. Many university and high school communities as well as libraries have launched efforts to collect oral histories of U.S. veterans. The results are rich community websites in which living people reported their own experiences of the war. Oral histories can also be present in the classroom. The best history teacher I ever had was a man who had been a prisoner of war during World War II. His first statement to the class was, "everything is relative." By that he did not mean that all belief systems were equal, but rather that historical events are connected to all kinds of forces – political, economic, cultural, religious – and so the understanding of history deepens when these dependencies are understood.

Biographies and autobiographies provide emotionally rich sources for understanding how and why things happened in the past. Examples include *The First American: The Life and Times of Benjamin Franklin* by H. W. Brand, *and A Son of Thunder: Patrick Henry and the American Republic* by Henry Mayer, and *Marie Curie: A Life* by Susan Quinn.

We all know the old saw, "History is written by the winners." In that context alternative views take on even greater importance. When I was young, colonialism and Westward Movement were presented only from the perspective of the dominant cultures involved. It has only been through my personal studies that I have managed to understand a little about the "losers" in such events. If we are not teaching about the Trail of Tears or Wounded Knee or The Burning Times in our studies of history, then we are exhibiting a harmful bias. The fact that actions have (often unintended) consequences is a key to understanding how the world works.

The following suggestions may be rejected out of hand by the dyed-in-the-wool historian. Historical fiction offers many opportunities to see the complexity of history in an entertaining context. Examples include *Huckleberry Finn* by Samuel Clemens (Mark Twain), *The Little House* series by Laura Ingalls Wilder, *Fever 1793* by Laura Halse Anderson, *Gardens in the Dunes* by Leslie Marmon Silko, and *Gates of Fire: An Epic Novel of the Battle of Thermopylae* by Stephen Pressfield. Looking even farther out, speculative historical fiction stretches the mind and helps us see that history was not inevitable or pat. For those willing to try it, I would recommend the first three books of the Alvin Maker series by Orson Scott Card, in which an alternative America is depicted from the days of the revolution forward.

Of course, confusion about truth value is the danger of using such materials in the classroom. By suggesting these unorthodox ways of approaching history, I am also highlighting the importance of historical methods as part of any history curriculum. What are the kinds and classes of evidence? What is a primary source? What forces are at work "behind the scenes"? What are the key points of view? How is historical fiction different from historical "fact"? How do we assess truth value? Why should I learn about history?

By incorporating primary materials like personal stories and oral histories, students of history can experience a greater sense of relationship with the materials and their authors. Through empathy and the engagement of point of view, personal relevance can also be enhanced. Despite our best efforts, the usefulness of history as informing the exercise of citizenship (strategy and outcomes) may not increase greatly without changes in other disciplines (e.g., journalism and political narratives). By admitting of a diversity of points of view while also reporting the known facts of history, its truth value is actually enhanced. Teens in particular perceive the examination of multiple points of view as more "fair" than their erasure.

At the end of the day, young citizens should be able to question pop-culture representations of history as well as journalistic reports. A quick review of how "war movies" have changed from the days of "With a Wing and a Prayer" to "Blackhawk Down" would provide an excellent lesson in this regard. It is the citizen (or student) who must judge the import and truth value of representations of history and decide whether and how to use it in work, personal beliefs or philosophy, and the exercise of citizenship.

Journalistic narratives

Recently, the *Wall Street Journal* changed its look to incorporate more color, giving the stodgy old paper something of a pop-culture facelift. But I suspect that the authority or truth value of the paper has been undermined for its most loyal readers. The designs of such venerable papers as *Wall Street Journal* and the *New York Times* have become emblematic of a certain truth value and journalistic standards. By contrast, *USA Today*, close to becoming the most widely read paper in the world, has always had a design style that borrowed from tabloids. Its "news" colorfully skims the top layer of "what's going on." It reports events, emotional moments, and pop-cultural phenomena. *USA Today* is not deep. Rather than journalistic exploration, it provides the reader with a snapshot of the state of the world – one that is highly colored, not only by its style, but by its pro-American bias.

The *Jerusalem Post*, the *Jordan Times* or even the *London Times* are not on the average person's reading list, although all are readily available on the Web. Here again we come to the problem of objectivity vs. point of view. Just as a person does not perceive that he or she has an accent but thinks that others do, American journalism has its own invisible biases. *Everything has a point of view*. Reading newspapers like the *San Jose Mercury* or the *Chicago Tribune*, most U.S. citizens are fooled into thinking that an American point of view is unbiased and that most foreign newspapers, especially those of the Arabic world, are "propaganda." Whether and why that may be true can only be learned by reading them. Through this exercise, we may identify the points of view embedded in our own news sources.

Television and radio news are problematic in somewhat different ways. As a recovering television addict (clean for 10 years, except for *Star Trek*), I rely on the radio for news. A friend has referred to my radio news habits as "NPR Poisoning." Yet of all news sources available to me, I find NPR (including their programs from the BBC) to be the broadest and boldest news sources in their diverse programming and treatment of point of view. Wishing not to be distracted by the spectacle, I choose radio as a more thought-provoking medium for news. During the 9/11 attack, I listened to the radio. I saw only one image of the World Trade Center bombing quite by accident in a hotel lobby. Later, I selected images from the Web to contemplate without the jabber of an excited reporter cluttering up my response.

A regular news source for me is "The News Hour" with Jim Lehrer, which is broadcast on public radio. Parenthetically, PBS seems unaware that many listen to the program on the radio, as their big sponsorship ads often do not speak the name of the sponsor. One night in a hotel room I decided to watch "The News Hour" on the television. I was appalled by the garish red-white-and-blue set and the overall visual design of the program. It was tacky. I don't think that "The News Hour" is tacky, but I do think that they are working in an arena where entertainment value is increasingly important. They must compete with the text-and-image extravaganzas of CNN and ABC. The recent flap over the potential replacement of Ted Koppel with David Letterman provides the best demonstration of the devaluation of news and the intrusion of entertainment into the journalistic sphere.

The ethics of journalism call for factual reportage and thoughtful analysis. Yet reporters are human beings, responding to the events they are covering. Typically, they withhold personal responses and judgments from their reportage. This is one of the ways in which NPR and other independent news sources are distinct from mainstream news. Reporters often include their personal experiences or responses, making it clear that it is a personal commentary.

The personal dimension is highly desirable; what we want from field reporters is a picture, not only of the events they are witnessing, but also something about the experience of being there, which includes personal response. Field journalists are heroes. They are sometimes killed in action. They go to dangerous places to help us find out what is going on "on the ground" and to give us context for understanding events. As we honor them and their experiences, so we also honor their profession.

Typical publishers concentrate primarily on sales and advertising revenue, so they are extremely concerned with strategy and outcomes in terms of business. Their relationship to journalists is typically less meaningful than their business concerns (although the occasional Pulitzer boosts business). Often, surrogate personal relevance is used as a sales tool, reporting stories of sensational crimes or emotionally charged political conflicts. The truth value of American journalism is currently being eroded by the move toward entertainment and the unspoken bias toward American policy.

The general public typically feels a weak relationship with journalists. Much of the news seems irrelevant to many people. The degradation of journalism by sensationalism and entertainment value tends to erode truth value and rarely provides the general public with actionable information (Postman 1985). Children and teens tend to find the news to be part of an adult world that they are not ready to enter. Teens think about aspects of the news with a critical eye and often with an activist or utopian bias – characteristics that are extremely useful in shaping our future.

The lack of young voices in the news is one more way in which contemporary American journalism misses the mark. NPR has sponsored "Youth Radio," a weekly news show that is entirely produced by young journalists. Sadly, there are few other sources of news from the world of the young. There is some activity of this sort on the Web – for example, www.kidnews.com (news created by young children around the world) and www.wiretap.com (teen-created news). Student newspapers can be a vibrant source of information about the interests and concerns of teens. The student paper at Los Gatos High School, for example, devoted two full pages to deep and thoughtful exploration of the causes and effects of the 9/11 attack. The paper delves into religion, popular culture, politics, and local news. School newspapers are great sources of information about how teens view the world and what is important to them. Teens may not read the *Times*, watch news on TV, or listen to news on the radio, but they are interested in news and they have a lot to say. By listening and creating more youth news outlets, we grow closer to our young citizens.

Political, patriotic and geopolitical narratives

As the decline in our participation shows, Americans are displaying growing cynicism about the political process. The combination of campaign financing and negative campaigning support the jaundiced view that a citizen's vote makes little difference. The debacle of the last presidential election lent credence to that opinion. Of the cynical Americans, young people have the lowest voter participation. Although they enjoy the spectacle of films and television, they are wise to the spectacle of politics. Interestingly, while youth voter participation is at an all-time low, volunteerism in the teen demographic is on the rise. This suggest that teens are indeed engaged in making the world a better place, but they don't necessarily trust government to do it.

The 9/11 attacks brought on a wave of patriotism. Flags are everywhere, as are patriotic bumper stickers, and tight-lipped news display the new patriotism. "United We Stand" is declared on cars, storefronts, and billboards. But there is something ominous about this phrase. The unspoken rejoinder is "Divided We Fall." The hidden message is that a true patriot is unquestioning during times of war. To question military policy or to attempt to articulate alternatives to the solution strategies in play is unpatriotic and dangerous. But the lack of public discourse on policy runs completely counter to the spirit of the original United States. The backbone of a democratic republic is the freedom – even duty – to engage in discourse and disagreement. An amended slogan might be, "united we stand around; divided we might have a useful conversation."

In my view, the narrative of patriotism has been over-generalized so that it interferes with the freedoms that our Constitution set out to guarantee. Certainly, patriotism is a good thing. Our judges and governmental officers and military personnel take an oath to "preserve and protect the constitution of the United States against all enemies, foreign or domestic." Notice that this is not the same thing as agreeing to support "my country right or wrong." Patriotism is, or should be, a commitment to be an active and responsible citizen, not to be a passive supporter of the political flavor of the day.

The Civil Rights movement, for example, was intensely patriotic, as it drew its strength from a Constitutional principle. Institutionalized racial discrimination grossly undermines the spirit of the American democratic republic. Yet in the 60s, Civil Rights workers were labeled "unpatriotic" and were pushed around by those who believed that the status quo was equivalent to the "right way" to do things. One of the remarkable strengths of the Constitution is its humanistic spirit. The founders foresaw a country that would undergo much growth and change. The Constitution was a guide for accomplishing change in a democratic way. It detailed both the rights and the responsibilities of citizens. It even set up ways for citizens to amend it. In our day, the responsibilities of citizenship are in danger of fading from consciousness.

A good patriotic narrative makes citizenship its centerpiece. If we have a corrupt campaign financing process, only active voters can change it. If we have low voter turnout, only citizens can change it. If we have domestic or global policies that deprive people of the rights that our Constitution proclaims as belonging to every person, only active participants in our democratic republic can change them. If political spectacle and sloganism displace meaningful discourse, only the people can change it. We need patriotic narratives that can revitalize citizenship. Such narratives may also be understood as actionable narratives. In the realm of citizenship, the actions to be promoted are engaging in discourse, becoming well-informed, communicating with your representatives, voting, and exercising the right to peacefully assemble.

Other actionable narratives that are much needed in our day would help people understand what actions they can take to improve their world. These include activism, volunteer opportunities, ways to become involved in their children's educations, honesty-based techniques for persuading others, and ways to advance unpopular causes without falling into the trap of self-marginalization. The narrative of education, for example, revolves around explanation of the status quo. "This is history. This is science. This is mathematics." To transform this view of education into an actionable narrative, a 180° turn is needed: the focus must be shifted from what is known to what is to be discovered, and what can be done with present and future knowledge. Life is change. The best education helps a student to survive, manage, and direct change.

Finally, we come to the category of geopolitical narratives. These are the "grand narratives" that percolate through all our stories and actions. Among the most damaging, in my view, is the narrative that links business and economic growth to prosperity. We demand "growth" from our economy and our business institutions every year. Futurist and writer Paul Hawken sees growth in a different light – instead of prosperity for the whole citizenry, he believes, growth typically transforms natural capital into wealth for the usual suspects – a wealthy elite (Hawken et al. 1999). If a steady stream of disastrous oil spills or clear-cutting the Pacific Northwest is required for growth, then the narrative of business, economic growth, and prosperity wins.

This narrative is challenged from several directions – by those who see the wisdom of sustainability, by those who prefer harmony to growth at the expense of the environment, by those whose cultures and livelihoods are displaced by "growth." Hawken observes that we have lots of people and not enough work, and that it therefore might be better to de-industrialize some economies. Contrariwise, the invention of the triode by Lee de Forest created new wealth without consuming natural capital; therefore, knowledge may create new economies and new wealth. How that wealth becomes general prosperity remains problematic.

People sometimes invoke the building of the railroads or the industrial revolution as examples of enterprise making life better for everyone. But today, we are much closer to the edge of environmental collapse. Different criteria should be applied. Overfishing, excessive logging and unsound logging practices, overpopulation, air pollution, water shortages, industrial and nuclear waste, and global warming are all parts of the picture. Even when faced with compelling scientific evidence, the business world would like to call "global warming" a "hypothesis" until the very last minute. The truth is that the very last minute may have already arrived. A global narrative is coalescing around the idea that husbandry of the Earth so that it remains a good home for humanity is more important than growth and development.

At the same time, we have much to teach the rest of the world, if we can do so with respect and follow-through. "Cultural imperialism" is anathema to the left, while cultural relativism is in vogue. Some people worry about the corrosive effects of American media (and American economic development) on the cultures of other peoples, but Hollywood earns most of its dollars from international distribution of American films, and so these concerns are typically overridden by the narrative of business and prosperity. Our solution to governmental and cultural practices in other countries with which we disagree is to ignore them unless they have economic implications for America. From female excision to the state-supported incursion of processed foods into the agricultural culture of India, we turn a blind eye to those problems that don't apparently damage our prosperity.

But if we think of ourselves as humans living together on a single planet, a different narrative comes to light. The alternative looks a lot like cultural imperialism. It involves exporting the form of the democratic republic to other countries and, increasingly, working with the United Nations to deal with global issues. The form of the democratic republic slots into tribal cultures particularly well, and is most certainly preferable to military governments or dictatorships in terms of public well-being and prospects for peace. Critics of this view say that we tried exporting the democratic republic to Russia and it failed. The typical explanation is that the Russian people have no experience with democracy and are inclined culturally to prefer other forms of government.

Yet if we believe that the form of the democratic republic is the best way to guarantee freedom and participation for citizens, then we are obligated to try to export it, as we did to Japan and Germany. The difference between the Russian experience and the narrative of democracy is *follow-through*. Teaching, learning, and aid make the difference between a half-hearted declaration that ours is the best way and a full conversion of less satisfactory forms of government to the ideal of a democratic republic. And when we have accomplished that, we may find that we are all global citizens, and that the freedoms inherent in the definition of democracy are adequate to accommodate – and even to celebrate - cultural differences.

At the end of the day, I believe that the best geopolitical narrative is one that honors individuals, cultures, the idea of the democratic republic, and – most of all – sustaining and protecting the Earth. Our government backed out of the Kyoto protocols. They backed out of support for various UN health initiatives. They opted for policies that support the old game of nationalism, growth, and exploitation. They are sneaking around the Constitution to surreptitiously weaken the separation of church and state. They routinely privilege business over the health of the planet and the rights of other human beings to determine their own livelihoods. To construct a new story, I think we need to give our narrative intelligence the strongest possible workout. We need to absorb many other kinds of narratives and understand the kinds of truths they have for us. May our next geopolitical narrative be devoted to establishing freedom, harmony, balance, and sustainability. If we can construct the story, we can find a way to make it so.

References

Adler, Margot (1979). *Drawing down the moon*. New York: Penguin/Arkana.

Campbell, Joseph (1986). *The inner reaches of outer space.* New York: Harper and Row.

Fauvel, John, Raymond Flood, Michael Shortland, & Robin Wilson, (Eds.) (1988). *Let Newton be! A new perspective on his life and works*. New York: Oxford University Press.

Gould, Stephen Jay (2002). *The structure of evolutionary theory*. Cambridge, MA: The Belknap Press of Harvard University Press.

Gould, Stephen Jay (1989). Wonderful life: The Burgess Shale and the nature of history. New York: W. W. Norton.

Hawken, Paul, Amory Lovins, & L. Hunter Lovins (1999). *Natural capitalism: Creating the next industrial revolution*. Boston: Little, Brown and Company.
Joseph, Lawrence E. (1990). *Gaia: The growth of an idea.* New York: St. Martin's Press.
Laurel, Brenda (2001). *Utopian entrepreneur*. Cambridge: MIT Press.
Laurel, Brenda, Tim Oren, & Abbe Don (1990). Issues in multimedia interface design: Media integration and interface agents. *Proceedings of CHI '90* (ACM SIGCHI). Reprinted in *Multimedia interface design*, M. Blattner and R. Dannenerg (Eds.), ACM Press/Addison-Wesley, 1992.
Lovelock, James (1979). *Gaia: A new look at life on earth*. New York: Oxford University Press.
Margulis, Lynn & Dorion Sagan (1987). *Microcosmos: Four billion years of microbial evolution.* New York: Simon and Shuster.
Popper, Karl (1992). *Logic of scientific discoveries*. New York: Routledge.
Postman, Neil (1985). *Amusing Oarselves to Death: Public Discourse in the Age of Show Business*. New York: Viking.
Schemo, Diane (2002). Students, especially 12th graders, do poorly on history tests. *The New York times*, May 10, 2002.
Vernadsky, Vladimir I. (1997). *The biosphere: Complete annotated edition*. New York: Nevraumont Publishing Co.
Watts, Alan (1999). *Tao of philosophy*. Public Lectures, Electronic University, Mill Valley, 1963–1969. Paperback edition, Charles E. Tuttle Co.

Chapter 6

We are what we tell

Designing narrative environments for children

Marina Umaschi Bers
TUFTS University, Medford

Introduction

"Who am I?" "What are the values I hold and cherish?" "Which is my place in the world?" Young people frequently ask these questions regarding identity and values. And they use different kinds of narratives to answer them: personal stories, popular tales, cultural myths. Computational systems can support young people to tell and listen to stories in order to learn about themselves and others. I coined the term *identity construction environments* to refer to technological tools specifically designed to allow children to learn about different aspects of the self through storytelling and computation. While their fundamental mission is to help young people construct a well-grounded sense of self by engaging in the exploration of personal and moral values, they also serve other educational goals. On the one hand, they support the cultivation of narrative intelligence by engaging in storytelling. On the other hand, they foster the development of computational intelligence by providing an opportunity to explore the power of design and programming.

In this chapter I will first present the concept of identity construction environments. Then I will describe three prototypes that I designed and tested with children and teenagers: the SAGE authoring environment, the web-based Kaleidostories and the 3D graphical multi-user environment Zora. I will briefly describe the technologies, the design principles and the use of each of these environments by young people in the real world. I will also share lessons learned with each one.

Identity construction environments

I coined the term *identity construction environments* to refer to technological tools purposefully designed to afford opportunities for exploring identity and engaging in reflection and discussion about personal and moral values. Given this definition, six design principles distinguish them from other technological tools for learning:

1. They are *purposefully* designed to help young people learn about their identity, particularly personal and moral values.
2. They are designed upon a theoretical model that understands identity as a complex and dynamic construction composed by conflicting values.
3. They afford opportunities for learners to engage in the design and creation of computational objects. These objects represent aspects of the self and can be created and programmed in a playful way.
4. They integrate the use of objects and narratives. For example, computational objects are described with narrative attributes and storytelling behaviors.
5. Their design is informed by the constructionist learning theory (Papert 1980), theories of identity formation (Erikson 1950) and Kohlberg's theory of moral development in a just community (Kohlberg 1982).
6. They support the creation and participation in a community. No sense of self develops in a social vacuum.

In the same spirit as other constructionist tools for learning, identity construction environments engage young people in a hands-on learning experience. They support the construction of knowledge by building personally meaningful artifacts that behave in the world. Identity construction environments are designed following the "construction kit" metaphor: an environment with a set of parts to be assembled and connected together. For example, structural or mechanical construction kits, such as LEGO, have parts from the world of engineering (e.g. bricks, gears, pulleys). Through the exercise of assembling them, young people can develop knowledge about mechanics. Other types of kits, such as computational construction kits (Resnick et al. 1996) are composed of parts from both the world of engineering and the world of computation (e.g. feedback loops, variables, control structures). For example Lego-Logo supports explorations of powerful engineering, robotics, computational and mathematical ideas. In the same spirit as these construction kits, identity construction environments provide dynamic building blocks focusing on identity and personal and moral values. These building blocks represent different aspects of the self and can be arranged and put together in a playful way.

Learners can design and program these building blocks with storytelling attributes and behaviors, thus exploiting the power of narrative (Polkinghorne 1988). Narrative is a fundamental component of identity construction environments. It serves a descriptive function because it supports the finding of coherence between the diverse personal experiences, thus allowing the telling of a coherent life story (Linde 1993). It also serves a constructive function because it enables, through external dramatizations, to play out diverse aspects of the self in "what if" situations. Both the descriptive and constructive functions of narrative are important in the process of identity construction and are supported by identity construction environments.

SAGE: Storytelling agent generation environment

SAGE is an authoring environment for children to create their own wise storytellers to interact with by telling and listening to stories. Children can engage with SAGE in two modes: 1) by choosing a wise storyteller from a library of already existing characters and sharing with him or her what is going on in their lives. The sage storyteller "listens" and then offers a relevant tale in response, and 2) by designing their own sages and programming the conversational interaction between storyteller and potential users as well as creating the database of inspirational stories offered by the storyteller in response to user's problems (Bers & Cassell 1998).

The LISP-based SAGE architecture has three parts:

- **Computation module**: in charge of parsing the user's story to extract nouns and verbs, expanding these keywords through WordNet, a hierarchical semantic lexical reference system (Fellbaum 1998), and performing a match between the user's personal story and an inspirational story in the database. This module does not have any knowledge about story grammar; it only deals with augmented keywords in order to find the story that is most like the user's story – which deals with the same themes.
- **Authoring language**: serves to design the personality of the interactive characters, the conversational flow between user and storyteller and the database of stories offered by the system in response to the user's story. A graphical user interface allows children to create conversational flows in a way similar to that by which they engage in pretend role play games, by

Figure 1. The interactive stuffed storyteller assistant.

planning *who* will say *what*, and *when*, while arranging graphical objects. In the database of stories users can write or record the inspirational stories that will be offered by the sage. They can index them with personal and moral values that they consider good descriptors of the main story points.

– **Interface**: is the layer through which a user communicates with the system, composed by the computer screen and a programmable interactive soft toy that behaves as the storyteller's assistant. The toy moves its ears, shakes its body and blinks its eyes to indicate attentiveness. These body movements can also be programmed by the children (see Figure 1).

Becoming sage by designing a SAGE

We conducted several studies to examine children's interactions with the sage storytellers, and their ability to build their own meaningful characters. Our earliest research showed that children engaged deeply with sages that we had designed, such as a Hasidic Rabbi and a Buddhist scholar (Umaschi 1996). When interacting with them, children revealed aspects of their inner lives and the problems they face everyday. For example, below is an extract from a conversation log between a ten year old and the Hassidic Rabbi:

SAGE: You know Abraham, a story is more than a story...a story has power. Nu, what is your story? what is in your mind ?
USER: Once I was learning math and I could never understand it.
SAGE: Oy vey... go on, Abraham.
USER: I kept on trying, but no matter how much I studied I could not get it.
SAGE: I see Abraham ... what else do you have to share with this old rabbi?
USER: I have the same problem with English, Hebrew and all other subjects.
SAGE: Well, your situation reminds me of a story. Remember that there is power on the stories. The Gerer Rabbi said: "Exile contains redemption within itself, as seed contains the fruit. Right work and real diligence will bring out the hidden reward. Abraham, do you see the point of this story related to your own story?
USER: yes, so now I see that I just have to keep on trying no matter what.
SAGE: I hope it was helpful. "Zie gezunt", be well. May the Lord, Source of Life, bless you Abraham with peace, and may you find your inner peace. Amen.

In later studies we conducted research with fourth and fifth graders who built their own meaningful storytellers, such as Mother Nature, Shaquille O'Neil and the Big Orange Fox (Bers & Cassell 1998). In building a sage, children designed that person to whom they wished they could turn with their problems. They also played with different notions of self, by creating or imitating the narrative voices they wanted or needed to hear. Children created storytellers as projections of fears, feelings, interests, and role models.

Using SAGE they were able to design and model abstract structures needed and used in both programming and storytelling. In order to do so, they created, manipulated and put together narrative building blocks such as speaker turns, scripts and communicational actions. This allowed them to explore the conversational storytelling genre while observing other people's interactions with their storytellers.

In later research, we explored how SAGE could be used for therapeutic purposes with chronically ill children who are particularly in need of telling the stories of their lives. A pilot study was conducted in the Cardiology Unit of the Children's Hospital in Boston (Bers et al. 1998). Young cardiac patients used the SAGE environment to tell personal stories and created interactive characters, such as Mrs. Needle or Mr. Tape, as a way of coping with cardiac illness, hospitalizations, and invasive medical procedures.

Lessons learned

The research done with SAGE showed that children in very different situations used this identity construction environment to explore aspects of their inner lives through the creation of stories and storytellers. While expressing their feelings by telling personal stories and listening to inspirational stories, children learned about themselves. While designing conversational interactions in which other people could participate, they also learned about others. SAGE's design engaged young people in learning about identity, as well as developing narrative and computational intelligence. However, it did not support further explorations of how a community shares narratives, nor how identity is constructed in a social context. Neither were kids engaged in exploring values through concrete actions. They only used them to label and categorize stories. Since all of these elements are important to develop a well-grounded sense of self, I decided to design a second generation of identity construction environments.

Kaleidostories: A web-based narrative experience

Kaleidostories is a web-based identity construction environment that focuses on the use of narrative to explore role models and values in the context of an on-line community. Every participant in the community is represented by a geometrical figure in the kaleidoscope displayed in the right top corner of the screen. The figure's color and shape changes according to how many role models and values are shared between the logged user and the other participants. The kaleidoscope allows visualizing community patterns of shared role models and values (see Figure 2). Kaleidostories runs in an NT Java-based Webserver and it is implemented in Java. Data entered by the children is stored and recovered from a database using Java servlets. The patterns visualized on the kaleidoscope are generated at run-time by queries to the database (Bers 1998).

The system guides users in the creation of a personal on-line portrait with narratives about the present – "who am I?" – and narratives about the future – "Who or what do I want to become?" It also guides them in the creation of role models. Children can either choose their role models from a library or create their own and add them to the already existing list. The system asks them to write stories involving role model's biographical information as well as narratives of personal identification, such as "why did I choose this person as my role model?" and "what are the values that I admire about him or her?" The

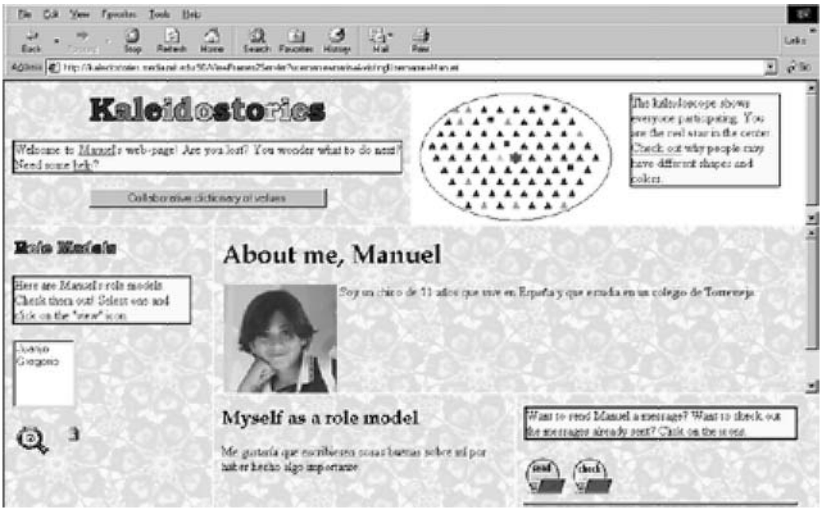

Figure 2. The Kaleidostories website.

system also invites users to link role models' stories with particular values (such as friendship and justice) and to define those values in a collaborative values dictionary. This dictionary has all the values that the Kaleidostories community holds as a group, as well as the personal definitions that each individual creates to ground those abstract concepts to concrete situations. At any point, children can look at the kaleidoscope, browse the creations of other participants and engage in a-synchronous communication.

Sharing stories across the world

I conducted two on-line pilot studies with Kaleidostories. First, I did a study with three bilingual sites (Spanish/English) in different parts of the world: a small bilingual class in a Cambridge public high school, an elementary school class in Torrevieja, Spain and a youth group from a Jewish Sunday school in Buenos Aires, Argentina. Second, I conducted a pilot study with only Spanish speaking sites: the same elementary school class in Spain, two rural schools with Internet connection in Colombia and a high school class in Argentina.

During both studies every local teacher decided to use the tool in a different way and with different goals. For example, the teacher in Cambridge integrated Kaleidostories into her "Spanish Literature and issues of adolescence" class. She focused on writing stories about role models in Spanish, a language which most of her students spoke very well but were not very comfortable writing. In Spain the teacher decided to focus on the values dictionary and did in-depth work with his class writing stories to express their most cherished values. The Argentinean high school teacher who participated in the second study taught psychology and sociology. She used Kaleidostories as a way to help her students ground their theoretical readings in a concrete personal experience. For example, as a final assignment, she requested her students to write a paper reporting how the on-line community evolved over time and what kinds of narratives of personal and social identity emerged.

In both studies children added their own personal role models to the library and very rarely used already existing ones. Sports players, popular singers and movie stars as well as family members, friends and well-known figures such as Mother Teresa of Calcutta were chosen as role models. Children also added their own values and definitions to the collaborative values dictionary. Friendship and love resulted, in both studies, as the most popular values with the major number of definitions. Some definitions were simple, such as "*Friendship is easy: two people meet and they become friends*" and others were more complex: "*They say that friendship is to be friends and that is it. But, the true friendship is to be faithful to your friends, in the good and the bad, and never betray them. In my opinion, true friendship is too demanding to be able to achieve it*". While reading the diverse definitions kids engaged in discussions about the different meanings that a same value might have for different people.

Kaleidostories provided a framework that encouraged reading and writing as fundamental tools for communicating with others. It helped bilingual kids to find a meaningful activity through which to express themselves in writing to an engaged audience of peers. Juan's story is a good example. Juan is a 17-year-old recent immigrant who did not yet speak English and who had severe problems writing in Spanish. He was a tough kid with discipline problems in school. With a lot of effort and many spelling mistakes Juan became very involved with Kaleidostories. It presented for him the challenge of learning to use computers and, at the same time, allowed him to open up about aspects of his inner life. Juan's kaleidoscope had lots of different colors and geometrical shapes representing the role models and values that he shared with others. As Juan became popular in Kaleidostories and exchanged more e-mails with kids across the world he started to care, for the first time, about his spelling. It was a barrier to being understood. He asked the teacher and his classmates to correct his writing. As time went by he started writing more complex stories and he converted into an expert user of the computer. Juan's development of narrative and computational intelligence helped him become a more confident learner and gain self-esteem. Juan's case shows how Kaleidostories fostered a social context that helped a teen change his sense of identity.

Lessons learned

Kaleidostories allowed young people to explore aspects of themselves such as role models and values through sharing stories. At the same time it provided a forum to share differences and similarities with others living in different parts of the world but sharing a language. As an identity construction environment it combined the power of computation to visualize community patterns and the power of narrative to express feelings and thoughts. However, Kaleidostories lacked the capability to include direct communication through real-time chat. It also lacked the flexibility to express a more complex sense of self. One of the most successful design features of Kaleidostories was the collaborative values dictionary. However, it only supported the expression of values as narratives and did not enable those values to be put to test through behaviors in the on-line community.

In order to facilitate the passage from moral knowledge to moral action, Kaleidostories' design was not enough. Although there was a sense of community, represented by the patterns of the kaleidoscope, the tools for self-organization and forums for discussion were missing. This is essential to form a responsible and just community (Kohlberg 1982) in which values are developed not only as narratives but also through action. Kaleidostories did not exploit the full potential of computation, as it did of storytelling. It limited computation to networking and visualization. On a different note, Kaleidostories was not fun enough to engage children to use it on their own for a long period of time. But issues of identity and values need a long time to be explored in depth. A big effort from the teachers was needed in order to keep students on track. The experiences with SAGE and Kaleidostories served me in designing a third generation of identity construction environment.

Zora: A narrative-based virtual world

Zora is a 3D graphical multi-user environment that provides the tools for young users to create a virtual city. As with the other identity construction environments mentioned before, Zora's design supports the exploration of identity and values through storytelling and programming. The name Zora was inspired by one of the cities that Italo Calvino describes in his book "Invisible

Figure 3. A personal home designed by a thirteen-year-old.

Cities": "*This city is like a honeycomb in whose cells each of us can place the things we want to remember... So the world's most wise people are those who know Zora.*" (Calvino 1972).

Users are graphically represented by avatars and can communicate via text or gestures. They can navigate around the 3D virtual city, converse with others in real-time and construct the city's private and public spaces: personal homes, community centers and temples. Temples are shared public spaces representing cultural traditions or interests. Users can populate these virtual spaces with computational objects and interactive characters representing role models and anti role models, which can be programmed to engage in storytelling interactions with other users (see Figure 3). Both personal homes and temples become autotopographies or spatial representations of identity composed by artifacts symbolizing intangible aspects of the self (Gonzalez 1995).

Zora is an object-oriented environment, meaning that users can make new objects by cloning existing ones and inheriting its attributes. Users can create the following attributes for their objects: (1) presentation attributes, *graphical appearance* and *motion*; (2) administration attributes, *ownership*, which determines who owns the object and therefore can edit it, and *permissions*, which set if the object can be cloned; and (3) narrative-based attributes, *textual description*, *stories*, *values* and *conversations*. Zora is implemented using

Microsoft's Virtual Worlds research platform, a software development kit for building distributed multi-user environments (Virtual Worlds Group).

There is a growing amount of work on virtual worlds (Turkle 1995). However, while most of the research looks at how community develops as such, Zora looks at how personal identity develops in the context of a community. The research is aimed at helping young people understand and affect the ways in which identity and values are constructed in the real world, as well as online. In the same spirit as other constructionist virtual communities such as the text-based MOOSE Crossing (Bruckman 1994) and the 2D Pet Park (De Bonte 1996), kids can program behaviors for their own creations. But in Zora, programming is limited to storytelling behaviors. For example, they can describe the underlying turn-taking rules between user and character as well as define the stories to be told in response to certain input. Like in the psychological novel, the engine of action is placed in the richness of the created characters and the resulting interactions rather than in the plot. As in Kaleidostories, users can create a collaborative values dictionary. But in Zora not only can they define its values but also put them to test through their actions in the community.

Kids designing their own virtual cities

I conducted two pilot studies in which young people used Zora: an intensive summer camp held at the Media Lab with a multicultural group of teenagers, and a five-month study with young patients in the Dialysis Unit at Boston Children's Hospital.

Despite their diversity in background and context, I chose these populations because both share a need and desire to explore identity issues. The first study explored how Zora could help young people from diverse cultural backgrounds to explore their identity while developing a sense of personal and moral values (Bers 2001). The second study focused on feasibility and safety of using the Zora virtual environment with young patients facing hemodialysis in a hospital setting. This includes the analysis of Zora's impact on children's understanding of their illness, and its potential to facilitate mutual patient support and interaction (Bers et al. 2001).

In both studies, participants built and inhabited a virtual city with personal homes and public spaces. For example, the summer camp participants built the Salsa and Merengue temple and the French Chateaux, while the dialysis patients built the Temple of Feeling Better and the Renal Rap room (see Figure 4).

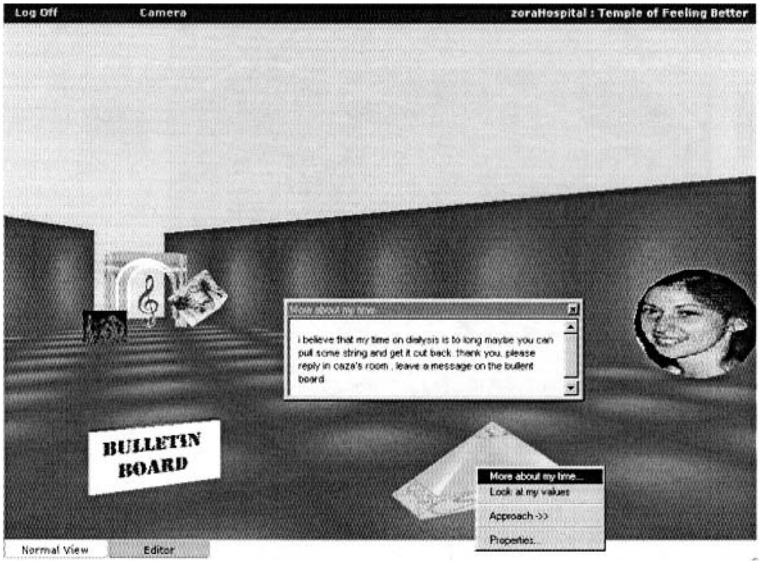

Figure 4. The temple of feeling better.

Zora engaged young people in the design of spaces and dynamic artifacts representing aspects of their complex selves. Kids used the Zora environment to explore personal identity and values in a community self-organized by democratic principles. For example, they held weekly meetings in the virtual City Hall and experimented with different on-line voting systems. As time went by they realized the need of laws to organize the social life of the virtual community. They agreed on basic laws such as "*no putting things in people's personal rooms*", "*set the properties of the objects placed in public spaces so others can use them if they like*", "*fess up to what you do*", and "*there will be no jail*".

By providing a social context for the development of self-government, Zora engaged young people in the creation of a participatory community in which values were discussed and put to test through behaviors. As time went by kids started to drop "cases" they wanted to talk about in the City Hall. Cases are special types of objects representing events or circumstances to be discussed and agreed upon. They require community members to take action to resolve them. During the summer camp experience most of the cases dealt with setting up the social organization of the virtual city. Examples of those cases are "*I think that people should not change or put things in other peoples rooms. Unless they have permission.*" or "*Anyone should be able to drop anything anywhere, but with a consequence*". Other cases were to discuss and raise awareness about controversial topics such as death penalty and current hate crimes reported in the news. During the experience at the hospital, cases raised awareness and discussion about particular situations regarding individual treatments and served patients as a way to voice their opinions and engage in informal interactions with doctors. For example, a seventeen-year-old boy left the following case in The Temple of Feeling Better: *"I believe that my time on dialysis is too long. Maybe you can pull some string and get it cut back. Thank you. Please reply in caza's room. Leave a message on the bulletin board"*. As a result, the patient engaged in a long on-line conversation with one of the doctors participating in the study.

During both pilot studies, kids engaged in five types of processes that supported learning about identity and values: creation, introspection, communication, participation, and perspective taking.

– **Creation:** Kids designed personal homes and temples, virtual autotopographies in which collections of symbolically meaningful objects and characters are displayed. The creation of these spaces supported the development of new insights about identity and values. Kids also created a participatory micro-community, a safe space where powerful conversations and self-government took place.
– **Introspection**: Throughout the experience with Zora, kids engaged in thinking about what types of places, objects, characters and stories best represent themselves as individuals and as a community. In this sense, Zora served one of the functions that has been attributed to the idea of the Sabbath: a time for reflection and self-examination (Heschel 1951).
– **Communication:** In Zora communication is both synchronous (learners converse with each other through their avatars in real-time) and asynchronous (learners post messages, read and write stories stored in their artifacts and engage in conversations with already programmed objects.) By communicating with each other kids not only expressed their sense of self and values, but also learned how to exchange opinions and debate.
– **Participation:** A sense of self doesn't develop in a vacuum but in constant interaction with others in a community. Zora engaged kids in self-organization and decision-making by supporting the creation of a participatory micro-community. Values became not only matters of narrative and introspection but also matters of behavior and taking action.
– **Perspective-taking:** Seeing the world as others do, understanding their motivations and actions, is a fundamental mechanism for exploring issues of identity and values. In Zora this type of experience was facilitated by kids

visiting each other's virtual homes and temples and, in SAGE's spirit, by engaging them in programming conversational interactions between their characters and other users.

Lessons learned

Zora's design was a result of my previous experience with both SAGE and Kaleidostories. In the same spirit as SAGE, in Zora children can program storytelling interactions for their characters to engage in conversations with the visitors. However, the natural language parsing is simpler and WordNet is not used to augment keywords. In Zora, as well as in Kaleidostories, narrative is the principal medium to form a community. Children can tell stories as well as contribute to the collaborative values dictionary. However, communication is both a-synchronous and synchronous. Real-time chat facilitates exchanging points of view in discussions.

The three-dimensionality and the navigation around the city have certain similarities with popular video games. The kids do not see Zora as educational software but as a captivating game. Kids had lots of fun with it, which is important to keep them engaged, and at the same time were able to explore aspects of their identity and values as well as discuss issues relevant to the Zora community and society at large.

Conclusion

As shown through the three identity construction environments presented above, SAGE, Kaleidostories and Zora, the integration of narrative with computation is a powerful tool to help young people explore identity and values. Narrative supports the construction of a sense of self by finding coherence between different aspects and experiences. It can also serve a healing function by allowing people to reflect back on their experiences and tell and re-tell their story (White & Epston 1980). Computation allows users to become designers of a context in which to engage in storytelling interactions, "what if" situations and real-time communication.

The research described in this chapter shows the potential of computational tools, particularly identity construction environments, for educational and therapeutic interventions that seek to foster self-awareness, personal cultivation and multicultural understanding. It also provides a new way of conceiving moral and civic education in the light of new technologies. There is a big potential in the integration of computation and narrative in the design of tools for learning about the inner world. This chapter hopes to provide some examples of how this can be done.

Acknowledgments

I am grateful to my advisor at the MIT MEDIA Laboratory, Seymour Papert and to Sherry Turkle, Mitchel Resnick, Joe Gonzalez-Heydrich and Edith Ackermann for insightful discussions about different aspects of my research. Thanks to Linda Stone and the Microsoft Virtual Worlds research group; to Claudia Urrea, Aaron Arakawa and Jon Chu, for their work on Kaleidostories, and Daniel Vlasic, for Zora. Also thanks to Justine Cassell who advised my work on SAGE, and to my husband, Josh Bers, for editorial comments on this chapter.

References

Bers, Marina, Joe Gonzalez-Heydrich, & David DeMaso (2001). Identity construction environments: supporting a virtual therapeutic community of pediatric patients undergoing dialysis. In *CHI'01 proceedings* (pp. 380–387). ACM.

Bers, Marina (2001). Identity construction environments: Developing personal and moral values throuch the design of a virtual city. *The journal of the learning sciences <i>, 10* (4), 365–415, Lawrence Erlbow.

Bers, Marina & Justine Cassell (1998). Interactive storytelling systems for children: Using technology to explore language and identity. *Journal of interactive learning research, 9* (2), 603–609.

Bers, Marina, Edith Ackermann, Justine Cassell, Beth Donegan, Joe Gonzalez-Heydrich, David DeMaso, Carol Strohecker, Sarah Lualdi, Dennis Bromley, Judith Karlin (1998). Interactive storytelling environments: Coping with cardiac illness at Boston's Children's Hospital. In *CHI'98 proceedings* (pp. 603–609), ACM.

Bers, Marina (1998). A constructionist approach to values through on-line narrative tools, In *ICLS Proceedings*, AACE.

Bruckman, Amy (1994). *MOOSE Crossing: Construction, community and learning in a networked virtual world for kids.* Ph.D. dissertation. MIT Media Lab, Cambridge, MA.

Calvino, Italo (1972) *Invisible cities*. NY: Harcourt Brace Jovanovich.

De Bonte, Austina (1998). *Pet Park: A graphical constructionist community*. MS Thesis. MIT Media Lab, Cambridge, MA.

Erikson, Erik (1950) *Childhood and society*. NY: Norton.

Fellbaum,Christiane (1998) *WordNet: An electronic lexical database.* Cambridge, MIT Press.

Gardner, Howard (1983) *Frames of mind: The theory of multiple intelligences.* NY: Basic Books.

Gonzalez, Jennifer (1995). Autotopographies. In G. Brahm Jr. & M. Driscoll (Eds.), *Prosthetic territories. Politics and hypertechnologies*. SF: Westview Press.

Heschel, Abraham Joshua (1951) *The Sabbath: Its meaning for modern man*. NY: Farrar, Straus & Giroux.

Kohlberg, Lawrence (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior*. NY: Holt, Reinehart & Winston.

Kohlberg, Lawrence (1982). The just community approach to moral education in theory and in practice. In *International conference on moral education*. Fribourg, Switzerland.

Linde, Charlotte (1993) *Life stories. The creation of coherence*. Oxford: Oxford University Press.

Papert, Seymour (1980) *Mindstorms: Children, computers and powerful ideas*. New York: Basic Books.

Polkinghorne, Donald (1988) *Narrative knowing and the human sciences*. State University of NY Press.

Resnick, Mitchel, Amy Bruckman, & Fred Martin (1996). Pianos not stereos: Creating computational construction kits. *Interactions, 3* (6), 41–49.

Turkle, Sherry (1984). *The second self: Computers and the human spirit*. NY: Basic Books.

Turkle, Sherry (1995*). Life on the screen: Identity in the age of the internet*. NY: Simon & Schuster.

Umaschi, Marina (1996). SAGE storytellers: Learning about identity, language and technology. In *ICLS 96 Proceedings* (pp. 526–531). AACE.

Virtual Worlds Group/Social Computing Group, Microsoft Research, Microsoft Corporation, http://research.microsoft.com/scg/

White, Michael & Epston, David (1980) *Narrative means to therapeutic ends*. NY: Norton.

Part II

Story Generation

Chapter 7

The Dr. K– Project

Brandon Rickman

Introduction

Dr. K– was constructed as an MFA thesis project in design and media art. The project is an attempt to create a fabricated narrative environment, an environment where nothing exists except that which is visible to the user. While a fabricated environment can be presented in varying degrees, Dr. K– was constructed with an extreme degree of fabrication.

Dr. K– operates as a text-based interface. The user is presented with a screen of text:

Office. A bench. A coat rack. Some furniture.

The screen updates as a result of user interaction:

Office. A bench. A coat rack. A desk. A character enters.

All Dr. K– stories begin with:

A place. Some scenery.

The narrative that follows is a unique virtual reality experience constructed out of historical elements from Edinburgh in the 1820's.

Operation

The user interacts with the Dr. K– system through a text-based interface. A set of elements is presented to the user on a screen, and these elements can be selected with a mouse. Because the narrative is constructed through interaction with the user, navigation occurs in one direction only – there is no way to go back to a previous page of the narrative. At the same time, even though the user is advancing through the narrative, there is no sense that the narrative itself is advancing in time. It is not a narrative developed in "real time", but in a lurching and discontinuous fashion.

This operation is analogous to an audience continually asking questions to a storyteller while the storyteller performs. The audience may be interested in elements that have been given only a cursory mention by the storyteller, and they ask the storyteller to elaborate on those elements. The storyteller can provide a direct answer, or be evasive in a number of ways. Here are three sample situations to help reveal some of the system operations.

Situation 1

A place. A plant. Some scenery.

The user selects "some scenery"; the user is asking for more information about the scenery. Because there are few elements in the scene, and the element selected is an archetype, the element will transform into a more specific state:

A place. A plant. A curtain.

Situation 2

Office. A bench. A coat rack. A desk. A character enters from a portal.

The user selects "a character enters from a portal." Because the scene does not yet contain either "a character" or "a portal", one of these elements will be added to the scene:

Office. A bench. A coat rack. A desk. A portal. A character enters from the portal.

Situation 3

Armory. A wall. The wall is old and warped. A sound. A man. Something happens.

The user selects "a sound"; the user is asking for more information about the sound. Because this scene already contains a number of details, a preference will be given to incorporating "the sound" into an action, by way of making the action "something happens" more explicit:

Armory. A wall. The wall is old and warped. A sound. A man. A character interacts with the sound.

Structure

There are four structural elements in Dr. K–. These are props, scenes, actors, and actions. Within each element type, there are a number of classes that define sets of those elements which share certain traits. For each class there is an archetype element that represents the abstract incarnation of that class. An archetype may also belong to a number of other classes.

Because all four element types have a class and archetype structure, elements of any type can undergo certain kinds of transformations. Elements can flux into something more specific, from an archetypical state into a more concrete state, or flux into something less specific, from a concrete state into a more archetypical state. An element changing to a more specific state can be described as coming into focus, and an element reverting to archetypical form is becoming unfocused.

Props

A prop element is a piece of scenery, a self-contained object or environmental component. Some sample props and the classes they belong to (classes are denoted with a ∼, archetypes with a *):

rug: [∼floor/ground, ∼cloth, ∼scenery]
potted plant: [∼plant, ∼scenery]
pickaxe: [∼tool, ∼weapon, ∼prop]
curtain: [∼cloth, ∼wall, ∼scenery]

Here are some classes and the props they contain:

∼floor/ground: { *floor/ground, rug, the floor }
∼cloth: { *cloth, rug, curtain }
∼scenery: { *scenery, rug, potted plant, curtain }

Note that the classes are not hierarchical: although the ∼plant class is seemingly more specific than the ∼scenery class, the "potted plant" prop belongs to both. Elements are never more than one step away from any of their archetype forms.

A prop that is the focus of the user's attention tends to flux into a more specific state. Props that have been neglected by the user tend to revert to more archetypical states. Thus it is possible that a rug, after a period of inactivity, may revert to *cloth. But if the user then focuses on the *cloth, it may transform into a curtain. An element that was once a type of ∼floor/ground is now a type of ∼wall!

There exist classes and elements for which there is no easy nomenclature. These elements, like the ∼floor/ground class above, are given a hybrid name that is (hopefully) descriptive of that class. In other words, the ∼floor/ground class contains props like "a rug", or "the ground", without signifying any additional qualities, such as whether that element is an "indoor" or "outdoor" prop.

Scenes

A scene element is a named location, a construction of place. Scenes are defined by the collection of props that they archetypically contain. Here are some sample scene elements, the classes they belong to, and the props they contain:

office: [~business, ~scene] { desk, wall, coat rack, machine }
bedroom: [~house, ~scene] { bed, pillow, wardrobe }
street: [~scene, ~outdoor] { pile of dirt, *smell, litter }

Some scene classes and their elements:

~scene: { *scene, office, bedroom, street, tavern, cemetery }
~business: { *business, office, tavern }
~house: { *house, bedroom, kitchen, entry hall }

Scenes are never the direct focus of the user. Scenes tend to become more specific, and do not generally return to the archetypical state.

Scenes are a byproduct of user interaction. As the user explores the environment, a collection of props is established. Some of these props may belong to one or more defined scenes. Upon reaching a certain threshold (a number of props suggest a number of possible scenes) the scene transforms to a more focused state.

Actors

Actors are much like props, but they can be mobile or autonomous. Some sample actor elements:

scoundrel: [~criminal, ~character]
male character: [~adult, ~male, ~character]
mob: [~mob, ~character]

Some actor classes:

~character: { *character, scoundrel, male character, mob }
~criminal: { *criminal, scoundrel, assassin }
~adult: { *adult, male character, female character }

Actors are subject to transformations similar to props. They change to a more specific state when they are the object of attention, and they revert to a more archetypical state when ignored.

Actor classes are determined by a mixture of physical characteristics and character roles.

Actions

Actions are the final element type. An action is an internal event, an action occurring within the scene. This is distinct from actions external to the scene, such as the transformation of an element to a different state as a result of user interaction. External actions are not represented by action elements.

Actions can have a number of props and actors associated with them. These associated elements are not explicit elements, rather they are denoted by the set of classes to which they belong. Actions also provide some simple grammatical elements used to construct English text. Some sample actions and their associated elements:

walk to: [∼action, ∼blocking] ([∼actor, ∼character], "moves towards", [∼prop, ∼furniture])
attack: [∼action, ∼harm] ([∼actor, ∼character], "attacks", [∼actor, ∼victim])
play: [∼action, ∼gambol] ([∼actor, ∼child], "is playing")

And some action classes:

∼action: { *action, walk to, attack, play }
∼blocking: { *blocking, walk to, walk to actor, walk from actor }
∼harm: { *harm, damage, attack, attack with }

Like props and actors, actions are subject to transformations into different states. At the same time, the props and actors associated with an action may change state.

Actions do not have specific temporal associations. A set of actions does not necessarily occur in the order they are listed.

At this point it should become clear to the reader that the framework provided by these different element types can potentially result in a number of troublesome situations. These situations include:

– An action which requires the existence of a prop or actor element that does not currently exist. A character enters from a doorway that is not present in the scene.

- An actor or prop may transform to a new state that does not satisfy the class requirements of an associated action. A character enters from "a doorway", but the doorway subsequently turns into "some scenery"
- Actions have no causality. A character may be dancing with a second character while attacking that character at the same time.

It is certainly possible to constrain the system to avoid these situations, or one can consider these situations as an emergent phenomenon of the system.

Mode of interaction

Given the three sample situations and the breakdown of story elements above, it should be clear that the form of interaction in this system is different in nature than the modal dialog used by many narrative environments. With a modal dialog an explicit list of user choices is displayed. In some situations the user is in control of a specific character who is engaged in a dialog with another character. The user may not know the specific result for any of choices, but she does know that by making a choice, she is actively choosing a specific mode of interaction between elements in the narrative. In the case of a character-to-character dialog, she is aware that the choice will result in one character "saying" something to another character.

When the user is in control of a character, that character is an avatar for the user. Dialog with the avatar can be interpreted as an internal element of the narrative, as well as an external dialog between the user and the system. As to the user awareness of the interactive process, the user in a dialog may expect other characters in the dialog to respond in kind, yet the exact nature of this response is not guaranteed or predictable.

In the case of Dr. K–, the user does not directly control any of the actors or elements. Here the choices are implicit – any fragment of the story can be selected, and the dialog between the user and the story elements, such as it is, takes place outside the story. Additionally, there is little user knowledge about how a selection will influence the story. Choosing to interact with "a character" may cause the character to change into "a child", or the issue of character identity may be ignored by the story engine entirely in favor of some other outcome. The user does not choose the mode of interaction.

Flux

As a result of continued user interaction, elements which are initially generic in form – "some scenery" – tend to become highly focused – "some litter, the litter is scattered all over" – with continued interaction.

Not only do elements come into focus, they can lose focus through neglect. When elements revert to more archetypal forms, "a desk" becomes "some furniture".

(Of course, since the user does not control the mode of interaction, attempts to manipulate the flux of the narrative may not have the desired result.)

Situation 4
A place. A tea chest. A body bag. A scrap of paper. An assassin. A character. Something happens.

The user selects "a character." This element will change to something more specific, but at the same time "a scrap of paper" will lose focus:

A place. A tea chest. A body bag. Some scenery. An assassin. A rascal. Something happens.

Because some elements must be in focus for other elements to lose focus, there is a gradual buildup of more-or-less focused elements as the story progresses. Highly focused elements are assumed to be of interest to the user, are considered to be of more consequence in the story, and are thus less likely to lose focus through neglect. In this way the elements of the narrative build up into remarkably evocative situations, despite the ambiguities and contradictions of the elements.

Visually, the flux of elements is represented by the opacity of the text in the interface. 80% black text indicates a very focused element, while 20% black text is almost invisible and indicates that the element is a candidate for change. Flux is measured by a combination of variables: the age of the element (how long since the element was last selected), how many times the object has been selected, the relevance of the element to the current scene (a table is relevant in a kitchen, a tombstone is relevant in a graveyard).

Weighted choices are made to determine when elements will change their degree of focus, and at what intervals new elements will be added to the scene. If a prop is selected and the number of props in the scene is above a certain threshold, actors are introduced into the scene. If an actor is selected and the number of actors is above a certain threshold, actions are introduced. When the number of actions is above a certain threshold, the engine resets and the user returns to exploring the first level of narrative.

Part 2: Simulation versus fabrication

In the process of interacting with a narrative environment, there is a tendency for the user to look for the limitations of the system. This may occur in a number of ways, depending on the kind of environment:

- In a 3D environment, the user may try to break away from the model, or find holes in the model that provide an external view of the environment.
- The user may try to exhaust all branches of a dialogue tree.
- The user may become intentionally unresponsive, to see how the environment operates without stimulation.

Once the user has explored these limitations, he can feel he has in some way mastered the environment, and will then proceed to explore the environment according to its internal logic.

The ability to perform these user experiments is often mistaken as a qualitative measure for interactivity. Indeed, in some cases a system must demonstrate a fundamental level of responsiveness if it is to be useful. In a narrative environment, however, this kind of exploration can lead the user to the Edge of the World. The Edge of the World is the place where the user has broken out of the environment without consequence.

If the user in unable to find an Edge, or the Edge of the World is elusive, then it is more likely that the user will become immersed within the narrative environment. In such a case, the user's path through the environment becomes significant. The user's path consists of a collection of character interactions, locations, and other notable elements. As particular points of interest, these elements form landmarks in the environment. Thus the user experience generates a set of landmarks, and those landmarks define a narrative. This is the ideal result of any narrative environment.

Characteristics of simulation and fabrication

A simulation operates on a mechanical model of the universe. A simulation may incorporate some randomness as a substitution for unknown or undefined variables in the system, but is otherwise made up of a collection of determinate properties: objects, physics, continuity, &c. Because simulations are frequently modeled on the "real" world, and because objects in the real world are not spontaneously generated or destroyed, the simulated world exists in a knowable state at any given moment. In other words, someone has built all of the virtual walls, floors, and lemon trees of the simulated world ahead of time.

Simulation lends itself to highly representational presentation – immersive 3D virtual reality.

A fabrication, such as that proposed by the Dr. K– project, operates on a potential model of the universe. There is randomness as a means to add diversity to the world, which would otherwise remain a homogenous world of potential. Exploring a fabrication creates an artifact, a fixed collection of objects and events, and the world is not otherwise knowable except by exploration. Fabrication lends itself to the constructive process of narrative, a pre-representational environment.

The qualities of experience in simulated and fabricated environments are quite distinct. These qualities are distinguished by a number of characteristics:

Repeatability: The same experience can be repeated multiple times by many users within a simulation. This is a scientific quality; repeating the experience is a useful way to test that the system "works". The fabricated experience is not prone to repeatability; each experience is intended to be unique for a particular user across repeated interactions.

Direct user control: The user can have control over one or more agents within a simulation. User control in a simulation will often "break" the simulation, much like tinkering with the innards of a clock may prevent it from working properly. The user has only limited control of agents within a fabrication. When the potential of the fabrication is exhausted, agent actions become determinate and thus out of control of the user.

Viewpoints: Because the simulated world is well-defined, there are a potentially unlimited number of viewpoints into the world. Virtual reality often conflates the high quantity of viewpoints with the notion of "interactivity". Because fabrication does not allow the user an unrestricted view of the world, a fabrication has only a fractional number of viewpoints.

Diversity of landmarks: A landmark is a memorable element in the experience. It can be a character, a location, a special event. In a simulation, the user may experience the same situation repeatedly. Only highly unique situations stand out from all others, resulting in a low number of true landmarks. Fabrication is concerned with the direct construction of these landmarks; fabrication fails if there are no distinguishable landmarks.

Representation: Simulation is dominated by high-fidelity representation. The higher the desired quality of representation, the longer the production cycle will be. If the simulated world is unclear or ambiguous there is a gap in user comprehension. With fabrication, ambiguity is the vehicle for exposition; to represent the world piece by piece is to tell the story of that world.

In the worst case scenario, the characteristics of a simulation result in a guided tour experience. So much effort is invested in creating the world, testing the world, and making the world look as good as possible, the author must ensure that the user does not miss any of the "good stuff". Indeed, the highly designed areas are the most aesthetically pleasing, given the otherwise mundane experiences that make up a majority of the simulated world.

In all fairness, the worst case scenario for a fabrication is unparseable nonsense.

Avoiding nonsense

There is one major pitfall to avoid when creating a fabricated world: the repeated generation of unparseable nonsense. Steps need to be taken to insure that the world will become more than a highly potential form, more than a random collection of elements.

Dr. K– uses a few different approaches to maintain some level of comprehensibility within the narrative.

The first approach is the choice of specific subject material for the story. Dr. K– is constructed around the historic account of William Burke and William Hare in 1820's Edinburgh. The elements of the project are drawn from this particular story, elements that when viewed in quantity create a underlying setting and mood for the story. (Burke and Hare were notorious criminals – they made a living by killing people and selling the bodies to an anatomy school.) This should not be considered a backstory, however. It exists more as a background motif, like a melody line for a set of musical variations. Users are not exploring an environment where they will literally discover Burke and Hare, but they may recognize some of their elements within the narrative.

Second, the project can be presented within a theatrical setting. The elements of the set – a desk, a bench, a tea chest – echo the elements within the story environment, reinforcing the place and mood of the story. Even when incongruous elements are presented by the environment, the theatrical setting helps to emphasize hidden relationships between those elements.

The project can be performed, or it can be arranged as a walk-in installation. A performance serves to engage an audience that may otherwise be too reluctant (or too jaded) to participate. At the same time, a performance may frustrate those who wish the explore the project on their own. These issues are characteristic of many virtual reality presentations. With any VR presentation, the choice of setting will affect the audience's perception of the piece, and will influence their tolerance for nonsense.

Finally, the story engine is tuned so that there will be some sense of progression when there is long term interaction. Scenes are realized as a result of certain props, actions are realized by introducing suitable elements. "Sensible" constructs are coded into the relationships between props and scenes, and between actions and their target elements. When these constructs appear, they reassure the user that there may be other associations hidden within the nonsense. Tuning also determines how likely the coded constructs are to transform into more difficult collections.

Closing comments

This project is documented and presented as an artistic exploration into interactive storytelling. As an author and artist I feel that the current emphasis on simulation and photorealistic representation in the realm of "interactivity" needs to be critically reevaluated.

The main source of "intelligence" in Dr. K– is the cognitive ability of the participant herself. The ability of a human mind to recognize and understand symbols, even when those symbols are distorted and occluded, is key to the operation of a minimal representation. This has led me to a difficult question:

Can one convey meaning [have connotation] without creating symbols [without denotation]?

Where is meaning generated within a narrative? One can manipulate the associations of objects without denoting the objects themselves: "There exists an entity that performed an action." If the action is resolved as being of a threatening nature, the entity will likewise be perceived as threatening, without any change in representation. Likewise, the action can become threatening without change in representation if the elements associated with it have threatening characteristics. But these subtle interactions are not possible with the typical models of representation – in particular with the current model of virtual reality and the efforts of photorealistic visualization. Virtual reality presents all associations at once, from any viewpoint. This leaves little room for new or alternative associations.

If these associations are made to be mutable, as has been attempted in Dr. K–, one can exploit that critical moment of interactivity when the unresolved associations are firmly realized – when a collection of objects becomes a place, when a character becomes a villain – and one can discover the moment in which is contained the seed of drama and the creation of narrative.

Chapter 8

The Rise and Fall of Black Velvet Flag

An "intelligent" system for youth culture documentary

Sheldon Schiffer
Georgia State University, Atlanta

> A record, if it is to be useful to science, must be continuously extended, it must be stored, and above all it must be consulted.
>
> Vannevar Bush

Making order and telling stories

I have a habit of spending as much time organizing my materials to make a film as I do actually shooting and cutting it together. I realized 10 years ago from the pleasure that I experienced that this organizational process was a creative one. I made complex libraries that would enable me to try hundreds of possibilities without losing a single image or sound clip. This was in the days when nonlinear editing was cost-prohibitive to the documentary filmmaker.

Depression set in when I had to dispose of many ideas and intriguing possibilities to serve the needs of a primary narrative. Brenda Laurel refers to the Flying Wedge model where even in game playing, as participants (or spectators) engage further, the scope of interest narrows to accommodate the more precise objectives of the players (Laurel 1992: 78). And so, as is the case in most media production, a lot of precious things (and useless things too) end up on the editing room floor because they do not serve the purpose of the author(s).

When I began to see that querying a database was a request to be told about something, I realized multimedia powered by database engines might allow for an autonomous reconstitution of narrative and the audiovisual materials on which it depends. In a sense, every "story" the system tells can be responsive to the unique qualities of the audience that asks for it.

Now that you have an idea of the context that attracted me to this discipline, I must confess that my enthusiasm needs to be coupled with some principles to consider. What I intend in this brief document is to describe some conditions where databases are useful to me in the construction and presentation of my own narrative expression, specifically an interactive documentary called *The Rise and Fall of* Black Velvet Flag. It is also important to reference some of the history of the predecessors whose work I have found useful. But before I can continue and give a more specific account of my creative strategies, I should define my use of some terms that I have repeated and will continue to use throughout.

How a query result might create "narrative"

Some digital media theorists have resisted the legacy of the concept of narrative, and celebrate a liberation from its imposing structures and systemization of temporal and spatial representation for digital media. Specifically, I am referring to Lev Manovich and his discussion of database culture and its resistance to the implied orderings of classical notions of narrative (Manovich 1998: 80–82). However, my experience with databases differs. While looking at a minimal list generated from a database query, there still remains from the representation of an object by its ordered title a set of loosely unanchored signs held in each field of a record. From these signs the perceiver may uncontrollably detect a pattern, a story perhaps, as to the logic behind the structure of the list. It is the structure and representation of the data that are the essential components that allow a list to be a narrative. The degree to which the data structure and query logic reveal themselves in the result returned is an indication of the comprehensibility of database-derived narrative.

David Bordwell's exhaustive inquiry into classical Hollywood cinema provides some useful parallels. As he used Vladamir Propp's taxonomic study of Russian folktales (Propp 1968) to understand the underlying functions of Hollywood films (Bordwell 1986: 18), I believe that a similar application to the database query is useful. Three important concepts are relevant in my definition of narrative that I will use from Propp. Fabula describes a sequence of events in a causal and/or chronological order that a perceiver can extrapolate from a text. Syuzhet describes the actual presentation of events in the text (some would call this the plot). Also there is narration – a cue to the perceiver to construct chronology or cause-and-effect relationship by recognizing patterns of events represented in the text.

A simple list resulting from a database query may demand some intellectual labor from the perceiver to make any meaningful associations. If a search mechanism is programmed to evaluate thematic or chronological data for each record, then the three elements of narrative that I use from Propp can be present. That is the premise from which I create my database-driven documentary.

I anticipate one problem; it is implied from the notion that narrative does not exist outside the mind of the perceiver. Narrative in a documentary is the result of a series of pre-planned understandings resulting from cues of narration that fit remembered patterns of logical or rhetorical structures – the syuzhet. With database culture still in its infancy, we have very few database-configured patterns embedded in our collective memory. Therefore I rely on other forms of expression. Cinematic and televisual patterns of narration are useful in this inquiry.

Another term must be clarified in order to proceed. Documentary is certainly a term with which recent cinema and its critics have come to no agreement. For the purposes of this discussion I will refer to the idea of documentary as the expressive form that utilizes physical objects that were not fabricated for the purpose of the immediate presentation, but have independent signification in a distinct and separate context.

Using a database for the presentation of interactive documentary

To address this issue, let me first say that creative work is not necessarily educational. Documentary has much to offer the student or researcher of any subject. An educational outcome is not my only objective. I intend for my documentary films and videos to entertain in a similar way as fiction films. My interest is to affect the viewer with a memorable experience rather than instruct on a particular concept. But unlike many fictional dramatic films, my documentary films are not scripted. They are loosely defined in their cinematic or videographic plan. And the material I bring to the editor can be, and has been, restructured many different ways to suit very distinct purposes.

What makes a nonfiction documentary film, video or multimedia project distinct from its fiction cousin is that the objects used in the making of the finished work often have a previous life of their own, a value to another context, another system of meaning. The photograph on my license serves one context:

it enables me to drive legally. And yet, that card could be a useful document in the making of a documentary about the construction of identity.

This other context is what makes any conscientious creator feel the arbitrariness of giving a narrative function to any document in a film. Two questions a responsible author must ask themselves repeatedly in every project: "Is my use of this document respectful of the meaning it evokes outside of my film's context?" and "Am I making the most effective use of this document for my own narrative?" (The answers to these questions often contradict.) A photograph or film clip in one of my projects may serve an insignificant need if in the end I must tell only one story, told one way, in a fixed order, by only me. But, given the freedom to re-purpose my material to accommodate a variety of other possible narratives, I may be able to exhaust more of the possible significations any photograph or film clip may offer. Though a documentary can have only one (sometimes unknowable) fabula (representation of time) to reference, multiple syuzhet (plot and rhetorical structure) can be the result of numerous restructurings and revisions of the author's narration. But with multiple syuzhet, the perception of the order of events can progressively change in the mind of the viewer making a single "truth" difficult to validate. The John F. Kennedy assassination and its panoply of documentaries and re-enactments represent a single and elusive fabula exemplary of this conundrum. Akira Kurosawa's film *Rashomon* (1950) is another example that constructs its meaningfulness on this intractable problem.

Multimedia enabled creators to design multi-threaded narratives. But we have crossed a junction where many authors realize that branching structures still have profound limitations. While we have experienced branching story structures in many nonfiction multimedia works, an accounting of each possible version can be made with a matrix and a linear equation. Hence in branching stories, the narrative trail gains more possible tributaries and alternate routes, but adds no responsive mechanism to the potentially dynamic intentions of its authors and the changing interests of its audience. Only intelligent systems developed to accommodate this change allow autonomous, dynamic and even modular characteristics to be realized. Only systems that use aleatoric mechanisms can produce an experience that represents the complexity and arbitrary characteristics of a documented reality.

One such system that explored a similar set of concerns in the making of documentary is the "Autonomist storyteller system" (Davenport & Murtaugh 1997: 446). These authors saw the limitations of affecting the viewer with the television documentary, and wanted to find another way to engage the viewer interactively while allowing for recombinant narrative structures to be determined by keyword matches. The resulting experience for the viewer is unpredictable, yet by carefully coding the video clips with attached alphanumeric values, narration is maintained in its presentations of sequenced material.

The effect of the "Autonomist storyteller system" as a communicative work indicates that the documentary is a potentially appropriate narrative genre from which to contribute to this developing expressive form. Each document can be catalogued and coded in a database for the service of multiple narratives. If the database is programmed to organize its records to make narrative construction possible, than a query to the database is, in a sense, a request for it to tell a story. Or, a query more specifically is a temporal and spatial ordering of records sensitive to the logical systems created for the type of narrative the author and the audience intend. Furthermore, the presentation interface is an environment where the author and audience develop objectives for evolving a story. Or more specifically stated, multiple queries are given a context where they can have value in relation to each other. Although I suggest here a rudimentary use of a database system for an intelligent documentary narrative engine, the iterated possibilities that I am using for the project based on the content described below are more complicated.

The use of narrative intelligence to document youth-culture phenomena

The introduction of databases to the creative author does not eliminate the viable use of linear moving images as an effective means of communication. Some subjects of inquiry, however, are exceptionally appropriate for exploration.

Youth culture provides two conditions that make it ideal terrain for database documentary form. First, it has some stable elements that the host adult culture imposes as youth labor to construct identity. Young people since World War II have spent their leisure time and money exploring and creating fashion, music, literature, cinema and drugs as symbols to differentiate themselves from adults. In contrast to the stable existence of these elements, their content is dynamic. A progressive exploration through time reveals exceptional qualitative variances of selection. Hence the second important condition is that these elements contain variables that change – dynamic data. This dynamic data is the collection of documents that a film or video maker uses to construct the syuzhet, a logical or rhetorical structure that shapes the experience of the viewer.

However, the data in the database is not the resulting cultural artifact we might call a digital movie. It is only a collection without a means of seeing it. The purpose for collecting, ordering and retrieving documents is revealed in the design of a query system. The series of seven films produced in the 1940s by Frank Capra, *Why We Fight*, was developed to persuade the population of United States population to support the war against Germany and Japan (Sklar 1993: 265–266). Hence, the films are query results from a collection of moving images and sounds that followed a politically determined algorithm. Certainly no machine-readable code was written to carry out this enterprise in propaganda, but the functional behaviors of the process are parallel. A body of the United States government created a policy and an agency to produce these films with a specific didactic intention. The moving images and sounds were collected, catalogued, ordered, selected and implemented for each film following an algorithm that was the policy of the war time government of U.S. President Franklin Roosevelt.

The example of the *Why We Fight* series provides a pre-digital model of the database as a mechanism for the production of documentary. The essential difference is that Capra had to cut his negative and commit his materials to one query result for each query – the finished films. The computer-hosted database provides the possibility of continuous re-constitution and multiple query results. It also allows the possibility of autonomous behaviors sometimes unintended by the author. Without the propagandistic algorithm pre-defined by the Roosevelt administration, other structures and hence other films constituted by the same collection of documents might make for another understanding of the experience of World War II.

Not unlike wartime journalism, documenting and understanding youth culture is an increasingly difficulty task. The 1970s was a decade when youth culture evolved fewer genre of expression. But the channels of cultural production were directed at more specific demographics than at any other time in history. Hippies and punks were once at binary poles with very little in between. Today the list of sub-cultural identities not only revisit past ones, but also the genera continue to propagate new species and subspecies every year. Ready to exploit, and even invent this dimension of cultural production are the major commercial forces that produce media for all – young and old, mainstream and "alternative."

Youth culture has, since World War II, and at moments before, when affluence and leisure time were abundant, been a leading force in the dynamic characteristics of the host culture of most every industrial nation. One explanation relates how the industrialized middle class family separates youth from the responsibilities and distractions of modern life typical of a working and consuming adult. A middle class young person has the free time to take in the media of the world, attempt to understand it, and create a material response to it. This process can proceed with few repercussions to the means of maintaining a livelihood – parents typically take care of that, even if their children have tattoos, piercing or colored hair. And even with some personal repercussions considered, the fact that young people are less obligated to property and its continuous acquisition allows a freedom of expression that usually is not revisited until old age, if ever again.

Youth culture responds to two social pressures. First from the young, there is the urge to acquire power in a social system that is abundant with contradictions. When the notion that a hard worker will be amply rewarded is frequently confounded, resistance to conformity has its first powerful rationale. Why conform to this notion if it is seen to be untrue? Sociologists Widdicombe and Wooffitt (Widdicombe & Wooffitt 1995: 17) answer the question:

> Subcultures offer a solution at a symbolic level. Subcultures solve at an imaginary level the problems which remain unresolved at the concrete material level, and this is why the solution is necessarily symbolic. Style enables the young person to achieve in image what they cannot achieve in reality.

Second, the dominant culture seeks to teach the young how to gain power. In a capitalist culture, this goal is accomplished by teaching the young how to buy, and more importantly, how to construct identity with what one buys. Once contradictions are realized, the incentives for resistance are discovered. From this resistance, a theme in the social history of urban post-WW II industrial society is realized – the handmade creation of identity through youth culture, its fashion and music. The materials, the artifacts and their symbolism are dynamic. Each generation finds its signifiers and modifies them constantly. Similarly, their meanings are modified (Hebdige 1979). Likewise, the young grow up. Many cease to resist. They accept the contradictions, or at least, they repress them.

The challenge of documenting this process is finding a way to accommodate massive change – change in the subject and in the audience. With few exceptions, most films I have seen on punk culture struck me as either too celebratory or too analytical. They either excessively subjectify or excessively objectify. But, perhaps the real culprit was me. At age seventeen, I was alienated by the objectification of punk in Penelope Spheeris' film, *The Decline of Western Civilization* (1981). But fifteen years later, I found it uncritical and accommodating. Can one film accomplish multiple objectives in the audience? Likewise, the evolution from these two poles was gradual. There was a place in-between when *The Decline of Western Civilization* was "just right" for me. I was ready for questions it asked and the answers it gave. The film contributed to my own evolution of consciousness about youth culture at different moments in time.

But the documents in that film are precious and unique. They cannot belong to another film without being taken out of context. How can Spheeris' film, *The Decline*, evolve without it being remade? Should it evolve? Or should it forever be a document symbolizing the ideas of those individuals of that time? If it were to be remade, could we save the original version so that we might compare and see the evolution? If we were to construct a mechanism that could integrate new content and accommodate an evolving author and audience, are we not creating a "context-controlled event-world"? New Media artist and theorist Peter Weibel describes the possibilities (Weibel 1997: 348):

> Another aspect of the variable virtual image is caused by the dynamic properties of its immanent system. As the system itself is just as variable it will behave like a living organism. It is able to react to the context-generated input, altering its own state and adapting its output accordingly.

Weibel continues his discussion by describing three specific characteristic elements of the digital image: virtuality (the way the information is saved), variability (of the image's object), and viability (as displayed by the behavioral patterns of the image) (Weibel 1997: 348). All three characteristics describe ways in which the image – and here I mean the cinematic, documentary image – changes. Hence the digital imaging system becomes a clear solution in the quest to find a medium that will represent change – in my case the evolution of youth culture as subject, youth culture as spectator, and ultimately, youth culture as author (myself and those who come after).

The use of the database as a creative medium is well argued by Manovich (Manovich 1997: 86):

> In general, creating a work in new media can be understood as the construction of an interface to a database... The database becomes the center of the creative process in the computer age.

I not only choose the database to achieve my creative objectives, I also feel as much commitment to providing a socializing experience. Sherry Turkle describes her work with young people in Multi-User Dungeons (MUDs). She surmised that networked gaming was a use of the computer and its correlated technologies not only "for thinking things through [but also for] working out personal concerns." (Turkle 1997: 356)

In this respect I come back to my original theme: the need for people to explore the contradictions discovered in the process of acquiring power from youth to middle age. My intention is to create a malleable system that will bring insight and identification as the generation that was punk comes to terms with its accommodations to that which it once rejected. And I intend to create systems that can generate new iterations from the intelligent responses it gathers from those who interact with it.

Why the Rise and Fall of Black Velvet Flag should be database-driven

The three members of Black Velvet Flag provide an unusual combination of characters and objectives. Two founding members work in advertising and were once punks who grew up in Southern California in the early eighties. Their adolescence was spent reacting to the pressures of conformity with LA punk culture. The ideology of the culture was critical of the political economy and the social values of the host culture, while paradoxically, punks enjoyed the freedom of expression and consumption it permitted. The third member was too young to know what punk was, and wanted to use the band to become a successful pop musician. The band's membership reflects the paradox that its music so profoundly illuminates. Together they performed hard core punk song lyrics that other bands wrote, while re-accompanying and rearranging the score to the lounge music of the fifties and early sixties – the music of the band's parents. The dialectic between the musical elements is disturbing, polemical and humorous. Most insightful listeners decoded the intent of the band. They learned something about the workings of media culture and its relation to youth culture while grooving to some very catchy tunes. Likewise, the audience contemplated their own passive behavior and politically anemic incentive to act on behalf of their conscience and convictions.

Historically and cross-culturally, youth in this century have been an active progressive political force to be reckoned with by every nation organized into a discrete political structure. For reasons that continue to fascinate me, youth in the United States since the mid-seventies have been among the most passive and narcissistic citizens of the world. What I set out to do in my work is to explore why and how this happens, and to find the examples and exceptions to this as well. What forces exist to co-opt and commodify the impulse to resist conformity, thereby making resistance an act of conformity? What forces exist that teach us to value private property, and then to conform and resist and conform again to a system that allows an individual to acquire and protect their

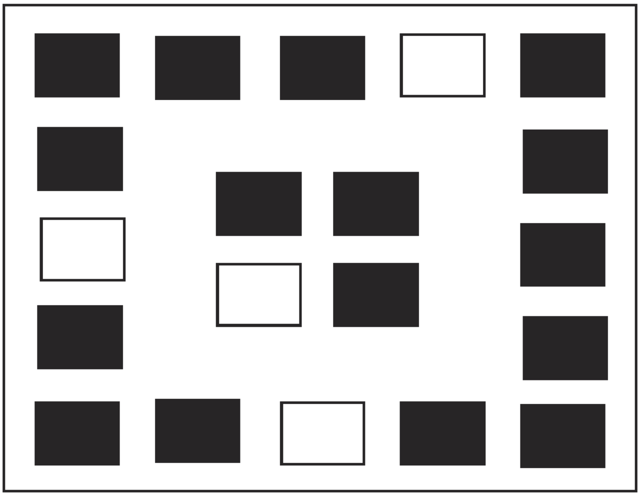

Figure 1. Static image selection screen.

acquisitions? It is the transition from the stage of questioning adult authority into assuming the power and role of an adult that is the subject of documented inquiry and my poetic response.

How the presentation system works: very simply

The user is presented with sixteen images randomly chosen that appear on the boundaries of the frame of the computer screen (See Figure 1). Of these sixteen, four among 32 different invisibly coded primary themes are represented – each image has its own theme. Each theme has four images representing it among the sixteen. The user must then select four images in any order by clicking on them. The user can pass the cursor over each image to hear a sound bite associated with the image to help decide to choose it. The images move to one of four positions in a rectangle in the center. The user can also make changes at will.

When the user has decided on the chosen four, s/he is taken to a screen where videos associated with these images play (See Figure 2). At the end of the viewing, the first screen returns and sixteen more images are selected. However, these are not all randomly selected. Twelve of the sixteen new images are selected by an algorithm. Their variables are provided by the theme data of the sixteen images previously chosen. Four of the new images are randomly se-

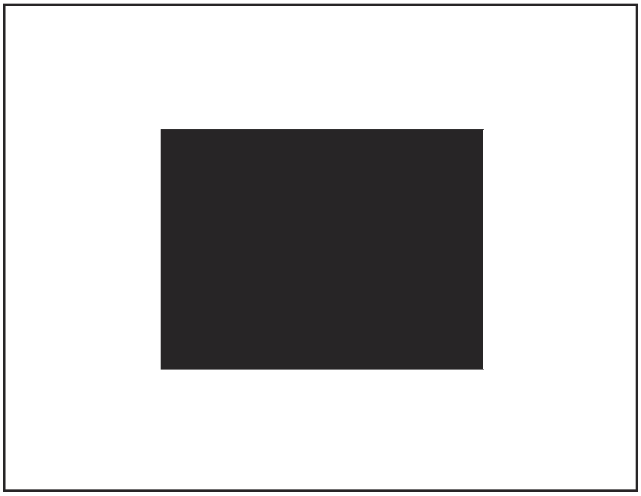

Figure 2. Video viewing screen.

lected. With each subsequent cycling between the two screens, two things happen: (1) the system collects and stores information from the choices the user makes, and repeatedly interprets this information to provide a constantly refined response to the user's interests, and (2) the user sees more and more of the audiovisual material in a variety of contexts at the service of multiple queries.

The four randomly chosen images invite the user to diverge from a path of inquiry that can become too narrow and predictable. The user can play for two minutes or many hours. There is no predetermined beginning or end. But some video clips are coded to occur as openers, others as closers. Most of the didactic or pedagogic intentions I may have for this film play themselves out in the algorithm. I am not saying that my control of the storytelling disintegrates; it never does. But the system allows a space for the user to subvert the narration to some degree. The degree of subversion remains to be seen.

References

Bordwell, David (1986). Classical Hollywood cinema: Narrational principles and procedures. In P. Rosen (Ed.), *Narrative, apparatus, ideology* (pp. 17–34). New York: Columbia University Press.

Davenport, Glorianna & Michael Murtaugh (1997). Autonomist storyteller systems and the shifting sands of story. *IBM Systems Journal*, *36* (3), 446–456. Reprint Order No. G321-5652.

Hebdige, Dick (1979). *Subculture, the meaning of style*. London: Methuen.
Laurel, Brenda (1992). *Computers as theater*. Reading: Addison-Wesley.
Manovich, Lev (1999). Database as symbolic form. *Convergence*, *5* (2), 80–99.
Propp, Vladimir (1968). *Morphology of the folktale*. Austin: University of Texas Press.
Sklar, Robert (1993). *Film: An international history of the medium*. Engle Cliffs: Prentice Hall.
Turkle, Sherry (1996). Constructions and reconstructions of the self in Virtual Reality. In T. Druckery (Ed.), *Electronic culture: Technology and visual representation* (pp. 354–365). New York: Aperture.
Widdicombe, Sue & Robin Wooffitt (1995). *The language of youth subcultures: Social identities in action*. London: Harvester Wheatsheaf.
Weibel, Peter (1996). The world as interface: Toward the construction of context-controlled event-worlds. In T. Druckery (Ed.), *Electronic culture: Technology and visual representation* (pp. 338–351). New York: Aperture.

Chapter 9

The recombinant history apparatus presents Terminal Time

Steffi Domike, Michael Mateas, and Paul Vanouse
Chatham College, Pittsburgh; Carnegie Mellon University, Pittsburgh, PA; SUNY Buffalo, Buffalo, NY

Introduction

Terminal Time is a history "engine:" a machine which combines historical events, ideological rhetoric, familiar forms of TV documentary, consumer polls and artificial intelligence algorithms to create hybrid cinematic experiences for mass audiences that are different every time. Through an audience response measuring device connected to a computer, viewing audiences respond to periodic questions reminiscent of marketing polls. Their answers to these questions allow the computer program to create historical narratives that attempt to mirror and often exaggerate their biases and desires. The engine uses the past 1,000 years of world history as "fuel" for creating these custom-made historical documentaries. By creating histories that clearly and instantly respond to changes in audience make-up, the project is intended to raise fundamental questions about the relationship of points of view to constructions of history particularly at the dawn of a new Millennium.

The audience interaction in relationship to the viewing experience is depicted in Figure 1. In the first question period, an initial ideological theme (from the set of gender, race, technology, class, religion) and a narrative arc (e.g. progress or decline narrative) are established. The second set of questions refines the ideological theme chosen in the first set, and possibly introduces a sub-theme (e.g. combining race and class, or technology and religion). The third set of questions further refines the theme(s) and introduces the possibility for a reversal (e.g. a decline narrative becoming a progress narrative).

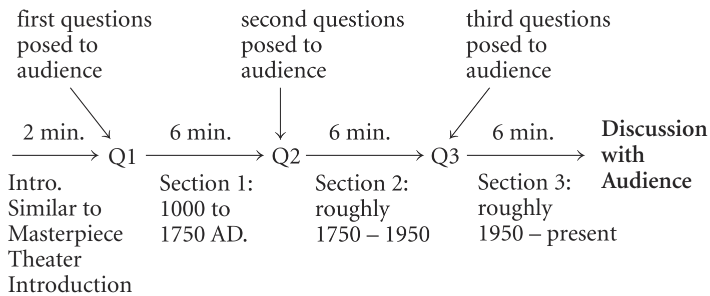

Figure 1. Audience interaction.

In the rest of this article, we will examine *Terminal Time* from three vantage points: its relationship to and examination of the popular historical documentary, its function as a critique of cyber-utopian navigation and control in "new-media", and issues of authorship and representation raised by using Artificial Intelligence (AI) techniques to operationalize ideological construction.

The "cookie-cutter documentary"

Ever since the first moving images were recorded, filmmakers have been aware of the power of their medium to effect historical meaning; the historical documentary became one of the first identifiable film genres. The popular model of this form in America today, most clearly exemplified by Ken Burns' "The Civil War," has the familiar structure of Western narrative: each program has a distinct dramatic arc, a beginning, middle and an end. The rhetorical structure – also familiar and now almost universally expected – invariably involves a crisis situation, a climax, and a clear resolution. Generally there is one prevailing narrative, one interpretation of the historical facts presented. Overall the tone set is one of progress. Usually the narrative is delivered to the audience by an unseen, yet obviously white, male narrator. So popular is this model that networks and cable channels, including the public television networks, rarely show programs that diverge from it - thus the form has become even more codified.

With *Terminal Time* we imitate the model of this "cookie-cutter documentary" with a machine that produces and reproduces it, until the model itself is revealed for the tool of ideological replication that it has become. Although dominant in popular media today, the cookie-cutter documentary is just one form of historical documentary. *Terminal Time* derives its impetus from the dominance of this archetype as well as from independent attempts to challenge the authority implied in the historical documentary and to posit alternative forms.

Terminal Time, as an exploration of the documentary form, has two points of entry. One is theoretical analyses of ideological structuring in mass media, in particular those made by Soviet filmmakers active in the early revolutionary period. In those early years of cinema, a great deal of experimentation took place; *Terminal Time* is indebted to those early pioneers of the film arts for their spirited quest to understand the ideological impact of their works. The second point of entry is the present homogeneity of mass media which reflects corporate ownership of media and domination over cultural institutions.

Historical roots for theoretical critique

V. I. Lenin, the leader of the Russian Revolution in 1917, encouraged the use of film as a political tool. Subsequently, Soviet filmmaking became established as an influential international model (Schnitzer, Schnitzer & Martin 1973). In those "astonishing and wonderful days" (Schnitzer, Schnitzer & Martin 1973: 13) of the early Soviet art world, filmmakers Lev Kuleshov, Dziga Vertov, Sergei Eisenstein, Esfir Schub and others created new visual languages. Within the context of the tumultuous expansion of Soviet art in general, and strong political support given to filmmakers in particular, the artists spoke and wrote about the theoretical challenges they faced, making important early analyses of the medium.

Filmmaker Lev Kuleshov is remembered for his deliberate tests of film editing the effect of on audience perception. He demonstrated that identical images can be used to mean very different things by pairing them with other imagery and narrative (Bordwell & Thompson 1997: 281). He found that it is the ordering of visual data that defines the meaning of mass media. The one who controls the order, controls the message. *Terminal Time* explores the "Kuleshov effect" even further by examining how meaning of imagery can be changed when juxtaposing it against different narrative texts. In fact, the Kuleshov effect makes this project possible. Imagery is plastic - a relatively small number of different video clips can illustrate a wide range of narratives.

Early newsreel producer Dziga Vertov challenged the medium with his original form of self-referential journalism; his goal was "to break out of the proscenium of the theater and to enter the arena of life itself." Vertov strove for kino-pravda, or "cinema-truth." He abhorred staged action and stated his mission to be "the creation of a new perception of the world" (Barnouw 1983: 58–

61). If made today, his work would not be called journalism at all, but would fall under the category of experimental documentary. Today's journalists cloak themselves in objectivity, intentionally distancing themselves from their subjects of inquiry. This technique has been instantiated throughout the media apparatus such that "news" is now universally presented as unquestioned truth, having no relationship to the person or corporate entity charged with its telling. *Terminal Time* further explores the notion of cinema-truth by creating an endless set of possible truths, all stemming from the expressed desires of the audience, who in this case are charged with choosing how the "truth" is told.

Esfir Schub's meticulous inter-cutting of footage of war, strikes, and other contemporary images of human suffering and struggle with home movies shot for Tsar Nicholas II instilled the home movies with new meaning. Images of the opulent lifestyle of the Romanovs were transformed and revolutionized through visual contrast. Barnouw credits Schub's editing work, in *The Fall of the Romanov Dynasty* (*Padeniye Dinasti Romanovikh* 1927) and two subsequent pieces, with advancing the genre of newsreel compilations (Barnouw 1983: 66). Shub also demonstrated that, original intentions notwithstanding, documentary film footage could be manipulated and its meaning re-contextualized to create powerful, alternative readings of history. *Terminal Time* is clearly indebted to Schub's work, as re-contextualization of historical materials is central to our endeavor.

Although Sergei Eisenstein's work was in historical fiction as opposed to historical documentary, he wrote extensively about the key role of montage in building film meaning and the power of film over perception. He stated that montage should be seen "as a means before all else of revealing the ideological conception [of the film]" (Eisenstein 1949: 244). In operationalizing Eisenstein's ideas, *Terminal Time* intends to expose to the audience how montage functions in this "revealing," thereby creating new perceptions of the world based on awareness of ideological conception.

Corporate media dominance

The media experience of today is primarily in the living room, as opposed to the theater or public arena of early Soviet times. The television is the screen. As Lenin realized eighty years ago, the moving image is a powerful tool for propaganda and political control. In today's world of television, he who controls the screen controls the content and form of the programming. While Americans and the U.S. political education system may interpret Lenin's interest in film and the power of mass media in light of his attempts to control the minds of his countrymen and extend the political power of the communist state, few extend that same critique to the contemporary corporate media apparatus which controls American mass media today.

In April 1997, the big four television networks, General Electric (NBC), Westinghouse (CBS), Disney (ABC) and Rupert Murdoch's News Corporation (Fox) were each given six megahertz of the digital broadcast spectrum, enough for each corporate enterprise to create four to six digital channels (McChesney 1997: 21). The cost to the four corporations is the return of their current analog broadcast spectrum to the Federal Communications Commission once the changeover from analog to digital systems is complete. These same corporate entities that control mass media in America today will continue to exert hegemony into the next millennium. Analog systems and broadcast spectra, which will revert to the "public domain," will quickly become obsolete due to new digital equipment standards. Citizen's lobby Common Cause reports that 98 percent of American homes have televisions, that most Americans get most of their news from TV, and asserts that the broadcast industry has the "ability to shape the national news agenda by controlling the messages that TV viewers will and will not see" (Common Cause 1997). It is not a great leap to conclude that the national news agenda and the public presentation of history reflects a combined corporate ideology.

At the same time, public television stations across the United States are struggling to survive. Most stations have already eliminated local production, becoming venues for uncritical social history documentaries, cooking shows and science and nature programming. Public television in America has not generally fostered community involvement or alternative points of view in show production and/or content. Media watcher Robert McChesney points out that public stations in the U.S. are far more inclined than public stations in Canada and Great Britain to reflect elitist culture due to their reliance on local and corporate underwriting (McChesney 1996).

Corporate control of broadcast and cable television has kept innovative and critical historical documentaries, as well as media of other genres, from reaching the public. With the "cookie-cutter documentary" model, the story of any particular moment is presented as the historical truth. Historical data, facts, quotes and imagery are carefully edited to seamlessly produce the narrative. Control of ideological messaging is exerted firstly through choice of subject matter and secondly through the style of narrative production. Both form and content, thus combined, have become codified as the "mass media method" for discussing historical issues.

Enforcement of the code is accomplished by the dominant media apparatus through control of funding and access. Commercial, public and cable television systems exclusively air works produced in the approved format, newspapers review and promote them, and they are favored by popular cinema distribution companies. Ken Burns' "The Civil War" exemplifies adherence to this code; Burns has been well rewarded for his compliance by corporate funding and network access (Litwack 1994: 16–18).

These funders and maintainers of the apparatus, here referred to as "The Generals", include General Electric, General Motors, General Foods and countless other high brass of corporate culture. Their backing, unlike the backing of labor unions, community organizations and issue-oriented groups, is deemed non-political by public agencies such as the public broadcasting system (Potter 1998). Support from "The Generals" insures broadcast and/or wide theatrical release of a media production. Support from tainted groups virtually insures marginality. For example, PBS has denied airplay to works supported by more than 50% by Union based organizations (Potter 1998). Ironically, organizations such as Mobil Oil are seen by organizations such as PBS as ideologically neutral.

Subverting the generals

In 1991, Steffi Domike (one of the three *Terminal Time* producers) and film partner Nicole Fateux turned their attention toward the 1892 Homestead Steel Strike, involving Pinkerton Guards, Andrew Carnegie, Henry Clay Frick and thousands of unknown and for a century unsung community members and workers. They wished to tell the story of the strike from the point of view of the striking works as opposed to the historically over-represented points of view of the rich and powerful Andrew Carnegie and Henry Frick. In 1993, with seed money from unions (United Steelworkers of America, Service Employees International Union and others), the Commonwealth of Pennsylvania and local granting agencies, they released the hour-long film, *The River Ran Red*.

The work intentionally mimicked the dominant form of the historical documentary in an effort to have this story of open and articulated class struggle broadcast to the nation. When the finished product was presented to the producers of the PBS series The American Experience, the producers of *The River Ran Red* were told that although the program looked and sounded very good, the station (WBGH, Boston) already had plans to tell the Pittsburgh story of that period through the life and accomplishments of none other than Andrew Carnegie!

Thus in 1996, three years after *The River Ran Red* was broadcast across Pennsylvania, WGBH's *The Richest Man in the World* enjoyed a national PBS release, using many of the same images, sounds, music, quotes and re-enactments demonstrated in *The River Ran Red*. Yet even using much of the same source material, the differences in editing and narrative structure made the message quite different. The different use of two images, described here, demonstrate how the perspectives of the two filmmaking groups molded the visual and historical record to suit their respective causes.

A portrait of young Andrew Carnegie with his brother Thomas, taken shortly after their arrival in America, is used in both films. *The River Ran Red* cut out Thomas altogether, zooming in slowly to a close-up of the youthful industrialist-to-be: "Carnegie was a poor weaver's son when he left his native Scotland in 1848. By the 1880s he had become one of America's leading industrialists." In contrast, *The Richest Man in the World* uses the same photo full-frame to illustrate the psychological pressures being placed on Andrew Carnegie by his mother in the lean years before migrating to America. The narrator discusses Margaret Morrison's utter embarrassment at her poverty and the failure of her husband to move the family up the local social ladder: "The Boy would have been extremely conscious of this. Andrew would feel the pressure of his mother's shame as well as the preference she showed his brother Tom."

A stereographic image of Homestead, with children in the foreground, the town and mill in the distance, is used in both films to introduce the town. *For The River Ran Red*, the town is introduced directly after a montage on industrial hazards and injuries. "Homestead was radically different. Work in the mill was just as hazardous, but steelworkers had built a powerful union which gave them a say in hiring, wages and how jobs were done." *The Richest Man in the World*, on the other hand, uses the image behind the following: "The town itself was foul. Garland wrote of 'great sheds out of which grim smokestacks rose, with a desolate effect–like the black stumps of a burned forest of great trees."' Interestingly, this exact quote, penned by novelist Hamlin Garland in 1894 (two years after the strike was broken), is used at the end of *The River Ran Red* to build a picture of the ultimate effect of Carnegie's policies on the town.

Clearly, images and words from the past can become re-coded to project whatever the filmmaker desires. With *Terminal Time* we intend to subvert the Generals by turning the "cookie-cutter" loose on the entire past millennium of human history. By incorporating audience feedback, *Terminal Time* allows the audience to manipulate the framing of the documentary and to interrogate its pose of objectivity. We invite the audience to join us in questioning the dominant, ideologically coded mode of producing history.

Interrogating "individual choice"

There is a great deal of industry hype surrounding interactive media and computing. Typically such experiences are promoted through a rhetoric of utopian navigation. According to such rhetoric, the computer provides unlimited access to information and experience, a pure source of empowerment that imposes no interpretation on the data that is processed. Other familiar tropes in this rhetoric include: real-time, immersion and virtuality – promising the thrill of reality or hyper-reality, without the effort, right from one's own PC. Microsoft's ads softly beguile us with the question "Where do you want to go today? ®"

Interaction leaves a trace. The flip-side of utopian navigation is demographic data collection. Especially as more computer-mediated interaction moves into networked environments (e.g. the Web), the very acts of user intentionality, those manifestations of the power of free choice lauded by information technology enthusiasts, have become the raw material for corporate data collection. By collecting, sorting, and categorizing acts of user interaction, corporations hope to sell users ever more precisely targeted products. "Where do you want to go today?" becomes "What do you want to buy today?"

Terminal Time is an exploration of both these dynamics, utopian navigation and demographic data collection. However, *Terminal Time* is not intended as a pure debunking exercise showing that all things interactive are bad. It is certainly the case that information technology has provided easier access to larger amounts of information. In fact, the producers of *Terminal Time* took advantage of the web in doing historical research for the project. Rather than debunking, *Terminal Time* is intended as an exploration of some of the unexamined assumptions and unintended side effects of information technology.

Utopian navigation

In the worldview of utopian navigation, the computer is seen as a value-free conduit, an executor of user agency. Even the use of the word "navigation" is telling - it moves the focus onto the user's movement in some data space and away from the system's active manipulation of that data. The computer is seen as pure communication device, pure medium. Of course in this post-

McLuhan age it is considered a given that a medium is not a passive pipe, but rather the active messenger of a worldview (McLuhan 1964). But the computer as medium has unique properties that can mask this understanding. Two such properties, identified by Janet Murray (Murray 1997), are the participatory and encyclopedic nature of digital environments.

The participatory nature of digital environments means that they take action in direct response to user input. Generally there is only a short lag time between user action and the system's response to the action; the user experiences an immediate gratification of the desire to a affect the system. But this immediate gratification can mask the recognition of the fact that the system's authors have determined the boundaries of this interaction. The system can only reflect the user's actions within the limits of the structures and processes envisioned by the system's designers.

The encyclopedic nature of digital environments means that they have vast capacity. The amount of information in digital environments often exceeds the amount the user can comprehend as a whole. It is impossible to access every record in a database, every document on the web. This enormous capacity is generally coupled to processes that enable access to the stored information, such as search engines and navigation interfaces. This combination of encyclopedic capacity and participatory access can imbue the user with a feeling of great power - all knowledge appears to be at one's fingertips. But the encyclopedic nature of digital environments can mask the recognition that the system's authors have excluded information from the system and prevents the user from asking why only certain forms of interaction are allowed.

Demographics in the electronic landscape

We use the term "electronic landscape" to refer to the immense corporate/institutional networks of interlinked technologies and databases that touch our lives. One need not look too hard for examples: video rental stores often keep digital records of a patron's entire rental history, as America learned during the Robert Bork confirmation hearings in 1987, or as we see on the other side of the U.S. political spectrum with Kenneth Starr's subpoena of Kramerbooks for records of all Monica Lewinsky's book purchases. Yet, it is not merely the individual institutions' usage of these records that is of concern. Data collection achieves its full power when the data is traded between companies, concatenating personal data from many sources into detailed, if Frankensteinian, digital profiles. Recently, the Metromail Corporation, which maintains and sells records from a detailed data base of over 90% of American households, has come under scrutiny for allowing such records to fall into the hands of convicted sex offenders. In the case of Metromail, one randomly selected individual was represented by over 900 pieces of data including address, income, ailments, marital status, hobbies, etc., as well as detailed purchasing habits (Bernstein 1994).

Such examples reveal that within the contemporary electronic landscape, each interaction concatenates to the regime of a virtual data-body, constructed and existing in virtual space. These virtual identities are constantly updated with information about credit ratings, spending habits, video preferences, ATM usage, medical history, driving records and numerous other bits of information. Artist Jeffrey Schulz calls this data space the "identity economy" and notes that "...every telephone call, every withdrawal of money from a bank account, every mail order, every magazine subscription, every visit to a doctor, etc., – creates a potential surplus of demographic identity information" (Schulz 1993: 160).

The internet, particularly the World Wide Web, provides an example of the relationship between utopian navigation and data collection. Network technology enables marketers to monitor a user's activities within a site, as well as terms entered in data-retrieval engines. Detailed web site "registration" processes allow sites to associate browsing behavior with personal information, thus making the information collected even more valuable to advertisers. The idea of uniquely identifying a user has even been pushed into the computing infrastructure itself. Both Intel and Microsoft have had to manage the corporate relations snafu arising from the revelation that Intel's Pentium III microprocessor, and Microsoft's operating system Windows 98 both broadcast a unique machine identifier when connecting to the network (Clausing 1999; Markoff 1999). While ostensibly put in there for "debugging" purposes, such an identifier certainly makes the task of automated demographic data collection easier.

The contemporary landscape is inhabited by many mechanisms to extract data from our pleasures and desires as well as presumably fears and dislikes. As advertisers begin to better utilize non-exclusionary marketing approaches based upon appropriated pluralist discourse and electronic, networked interfaces designed to process more sophisticated blocks of data, our culture approaches an interesting threshold. Here every action is an interface. Here every passing whim or building need may be immediately analyzed for the perfect commodified remedy, suggested by ubiquitous marketers perfectly in accord with our financial assets. At this threshold, all of our subjective interests serve to forcibly fix our position within a marketing database. *Terminal Time* explores this convergence of utopian navigation with demographic data collection by using audience polling to target market histories of the world which are not actually intended or desired.

A democratic, recombinant history

Utilizing indirect questionnaires as a user interface, the system essentially target markets each audience with an appropriate history. Rather than asking audiences what type of history they would like, or how they would like to navigate through history, they are asked questions about their own demographics and psychographics: their work status, what cultural trends they find most disturbing, how well they get along with others, etc. The resulting history holds a funhouse mirror to the audience, reflecting an exaggerated and distorted view of the audience's biases. A sample question follows.

What is the most pressing issue facing the world today?

A. Men are becoming too feminine and women too masculine.
B. People are forgetting their ethnic heritage.
C. Machines are becoming smarter than people.
D. Its getting harder to earn a living and support a family.
E. People are turning away from God.

The most unfamiliar and perhaps unsettling feature of the interaction is that audiences must publicly applaud for their given answers, changing a simple response into a public display. The applause meter was chosen as the input device for two reasons: ease of setup in different venues and the audience dynamic created by public applause. The applause meter requires no special setup in a theater. All that is required is a good quality directional microphone and a small mixing board. Alternative input devices, such as buttons or knobs placed at every seat, would be difficult to install. Such devices would effectively prevent *Terminal Time* from traveling to many venues. More importantly, applause metering enables interesting and entertaining audience dynamics. These interaction dynamics was originally explored in *The Consensual Fantasy Engine*, an interactive cinema piece by Paul Vanouse and Peter Weyhrauch (Vanouse and Weyhrauch 1995).

The applause interaction creates a collaborative, yet competitive relationship with other audience members. The interaction is collaborative in the sense that the phenomena is totally collective, yet competitive because the winning responses will inevitably change the ensuing representation of world history, the very basis from which ethnic, religious and ideological self-awareness has stemmed. With applause, the audience members can gauge how the audience as a whole is responding to questions. During the interactive polls, segments of the audience sometimes compete for control, clapping and shouting to make their choice the winner. At other times, the audience laughs when a choice meets with silence (no one wants to vote for it). Sometimes the applause grows into a groundswell of whistling and clapping as it becomes clear that certain choices are nearly unanimous.

Of course the audience experience is determined not only by the points of interaction, but also by the audience's reaction to the historical narrative produced. The audience recognizes that their interaction has an influence on the historical narrative, but, unlike a utopian navigation scenario, the resulting narrative is not a perfect, transparent response to their interaction. Rather, the narrative escapes their control, producing a story they did not intend, nor desire. As the history begins 1000 years ago, the audience should experience a comfortable sense of historical authority engendered by the familiar documentary form and the remoteness of the historical events. As the history unfolds, the effect of the periodic audience polls becomes more and more apparent. The increased bias evident in the history should begin creating a tension with regard to the veridicality of the history (a sense of "wait a minute, this doesn't seem quite right...").

In order to fully appreciate the piece, an audience should see it more than once. In a typical hour-long performance, an audience will be able to see two performances. In the second viewing, even if the audience answers the polls in exactly the same way, they will experience a different history. In the event that the polls are answered in the same way, the differences will appear in the specific events chosen and the text generated for those events, not in the ideological bias. Seeing two different histories back-to-back makes the effect of ideological bias in historical construction fully apparent. Typically, during the first performance, audiences respond to the questions truthfully, that is, actively trying to reflect their true beliefs in their answers to the questions. During the second performance they tend to respond playfully to the questions, essentially trying on different belief systems to see how this will effect the resulting history. While this could be seen as "game-like" psychographic tourism on the part of the audience, this reaction seems to indicate an understanding of the influence of belief system (as reflected in the answers to the questions) on the resulting history.

Authorship and representation

Terminal Time is informed by a conception of AI as an expressive medium (Mateas 1999; Sengers 1998). Expressive AI conceives of AI systems as cultural artifacts. The concern is not with building something that is intelligent independent of any observer and cultural context. Rather, the concern is with building an artifact that seems intelligent, that participates in a specific cultural context in a manner that is perceived as intelligent. Expressive AI views a system as a performance. Within a performative space, the system expresses the author's ideas. The system is both a messenger for and a message from the author. Expressive AI thus changes the focus from the system as a thing in itself (presumably demonstrating some essential feature of intelligence), to the system as a communication between author and audience. At the technical level of building the artifact, the technical practice becomes one of exploring which architectures and techniques best serve as an inscription device within which the authors can express their message.

As authors, we have specific artistic goals and audience experiences we are pursuing with *Terminal Time*. The project would lose meaning if we could not exert authorial control over the histories generated by the system. Of course, maximum authorial control would consist of writing a fixed set of canned histories; audience interaction would select one of these canned histories. But this extreme of control is inappropriate for this project on several grounds. Conceptually, the project depends on the machine "really constructing" the histories. The critique of the computer as a passive conduit of information requires that the computer actually take on an active role as a semi-cooperative genie, obviously responding to the choices voted on by the audience, but taking these choices to extremes. And on practical grounds, the number of possible histories resulting from all possible answers to all the questions is too large to build by hand. So, even if the conceptual purity of the piece did not demand it, practical necessity would require that the computer play an active role in story construction. As we reject the extreme of pure hand-authoring, we also reject the extreme of strongly emergent architectures, that is, architectures in which as little high-level knowledge as possible is given to the system, with all high-level behavior resulting from large numbers of statistical combinations of low-level elements. Such architectures by definition make authorship highly problematic. In a sense, they provide no authorial "hooks," no places within the architecture in which an author can inscribe her intention, can exert specific control. Much of the architectural work that went into the iterative prototyping of *Terminal Time* was a search for an architecture providing authorial "hooks" on the right level of abstraction: fine-grained enough to allow significant combinatorial possibilities and the capability for surprise, yet providing the appropriate authorial affordances to allow the exertion of authorial control over multiple levels of the story construction process.

The *Terminal Time* architecture

Terminal Time's architecture consists of the following major components: knowledge base, ideological goal trees (Carbonell 1979), rule-based natural language generator, rhetorical devices, and a database of indexed audio/visual elements primarily consisting of short digital movies and sound files containing music (for more architectural details than are provided in this chapter, see (Mateas, Domike & Vanouse 1999; Mateas, Vanouse & Domike 2000)). The architecture is depicted in Figure 2. The knowledge base contains representations of historical events. This is the raw material out of which the ideologically-biased histories are constructed. Examples of historical events are the First Crusades, the invention of Bakelite, and the rise of enlightenment philosophy. Ideological-goal trees represent the current ideological-bias being pursued by the narrator. The goal-trees consist of rhetorical goals ordered by subgoal and importance (to the ideologue) relationships. These goals are used both to select historical events to include in the story and to "spin" the event in an ideologically-consistent manner. The rule-based natural language generator (NLG) generates the narrative text once specific facts have been selected and connected to make a story. The storyboard serves as a working memory for processes that impose a narrative order on event spins created by the goal tree. Rhetorical devices are connecting pieces of text with accompanying constraints on story structure. These devices are used to create narrative connections between historical events. Finally, the multimedia database contains the audio/visual elements for the assembled documentary. Once a narrative track has been constructed, information retrieval techniques are used to match the "best" indexed multimedia elements to the appropriate pieces of text. Once the multimedia elements have been selected, the resulting documentary is displayed, layering text-to-speech synthesis of the narrative track, and the video and audio elements.

The audience's responses to the questions influence the machine by selecting and editing rhetorical goal trees, selecting a set of rhetorical devices, and placing constraints on the storyboard. In a sense, the audience response parameterizes the machine. The responses activate structures and processes; the machine then autonomously generates a biased history.

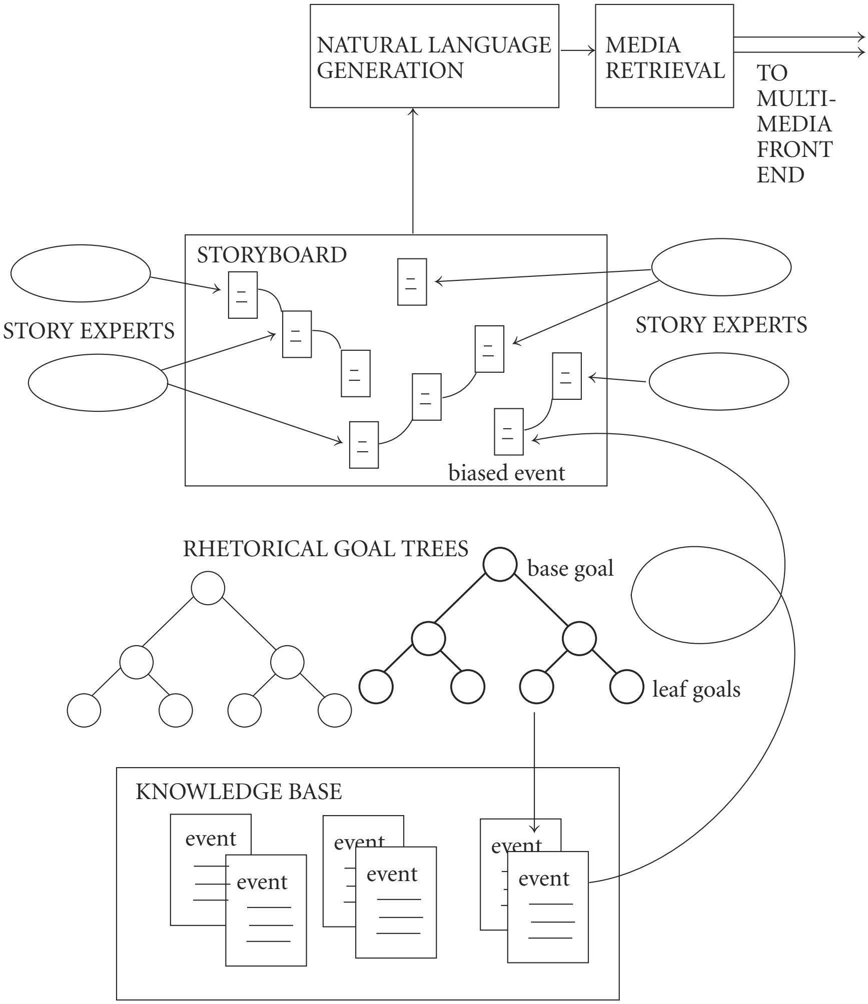

Figure 2. *Terminal Time* architecture.

The knowledge base consists of higher order predicate statements about historical events, definitions of ontological entities used in the historical event descriptions, and inference rules. *Terminal Time*'s ontology is based on the Upper Cyc Ontology, the top 3000 most general terms in the Cyc ontology (Lenat 1995). The upper ontology provides a useful set of distinctions in terms of which the more specific ontology needed by *Terminal Time* can be defined. Figure 3 shows the representation of the historical event "The Giordano Bruno Story."

```
($isa %GiordanoBrunoStory %HistoricalEvent)
($isa %GiordanoBrunoStory %IdeaSystemCreationEvent)
($isa %GiordanoBrunoStory %Execution)
(%circa %GiordanoBrunoStory (%DateRangeFn
(%CenturyFn 16) (%CenturyFn 17)))
($eventOccursAt %GiordanoBrunoStory $ContinentOfEurope)
($performedBy %GiordanoBrunoStory %GiordanoBruno)
($outputsCreated %GiordanoBrunoStory %GiordanoBrunosIdeas)
($isa %GiordanoBrunosIdeas $PropositionalInformationThing)
($isa %GiordanoBrunosIdeas $SomethingExisting)
(%conflictingMOs %GiordanoBrunosIdeas %MedievalChristianity)
($isa %GiordanoBrunosIdeas %IdeaSystem)
($performedByPart %GiordanoBrunoStory
 %TheRomanCatholicReligiousOrg)
($objectActedOn %GiordanoBrunoStory %GiordanoBruno)
```

Figure 3. Formal representation of the Giordano Bruno story.

The formal representation of historical events is manipulated by processes (described below) which select events for inclusion in a story, produced biased spins of events, link spins together into narratives, and generate narrative text.

Terminal Time organizes ideological bias with goal trees, adapted from the ideological reasoning program Politics (Carbonell 1979). These goal trees represent the rhetorical goals of an ideological story-teller. For example, the Hard-core Anti-religious Rationalist has as one of its top level goals *show that religion leads to evil*. Two subgoals are *show that religion causes war* and *show that religion causes oppression*. Audience interaction defines and modifies the current active goal tree, possibility adding, deleting, or changing goals. Two different ideological positions can be mixed by combining goals from two goal trees. For example, the audience's answers to the first set of questions may select the Hard-core Anti-religious Rationalist goal tree. Answers to the second set of questions may determine that racial equality (exaggerated as a homogenized "Bennetton commercial" multiculturalism) is a sub-theme. The goal tree is modified to include the Corporate Multiculturalist goals in addition to Hard-core Anti-religious Rationalist goals, thus producing a hybridized ideological narrative. Some responses to questions (particularly questions in the third and last set) modify the tree more subtly, adding and removing individual goals in the tree.

These goal trees scan through the knowledge base to select and produced biased spins of events for use in a story. The "spun" events are put into a conceptual container called the storyboard. Rhetorical devices then connect the event spins into a narrative structure. Rhetorical devices are sentences (actually declarations of NLG rules and arguments) that can connect episodes or collections of episodes together to create a story flow. For example, the sentence "Yet progress doesn't always yield satisfaction" can be used to connect several episodes describing the positive effects of technological progress and several episodes describing social or environmental problems arising from technological progress. Associated with the English sentence is a formal representation constraining the meanings that episodes before and after the rhetorical device can have. For example, "Yet progress doesn't always yield satisfaction" has constraints specifying that everything preceding the rhetorical device must be positive technological, artistic, or industrial progress, and that everything following the rhetorical device must be negative effects of progress.

Once a collection of spins has been connected together by rhetorical devices, the resulting story, which at this point still consists only of formal representations, is sent to the natural language generator to produce the actual narrative text. In addition to generating text, the natural language generator associates index terms with each generated sentence. These index terms are used to retrieve appropriate movie and sound clips from a term-indexed multimedia database. Even though the mechanisms linking images to narrative are less sophisticated than the mechanisms producing the narrative, the Kuleshov effect ensures that the resulting juxtaposition of image and narrative will still make sense to the audience.

This architecture was arrived at through an artistic as well as technical exploration. We desired an architecture that creates narratives rendering our authorial intent without necessarily portraying our own ideological viewpoint. Through such an architecture we can see stories created that might involve unusual causal relationships or unexpected conclusions, that, while satisfying us as authors, go beyond our own conceptions. Additionally, the history construction process captured in the software architecture is itself of conceptual interest. We see it as a caricature of ideological thought and "cookie-cutter" documentary construction, an explicit comment on the mechanical nature of shallow ideological reasoning. Our engagement with AI in the *Terminal Time* project is a concrete example of expressive AI. The AI architecture serves the needs of, and simultaneously informs, our artistic intent.

Conclusion

Terminal Time interrogates three cultural constructs: the naturalization of history in the historical documentary, the rhetoric of choice in cyberculture, and representations of knowledge and intelligent activity in Artificial Intelligence research. Our critique makes full use of the resources available in the very cultural fields under discussion. We explore the naturalizing tendency of the documentary using the filmic grammar of the documentary, comment on utopian navigation using interactive technologies, and point the way to an alternative conception of AI by building an AI program. The self-referential use of cultural resources and naïve hubris are the defining characteristics of the creative process employed in building *Terminal Time*.

References

Barnouw, Erik (1983). *Documentary: A history of the non-fiction film*. New York: Oxford University Press.

Bernstein, Nina (1994). Lives on file: Privacy devalued in information economy. *The New York times on the web*, June 12, 1–11.

Bordwell, David & Kristin Thompson (1997). *Film art*. New York: McGraw-Hill.

Carbonell, Jaime (1979). *Subjective understanding: Computer models of belief systems*. Ph.D. Thesis, Computer Science Department, Yale University, Research Report #150.

Clausing, Jeri (1999). Privacy groups seek recall of Intel chip. *The New York times*, January 29.

Common Cause (1997). Your master's voice. *WIRED*, August 1997, 45.

Eisenstein, Sergei (1949). Dickens, Griffith and the film today. In J. Leyda (Ed.), *Film form*. New York: Harcourt, Brace & World. Cited in Grindon, p. 22.

Grindon, Leger (1994). *Shadows on the past: Studies in the historical fiction film*. Philadelphia: Temple University Press.

Lenat, Doug (1995). Cyc: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, *11*, 32–38.

Litwack, Leon (1994). The Civil War. In S. Dolan (Ed.), *Telling the story: The media, the public and American history*. Boston: New England Foundation for the Humanities.

Markoff, John (1999). Microsoft to alter software in response to privacy concerns. *The New York times*, March 7.

Mateas, Michael, Steffi Domike, and Paul Vanouse (1999). Terminal Time: An ideologically-biased history machine. *AISB quarterly: Special issue on creativity in the arts and sciences, Summer/Autumn 1999(102)*, 36–43.

Mateas, Michael, Paul Vanouse, and Steffi Domike. (2000). Generation of ideologically-biased historical documentaries. In *Proceedings of AAAI 2000*. Austin, TX, 236–242.

Mateas, Michael (2001). Expressive AI: A hybrid art and science practice. *Leonardo: Journal of the International Society for Arts, Sciences, and Technology, 34* (2), 147–153.

McChesney, Robert W. (1997). The digital TV heist. *In these times*, *21* (13).

McChesney, Robert W. (1996). *Corporate media and the threat to democracy*. New York: Seven Stories.

McLuhan, Marshall (1964). *Understanding media: The extensions of man*. New York: New American Library.

Murray, Janet (1997). Hamlet on the Holodeck: The future of narrative in cyberspace. Cambridge, MA: The MIT Press.

Potter, Chris (1998). Too hard hat to handle? PBS's double standards with labor underwriting. *Pittsburgh City Paper*, March 11, 8.

Sengers, Phoebe (1999). Designing comprehensible agents. In *Sixteenth international joint conference on Artificial Intelligence: Vol. 2* (pp. 1227–1232).

Schnitzer, Luda, Jean Schnitzer & Marcel Martin (Eds.) (1973). *Cinema in revolution: The heroic era of the Soviet film.* New York: Hill and Wang.

Schulz, Jeffrey (1993). Virtu-real space: Information technologies and the politics of consciousness. In Simon Penny (Ed.), *Machine culture. Visual proceedings: The art and interdisciplinary programs of SIGRAPH 93*. New York: Association for Computing Machinery.

Vanouse, Paul & Peter Weyhrauch (1995). *The Consensual Fantasy Engine: An audience-driven interactive fiction*.

Chapter 10

Experiments with the theatrical Greek chorus as a model for interactions with computational narrative systems

Carol Strohecker, Kevin M. Brooks, and Larry Friedlander
Media Lab. Europe, Dublin; Motorola, Cambridge, MA; Stanford University, CA

Introduction

This chapter presents explorations of narrative systems with which users interact in manners reminiscent of the chorus in ancient Greek theater. We review structure and roles of the theatrical chorus, describe how these informed an interactive narrative prototype for a single-user platform, and conclude with a "what-if?" inquiry for systems in which multiple simultaneous interactors use tangible or otherwise palpable computational objects to facilitate and reflect actions and conversation.

The first experiment is largely structural: players' interactions unfold progressively finer detail about the story, but do not change the course of events or fundamentally modify the characters (Strohecker et al. 1999). Players interact mainly by querying representations of chorus members, who comment on the narrative from different perspectives. Here we consider how the chorus model might be pushed further, such that viewers *become* chorus members in some more direct sense, thereby contributing to the choral aspect of the narrative system as it grows over time. The discussion includes the psychological process of introjection and how it could guide thinking about systems that change based on players' interactions. The usefulness of considering this psychological process with regard to systems development has to do with humans' affinity for manipulating objects and imbuing them with meanings that can be shared or personal.

An important assumption is that the psychoanalytic tradition has helped to articulate processes of human thinking that need not, and often should not, be considered pathological. Many of the cases reported in the literature describe reasonable responses to unusual situations. Furthermore, Winnicott (1971a, b) and Fairbairn (1963) are careful to note that they believe the particular behaviors and thinking patterns cited here (i.e., the use of transitional objects and the phenomenon of object-splitting) to be among normal human developmental processes.

A similar qualification has to be made for the reference to Piagetian research. We move beyond Piaget's early notion of stage theory to approaches of post-Piagetians who focus on individual differences and the importance of social context in learning (Turkle & Papert 1992; Harel & Papert 1991).

Another assumption concerns the use of physical objects as devices facilitating interactions with computational systems. It would be problematic to assert that a physical object can be designed so particularly as to be obviously suited to a specific purpose: Norman (1988)'s thesis is countered by discussions of the *bricoleur* who collects objects and adapts them to various purposes as the need arises; the purpose may very well differ from that for which the object was originally designed (Levi-Strauss 1966; Pirsig 1974/1984; Turkle & Papert 1992). This human capability to reinvent mandates allowance of a range of creative, unpredictable uses for any given object, but does not deny that an object may indeed also be useful for whatever purpose its designer may have intended or desired. The rudimentary design principles considered here rely on variability of taste and concept as different people choose and use objects for interactions with narrative systems. Nevertheless the range may also include commonalities of interpretation among interactors. Both results would be useful and potentially interesting.

The theatrical chorus

Forms of dramatic chorus appear in contemporary works of various cultures. In performances by Ladysmith Black Mambazo, for example, the chorus address audience and actors, providing both narrative continuity and musical entertainment. In the film *Little Shop of Horrors* (Oz 1986), a chorus of singer/dancers appears from time to time, amplifying story events and characters' emotional states. Another film, *Annie Hall* (Allen 1977), embodies a chorus-like function in textual commentary rather than personifications. In the balcony scene, Annie and Alvy chat about wine and tennis while a literal subtext is displayed at the bottom of the screen, comically revealing inner monologues replete with worries that the other person will not find the speaker attractive. Such variations exemplify the wide range of roles that the chorus can play, providing a rich base of potential functions for computational narrative systems.

In ancient Greek theater, the chorus evolved through several eras and incarnations. Initially the chorus were separate from the actors in role, location, and appearance, but over time they merged more and more with the actors and the action. At first a collective, singing, dancing, and speaking together, chorus members gradually emerged as individual speakers. Often, a chorus leader spoke for the group or provided coherence for their diverse expressions. Originally anonymous, their uniform identities accentuated by masks, chorus members gradually appeared as personalities with particular views on issues and events. And, beginning as a theatrical device whose function was to express mood or tone, and whose contribution was to witness, comment, or clarify, the chorus gradually shared responsibility for delivering pieces of the narrative.

In Sophocles's *Antigone* (Corrigan 1965), the chorus speak both singly and collectively as they comment on the moral dilemma. Interestingly, their change of view through the course of the drama suggests the sort of character development that we would normally expect an individual actor to represent. The chorus of elders begin by describing the conflict between Oedipus's sons. The elders side with Creon in his decree that Eteocles should receive a decent burial but Polyneices should not. Creon is king, they declare, and his word should be the respected law of the land. Here the chorus speak separately but hold one view. Even as they assert it, though, they acknowledge the complexity of the issue by reserving the possibility that their view could be flawed, affected by vulnerabilities of old age. Indeed, they waver as it becomes increasingly clear that both Antigone's and Creon's arguments are substantive. Eventually realizing that Creon is behaving rashly, they continue to side with him but again allow that old age may predispose them to a certain foolishness. They begin cultivating a way out of their stated position by articulating two aspects of the dilemma: "You acted for the good," they tell Antigone," but disobeyed the law." "What moral law have I disobeyed?" she asks.

The Law, it seems, may have more than one face. Originally siding with human law, the chorus now understand and respect "good," the law of the gods. They urge Creon to relinquish his decree, but it is too late. Faced with sweeping tragedy, the elders condemn Creon and manage to frame the situation in terms of themselves – they, like Creon, have changed through the horrific events: "proud men in old age learn to be wise." They have managed to both reverse and maintain their original position. Still advocating compliance with Law, they emerge on the side of Good. Finally, their age no longer contributes to frailty, but to wisdom.

Thus the chorus may personify, clarify, magnify, subdue, transpose, interpret, retell, frame, or give perspective to the narrative action. Chorus members may fill in "holes" in the narrative, offer commentary, foreshadow the action, and reflect the action by re-enacting it in other modalities. The chorus may act as an intervening layer that protects the audience from incidents too horrifying to directly experience, such as the multiple deaths in *Antigone* (Friedlander & Strohecker 1995). In that awful tale, the chorus are the only survivors.

Tired of Giving In: An initial prototype

United States history offers a latter-day Antigone in the figure of Rosa Parks, a Black woman who refused to give her seat to a White man on a segregated public bus.[1] The time was 1955, the place was Montgomery, Alabama, and the event became a milestone in the American Civil Rights Movement.

One widely distributed textbook version of this story describes Parks, who was 41 years old at the time, as an aging woman who was so tired after a long day at work that she didn't want to give up her seat (Kohl 1995; Mayer 1995). This telling neutralizes her act of civil disobedience. Parks has addressed the misrepresentation by asserting, "The only tired I was, was tired of giving in" (Parks & Haskins). We emphasize her statement through the title of our retelling, *Tired of Giving In* (TOGI) (Brooks 1996; Strohecker 1996; Strohecker 1997 a, b).

Some versions of the story do acknowledge Parks for her courageous act but over-emphasize its individual nature. Many members of Montgomery's Black community had experienced discrimination, particularly on public transportation, and many had resisted in one way or another. The community was well organized through a network of churches and the efforts of dedicated activists. By 1955, members of the NAACP and other local groups were waiting for a legal case that could serve as a test of the segregation laws.[2] They nearly found one several months before the Parks incident, when a teenaged girl named Claudette Colvin was arrested for the same offense. However, because Colvin had resisted the police and was expecting a child out of wedlock, community strategists felt that her case would not be able to withstand the publicity associated with a trial of such importance (Blackside 1987; Robinson 1987). They decided to wait.

Some accounts acknowledge this bit of strategy while emphasizing the fact that Parks was secretary of the local NAACP and worked closely with its president, E. D. Nixon. It was Nixon who had decided not to pursue the Colvin case. These accounts suggest that Parks's resistance on December 1 was planned, as Nixon and others felt that the time had come to challenge Jim Crow,[3] and that Rosa Parks was a strong character who could withstand public scrutiny and represent the Black community well. Such accounts laud Parks but diminish her remarkable act of bravery. As a community activist, Parks was better prepared than most to follow through with such an act. She must have realized the likelihood of being arrested when she refused to give up her bus seat. Nevertheless, when the situation occurred – and it occurred spontaneously (Parks & Haskins 1992) – she handled it with quick thinking, dignity, and courage.

Our account attempts both to extol Parks as an individual and to explicate the power that had built up in the community through so many years of abuse, organization, and resistance. Ours is a story not just of one brave individual, but of many, who came to realize that by acting together they could change their world. Thus the TOGI chorus include members of 1955 Montgomery's Black and White communities, who are struggling with the issue of segregation. These are the members of Chorus Present. There are also two other groups in the TOGI chorus. Members of Chorus Past are Africans taken as slaves during colonial times, who look to the 1955 events and wonder how they could ever come to be.[4] Members of Chorus Future are young urban dwellers of today, who look to the 1955 events and wonder whether anything has really changed. Each choral category includes perspectives identified as "positive," "neutral," or "negative" with respect to the bus boycott.

TOGI's characters include notable members of the Black and White communities. The portrayals of Rosa Parks, Claudette Colvin, E. D. Nixon, Martin Luther King, Jr., Jo Ann Robinson (president of the Women's Political Council), and Fred Gray (attorney for Rosa Parks and later, for the Montgomery Bus Boycott), derive from autobiographies and other accounts of the boycott and related events. Other characters are more fictionalized, though they also are based on accounts of and by people in Montgomery at the time: Tacky Gayle (mayor of Montgomery), Clantello Bagley (manager of the City Bus Lines), J. P. Blake (the bus driver who had Rosa Parks arrested), and the arresting officers.

TOGI narrative structure

TOGI opens as the chorus members introduce themselves. Viewers can watch this introduction or skip to the first scene of the story. The narrative proceeds through four interactive scenes: the *TOWN* of Montgomery, the *BUS* on which Rosa Parks is arrested, the *JAIL* cell to which she is taken, and the Holt Street *CHURCH* where the boycotters congregate. Alternating with these scenes are refrains in which chorus members chant about moments in the story.

People often say that a story "unfolds," and that is the way we frame interactions with TOGI. By selecting areas of scene images and querying depictions of characters and chorus members, viewers unfold details of plot and character. The story can be revealed through six "folds":[5]

Fold 0. Viewers do not have to interact with TOGI. Viewers who let it play on its own hear the "default" story, told by a chorus member who sees the events from today's perspective. She thinks that Parks and others who participated in the Montgomery Bus Boycott set an example that remains relevant. Each of the four scenes opens with a bit of her narration.

Fold 1. Clicking part of a scene image, or module, reveals an associated set of characters and chorus members. The default story is put "on hold" as the characters begin talking with each other, revealing additional aspects of the story. If the viewer interacts no further, the characters' dialog plays out. Then the program returns to the point at which the viewer intervened, and the program continues as in Fold 0.

Fold 2. As the Fold 1 characters' dialog proceeds, the viewer may click on one or more of the available chorus members. Activated chorus members interject comments within the dialog. One of our concerns in scriptwriting was to key the meaning of each potential choral interjection to a corresponding chunk of the characters' dialog. If the viewer activates more than one chorus member, the multiple comments play out in queue. Then the Fold 1 dialog resumes. The tonal effect is colorful, a bit like talk radio, and the visual presentation in this early prototype is like an illustrated storybook. Still images and movie clips pertinent to the dialog come and go, superimposed over the pictorial background of the scene.

Fold 3. If, during the Fold 1 dialog, the viewer clicks a character rather than a chorus member, the character turns to the chorus member whom the viewer has queried most frequently up to that point in the program. This character and chorus member exchange remarks relevant to the current dialog, and after their remarks play out, the dialog resumes.

Fold 4. One chorus member reveals a "graffiti wall" associated with the scene. Here the program is most true to the chorus metaphor: by adding his or her own comments, the viewer in effect joins the chorus. In this early prototype the comments are typed; ideally they would be spoken. Comments are marked by glyphs that subsequent viewers can query to see the text that other viewers have added.

Fold 5. Another chorus member reveals references citing sources of information and media used in the presentation.

As TOGI plays and the viewer interacts, the program tracks the number of times each chorus member is queried. A second tally tracks the viewer's interest in a given perspective (positive, neutral, or negative). When the viewer activates a new module, the program determines which of the module's relevant chorus members to make available, focusing on those seen less often than others in the overall presentation. Less-seen chorus members, and their attitudinal counterparts in other choral sections, have a greater chance of being available for interactions. In this way the program tries to balance perspectives offered to the viewer. If a viewer continuously queries a chorus member who speaks against the boycott, for example, the program is likely to make available a chorus member who speaks for it (and vice-versa). Thus the history of interactions helps ensure access to a full range of views.

Further work

We are considering further experiments for multiuser systems, consistent with the chorus as "the human collectivity confronting the event and seeking to understand it" (Barthes 1985). This aim recalls the notion of "computers as theatre," a phrase that emphasizes the use of metaphor, so well understood in theatrical contexts, in developing graphical user interfaces and other experiential aspects of computational systems (Laurel 1991). In later work, Laurel experimented more explicitly with theatrical spaces supported by computation (Bates 1992; Laurel et al. 1994). Our designs reflect similar contexts for users whose collective interactions help shape the system as it grows over time, and whose interactions are supported by tangible or otherwise palpable objects.

One form for such experimentation could be room-scale spaces with handheld computational objects that mediate communication between users and to the narrative system. Another form could be virtual spaces with multimodal, abstract representations of objects. In either case the ability of the objects to facilitate interactions depends on a fundamental human process, "introjection." We consider *objects* as things that people use not just in the physical world, but also in their minds (Strohecker 1991). Developmental perspectives help to clarify this premise.

Thinking with objects

Psychological theorists describe the phenomenon of introjection (also called "incorporation," "internalization," or "identification") in various ways. In Freud's model, the infant's mind begins as the id, which later gives rise to the ego. Still later, the conscience, or super-ego, forms as the child's ego incorporates the parent's – that is, "takes [the parent's ego] up into itself" (Freud 1933/1965: 56). Fairbairn discusses "internalization of the object" in terms of an infant's coming to understand the alternating presence and absence of its primary caregiver. When alone, the infant keeps in mind some internal representation of the figure. The representation is simple at first, but since the external version of the object sometimes pleases the infant and sometimes does not, the object comes to be perceived as having a double nature. These two aspects "split off from the main core of the object" (Fairbairn 1963: 224). Fairbairn sketches a scenario of progressive splitting of internal objects as the external separation is enacted.

Klein emphasizes emotional connections between internal objects as the infant structures relationships between its ego and family figures (Sutherland 1989: 37). Winnicott also describes formation of emotionally charged inner objects, as the infant moves from a sense of being merged with the mother to a sense of autonomy as a separate individual (Winnicott 1971a, b). "Transitional objects" assist in the process. The child substitutes something tangible (such as a blanket, teddy bear, or favorite toy) for the physical closeness to another person, which by necessity lessens as the child grows. The substituted object is a sign of psychological incorporation of the person, which enables the child to tolerate loss of the external relationship. What develops in its stead is an *internal* relationship, with a representation of the loved one.

Objects can play a comparable role in situations that do not involve a significant loss. Papert describes the importance of gears in his early thinking:

> I became adept at turning wheels in my head and at making chains of cause and effect ... I believe that working with differentials did more for my mathematical development than anything I was taught in elementary school. Gears,

> serving as models, carried many otherwise abstract ideas into my head. (Papert 1980: vi)

Similarly, Piaget describes the importance of actions – observable, physical activities – as the internalized operational "glue" that holds together certain structured understandings (Beth & Piaget 1966: xvi).

Thus actions with objects – doing things to and with objects – may constitute a crucial aspect of human learning and thinking. Many people prefer thinking with objects, moreso or rather than with abstractions (Turkle & Papert 1992). Objects may facilitate transitions from one thought to another, or one emotional state to another, at any age. Perhaps those who prefer this style of thinking would be especially inclined toward commentary on narratives through the use of manipulable objects.

Chorus members' object-based interactions with narrative systems

Game and film manufacturers often produce supplemental media to extend their audiences' experiences of narratives. For example, Nintendo publishes magazines that include solutions, character descriptions, and accounts by skilled players. The experience of a game is not limited to sessions in which the player is actually engaged with the machine, but extends through reading and discussion at other times. Relevant media include game cartridges and magazines, even dolls or other facsimiles of characters and objects within the game. Similarly, the producers of *Toy Story* released a line of dolls and other toys to augment the cultural impact of the film (Lasseter 1995). Involvement with the narrative includes not just suspension of disbelief, identification with characters, and other processes related to film-watching, but creative processes involved in play with objects that can be held, moved, and transformed.

Strategies for extending a narrative context through use of a range of media tacitly acknowledge people's use of objects to mediate thought. They also pave the way for development of objects that augment a narrative context not just in players' minds, but between players and within a computational system.

Objects as stand-ins

The theatrical chorus are concerned primarily with commentary: the collective enriches the audience's experience of the narrative through verbal, often musical, responses to dramatic events. We could imagine objects associated with a narrative system as being recorders, transmitters, and/or players of spoken or chanted commentary. However, filtering and editorial functions would likely soon be needed in order to manage the accumulating data, and despite progress in the field of computational linguistics, the effort could be daunting and the results disappointing. Another approach could be to designate a live coordinator, like an emcee or editor, to orchestrate or constrain players' interactions (e.g., Laurel et al. 1994). Indeed, this role bears some resemblance to that of the traditional chorus leader. However, using physical objects as mediating devices suggests still another approach, in which carefully chosen or designed objects cue timing and content of comments.

Objects as interlocutors

Choral comments typically occur at a meta-level: the narrative proceeds as characters enact their roles within the structure of the plot; the chorus observe the enactments and comment on meanings and outcomes. This function is well served by alternating time and/or space, and lends itself to re-tellings of stories that the audience already knows. The chorus do not change the course of events, but enrich understanding of it through multiple perspectives. Thus the chorus may be best suited to stories in which many truths are possible, such as the moral dilemmas in tales like *Antigone* and the arrest of Rosa Parks (Strohecker et al. 1999). Discussion of many interpretations of such stories is essential to appreciating them. Sharing objects that pertain to the story could facilitate discussion, becoming "conversational props" that enliven participants' experiences of the narrative and emphasize the role of the chorus (Bellamy et al. 1994).

Objects could also be the means through which chorus members engage with the narrative system. The timing of an object's appearance could create shared moments of interest that help to coordinate interactions, and the form and function of the object could evoke certain kinds of responses. For example, coins, a bus driver's hat, and a policeman's badge are objects that could facilitate changes of perspective within the Rosa Parks story. An interactor holding the police badge may shift focus from the morality of personal rights to the morality of civic responsibility. An interactor with the coins may take the opportunity to deposit them at the front of the bus, only to be told to exit and re-enter from the back, thereby dramatizing an important theme of the story.

Objects as transformers

Imagine a scenario in which chorus members are represented as masks that interactors adopt as they enter the computational domain. While not mandating any particular comment, the representational quality of each mask could suggest a view or mood to which the player would respond. The appearance of the mask could change through successive enactments, as interactions accumulate. For example, a player's speech prosody may affect the facial expression of the mask (e.g., Cahn 1999). It may even cause a new digital mask to spin off as a prop for future chorus members.

Interactions could register in various forms: for example, the system may associate manual gestures or facial expressions with certain emotions (e.g., Pinhanez 1999; Wilson & Bobick 1999; Wren 1999). Objects could collect and transmit information about who is using them, how, when, for how long, etc. (e.g., Brave et al. 1998; Resnick 1998).

Multiple interactors could convene in a real-world "smart room" sort of theatrical space, in a local-net multiuser virtual domain, or via a web-based graphical environment. Each of these milieus implies a different scale in terms of the number of simultaneous participants, and each scale would necessitate particular strategies for coordinating the multiple inputs, but the basic modes of interaction could apply to each.

Ultimately, the narrative system would reflect a plurality of influences. Its choral aspect would become richer over time, as cumulative feedback formed a collective construction, perhaps reminiscent of phenomena like the AIDS quilt. Interestingly, this notion of multiple voices shaping the system also reflects contemporary models of mind (Bakhtin 1981; Minsky 1986; Mondykowski 1982; Wertsch 1991).

Acknowledgments

Thanks to Marc Davis, Lise Motherwell, Warren Sack, and members of the Narrative Intelligence reading group at the MIT Media Laboratory, as well as Adele Goldberg, David Liebs, and others at Neometron, Inc., for helpful discussions. Michael Mateas and Phoebe Sengers provided feedback that helped to improve this paper.

Notes

1. Whereas current vernacular favors identification of groups by culture or nationality of origin, our designations of racial groups reflect usage during the time of the Civil Rights Movement.

2. NAACP stands for "National Association for the Advancement of Colored People."

3. "Jim Crow" is a term for the system of segregation laws.

4. We struggled with how best to represent and balance race, gender, and culture in the three sections of the chorus. In our absorption with showing varying responses to enslavement, for example, we omitted another representation that would have been appropriate for the Chorus Past, that of a White slave trader. But in other cases we were able to incorporate a more complete picture. For example, our characterization of SallyJo as "ambivalent" enabled allusion to a role that many White women played in the Montgomery Bus Boycott. Some were genuine sympathizers, and others just wanted their housemaids to continue working, but for whatever reasons many surreptitiously offered car rides to pedestrian boycotters. SallyJo waffles but ultimately changes her attitude during the presentation, joining the ranks of the pro-boycotters. Similarly, our characterization of Jonah as "systemized" enables a more fluid treatment of his personage: at the critical moment just before Rosa Parks is arrested, he flees the bus (as people sometimes did under such circumstances).

5. We have numbered the folds for purposes of designation, but as the text explains, the folds are not necessarily sequential.

References

Allen, Woody (Dir.) (1977). *Annie Hall.* United Artists.

Bakhtin, M. M. (1981). Michael Holquist (Ed.), Caryl Emerson & Michael Holquist (Trans.) *The dialogic imagination: Four essays by M. M. Bakhtin.* Austin: University of Texas Press.

Barthes, Roland (1985). *The responsibility of forms: Critical essays on music, art, and representation.* Berkeley: University of California Press.

Bates, Joseph (1992). Virtual reality, art, and entertainment. *Presence, 1(1),* 133–138.

Bellamy, Rachel K. E., E. B. W. Cooper, & Richard D. Borovoy (1994). Supporting collaborative learning through the use of conversational props. *Proceedings of the east-west conference on human-computer interaction*, 181–191.

Beth, Evert Willem & Jean Piaget. (1966). W. Mays (Trans.) *Mathematical epistemology and psychology.* Synthese Library, vol. 12. Dordrecht-Holland: D. Reidel.

Blackside, Inc. (1987). Program 1: Awakenings. *Eyes on the prize: America's civil rights years, 1954–1965.* Boston: PBS Home Video.

Brave, Scott, Hiroshi Ishii & Andrew Dahley (1998). Tangible interfaces for remote collaboration and communication. *Proceedings of CSCW '98* 169–178.

Brooks, Kevin M. (1996). Do story agents use rocking chairs? The theory and implementation of one model for computational narrative. In W. Hall & T. D. C. Little (Eds.), *Proceedings of ACM multimedia'96,* 317–328.
Cahn, Janet (1999). *A computational memory and processing model for prosody*. Doctoral dissertation, Media Laboratory, Massachusetts Institute of Technology.
Corrigan, Robert W. (Ed.) (1965). *Sophocles.* New York: Dell Publishing.
Fairbairn, W. R. D. (1963). Synopsis of an object-relations theory of the personality. *International journal of psycho-analysis, 44*, 224–225.
Freud, Sigmund (1933/1965). Lecture XXXI: The dissection of the psychical personality. In Strachey, James (Trans., Ed.), *New Introductory Lectures on Psychoanalysis* (pp. 51–71). New York: Norton.
Friedlander, Larry & Carol Strohecker (1995). The Greek chorus as a model for agents in interactive stories. *Working notes, AAAI spring symposium, Interactive story systems: Plot and character*, 129–133.
Harel, Idit & Seymour Papert (Eds.) (1991). *Constructionism.* Norwood, NJ: Ablex.
Kohl, Herbert R. (1995). The story of Rosa Parks and the Montgomery Bus Boycott revisited. In *Should we burn Babar?: Essays on children's literature and the power of stories.* New York: New Press.
Lasseter, John (1995). *Toy Story*. Walt Disney Productions and Pixar Animation Studios.
Laurel, Brenda (1991). *Computers as theatre.* Reading, MA: Addison-Wesley.
Laurel, Brenda, Rachel Strickland & Rob Tow (1994). Placeholder: Landscape and narrative in virtual environments. *Computer graphics, 28* (2), 118–126.
Lévi-Strauss, Claude (1966). *The savage mind.* Chicago: University of Chicago Press.
Mayer, Henry (1995). Turning fact into fable: Herbert Kohl's essays examine the politics of storytelling. *New York times book review,* August 20. Late Edition Section 7 p. 16 Column 3 978 words.
Minsky, Marvin (1986). *The society of mind.* New York: Simon and Schuster.
Mondykowski, S. M. (1982). Polish families. In M. McGoldrick, J. K. Pearce & J. Giordano (Eds.), *Ethnicity and family therapy.* New York: Guilford Press.
Norman, Donald A. (1988). *The psychology of everyday things.* New York: Basic Books.
Oz, Frank (Dir.) (1986). *Little shop of horrors.* The Geffen Company.
Papert, Seymour (1980). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books.
Parks, Rosa & Jim Haskins (1992). *Rosa Parks: My story.* New York: Dial Books.
Pinhanez, Claudio (1999). *Representation and recognition of action in interactive spaces.* Ph.D., Media Laboratory, Massachusetts Institute of Technology.
Pirsig, Robert M. (1974/1984). *Zen and the art of motorcycle maintenance: An inquiry into values.* New York: Bantam Books.
Resnick, Mitchel (1998). Technologies for lifelong kindergarten. *Educational technology research and development, 46* (4), 43.
Robinson, Jo Ann (1987). *The Montgomery bus boycott and the women who started it.* Knoxville: University of Tennessee Press.
Strohecker, Carol (1991). *Why knot?* Ph.D. diss., Epistemology and Learning Group, Media Laboratory, Massachusetts Institute of Technology.

Strohecker, Carol (1996). *Tired of Giving In:* Experimenting with the Greek chorus as a model for interaction with stories. In *Nov'Art: Les états généraux de l'écriture multimedia* (pp. 94–97). Paris: ART3000.

Strohecker, Carol (1997). An interface metaphor and mechanism for learning history through multimedia stories. Working Paper 97-02. Cambridge, MA: MERL – A Mitsubishi Electric Research Laboratories.

Strohecker, Carol (1997). A case study in interactive narrative design. *Proceedings of designing interactive systems* (pp. 377-380). New York: ACM Press.

Strohecker, Carol, Kevin M. Brooks & Larry Friedlander (1999). *Tired of Giving*. In: An experiment in narrative unfolding. Technical Report 99-16. Cambridge, MA: MERL – Mitsubishi Electric Research Laboratories.

Sutherland, John D. (1989). *Fairbairn's journey into the interior*. London: Free Association Books.

Turkle, Sherry & Seymour Papert (1992). Epistemological pluralism: Styles and voices within the computer culture. *Journal of mathematical behavior, 11* (1), 3–33.

Wertsch, James V. (1991). *Voices of the mind: A sociocultural approach to mediated action.* Cambridge, MA: Harvard University Press.

Wilson, Andrew D., & Aaron F. Bobick (1999). Realtime online adaptive gesture recognition. *Proceedings of the international workshop on recognition, analysis and tracking of faces and gestures in real time systems '99.* 270–275.

Winnicott, D. W. (1971a). *Playing and reality*. London: Tavistock Press.

Winnicott, D. W. (1971b). *Therapeutic consultations in child psychiatry*. London: Hogarth Press.

Wren, Christopher R. (1999). *Understanding expressive action*. Doctoral dissertation proposal, Media Laboratory, Massachusetts Institute of Technology.

Chapter 11

Assumptions underlying the Erasmatron storytelling system

Chris Crawford
Jacksonville, Oregon

Introduction

Erasmatron technology consists of three components: the storytelling engine, the front end, and the editor. This technology is predicated on three fundamental assumptions; I intend to demonstrate that, if those assumptions are accepted, then the Erasmatron technology is an easily-understood consequence.

Assumption #1: "Interactive storytelling"

I set out to build a technology to implement interactive storytelling. I did not build digital storytelling (the use of digital media to present conventional stories), nor did I set out to build computer games embellished with storytelling. My goal was to realize the explicit meaning of interactive storytelling: something that is both "storytelling" and "interactive". This pedestrian attitude precludes a great many of the abuses now presenting themselves as interactive storytelling.

Interactivity comes first

How much emphasis should be assigned to each of these two components (interactivity and storytelling)? We will all agree that the final result should give equal emphasis to both, but the more pressing issue is how much effort and energy we should devote to each of the two during development, and the priority we assign to tackling each of them. Clearly, we all understand storytelling more than we understand interactivity. Many designers therefore assume, in deference to the Aristotelian admonition, that the proper approach is to proceed from the more knowable (storytelling) to the less knowable (interactivity). This is a mistake. The admonition to proceed from the more knowable to the less knowable is intended for application to the search for understanding. We best learn about the world by basing new discoveries on a firm foundation of current knowledge. But learning is not at all akin to designing; we have no reason to believe that successful methodologies for intellectual growth can be applied to the utilization of knowledge through design.

Another, less noble reason to begin our efforts with storytelling is simple human laziness: if we start by building what we know well, then we will make rapid progress and can later turn to the difficult phase of the design, confident that we are already halfway to success. That is what we tell ourselves – but we are wrong when we think this way.

In practice, the thoroughness of our understanding of any particular aspect of a design permits us more ways to modify the design of that aspect. For example, a cathedral designer who understands foundations well but flying buttresses poorly can readily design many workable foundations, but few workable flying buttresses. This architect would be tempted to begin his task by designing a clever foundation first, deferring the more difficult task of designing flying buttresses. But such a strategy would surely fail, for requiring flying buttresses to fit an already-designed foundation imposes an additional constraint on an already-difficult task. Our cathedral designer would do better to begin with the most difficult task first: designing flying buttresses that work. Once this problem has been solved, it is a simple matter for a talented foundation-designer to whip up a foundation to fit the flying buttresses. It is easier to deform the well-known task to fit the constraints of the poorly-known task than vice versa.

There are many examples of the failure of this strategy. Most offer the user a conventional story with an iota of interactivity tacked on. Since the result is 99% story and 1% interactivity, it is automatically compared with similar products that are 100% story and 0% interactivity: movies and novels, for example. The purportedly interactive story comes out badly in any such comparison, leading many to conclude that interactive storytelling is a futile enterprise. Like castaways on an island, we will never convince ourselves of the desirability of swimming to a nearby island if we stride knee-deep into the water and observe that our situation has not improved.

Thus, the design of interactive storytelling must begin with the interactivity, applying the knowledge of storytelling to conform to the requirements of interactivity. At some future date, when we understand interactivity almost as well as we understand storytelling, we can dispense with this rule, but for the time being it must reign supreme.

Interactivity mandates choice

If interactivity must be our starting point, we must solidify our understanding of interactivity before we can design with it. Irritatingly, we are obstructed from this goal by the universal misuse of the term. I have in my possession a bottle of "interactive shampoo" – or so the manufacturer claims. I also have an "interactive rug" and an "interactive candy bar". This paper is not the place for me to discourse on a topic as broad as interactivity. Happily, for the purposes of this paper, I need only assert one characteristic of interactivity: it mandates choice for the user. Every interactive application must give its user a reasonable amount of choice. No choice, no interactivity. This is not a rule of thumb; it is an absolute, uncompromising principle.

Those who chafe at uncompromising principles can take comfort in my concession that the amount of choice necessary to achieve interactivity is not carved in stone; it depends on the design situation and as such is subject to some interpretation. I can offer two ways to estimate the amount of choice appropriate for interactive storytelling. The first is to consider the number of choices implicit in a typical story. Since every action the protagonist takes, every word s/he speaks, is the result of a conscious choice, we can safely conclude that hundreds or thousands of choices are made in the progress of a story. Of course, many of those choices, such as including the grammatically required definite articles in spoken sentences, are dramatically insignificant. We can prune our estimates of the number of choices made during a story, but there remain scores or hundreds of dramatically significant choices in the typical story.

A second way of estimating the number of choices required for adequate interactive storytelling is to consider a general criterion for interactivity: the ratio of accessible states to conceivable states. For example, a word processor permits us to realize the great majority of the documents that we can imagine; we are therefore satisfied with the performance of our word processor and feel little need for a better one. If, by way of contrast, our word processor broke and lost its ability to change fonts on command, a great many of the documents we could well imagine would no longer be accessible, and we would consider our word processor to be much less satisfying.

The same criterion applies to interactive storytelling. With any given dramatic situation, our user can imagine a great many possibilities; our task is to provide as many of those possibilities as is reasonable. For example, if Pierre slaps Vincent with his glove, we would surely expect Vincent to be able to challenge Pierre to a duel, or slap Pierre, or verbally or physically assault him. Were we to provide all of these options, then our ratio of accessible states to conceivable states would be fairly good. We could increase that ratio by including some of the less obvious options, such as Vincent skulking away or bursting into tears or falling on his knees and begging forgiveness. A truly heroic effort would further include such unlikely options as Vincent picking his nose or reciting a Tennyson poem – but the "imaginability" of such options is so low that little would be gained from this. In any case, we can agree that a fully developed interactive storytelling product would surely include hundreds or thousands of choices for the user.

Choice means verbs

Thus, our goal is to provide the user with many choices. What, precisely, do we mean by "choice"? In particular, how can we express the vague concept of choice in a computationally accessible data structure? My answer is to reduce the broad concept of choice to the more specific concept of a verb ensconced inside a sentence. Thus, a choice is a selection of one of several plans stated as sentences:

"I will slap Pierre with my glove."
"I will challenge Pierre to a duel."
"I will insult Pierre."
"I will punch Pierre in the nose."

The core concept of each of these choices is the verb (slap, challenge to duel, insult, or punch in the nose), but that verb must be given context with the inclusion of a subject and a direct object to form a sentence.

Discrete versus free-form verb choice

In the ideal storytelling engine, the user will be free to choose any verb imaginable. However, the effort required to dramatically process so many verbs prohibits such an arrangement. For the moment, we require a simpler arrangement that saddles the storyteller with a manageable amount of labor. If a *storyworld* (my term for the data structure fed into the storytelling engine) is to function with, say, only a thousand verbs, then the verbs available to the user at any given point must be confined to the dramatically most significant options. Such a short list of options is most conventionally handled with a menu structure. Fortunately, this does not force us into the classical and unacceptable branching tree structure used in so many previous attempts at interactive storytelling. By taking advantage of the generality afforded by an extended sentence structure, individual verbs can be designed to handle a wide variety of situations and can be re-used many times in a story without imposing tedium.

Consequent structure of the engine; embellishments

The core design of the storytelling engine arises almost spontaneously from these considerations. Starting from an event, expressed as a sentence in the form above, the direct object considers each of the options available and chooses one, at which point it becomes a plan. The plan is then executed and becomes an event, which starts the cycle all over again.

Little percipience is required to compile a long list of flaws in this basic design; the Erasmatron storytelling engine is therefore burdened with an equally long list of modifications, embellishments, and special considerations so that it can address such flaws. For example, the drowsiest reader will immediately note that this design imposes a two-character alternating pattern on the storyline; additional characters can play no role in the story. This killing problem is easily addressed with a concept I refer to as a *role*. As a data structure, a role has a boolean expression asserting the requirements a character must satisfy in order to play the role. It also includes a list of options – verbs – available to the character playing the role. Whenever an event takes place, the engine polls each of the characters who witness the event, starting with the bystanders. Each bystander is tested against the role requirement; if the bystander meets the requirements, then that character selects one of the offered options. In doing so, the character has the option to *hijack* the storyline, precluding any other characters from reacting to the event. Thus, in example presented above, Antoinette could intervene before Vincent reacts to Pierre, hijacking the storyline and precluding Vincent's reaction.

Additional embellishments include provisions for time-deferred execution of plans, and plans deferred until appropriate audience requirements have been met (a man and woman deferring sex until they are alone, for example).

The engine also provides a set of stages on which action can take place; this insures that dramatically necessary subsets of the cast can be assembled at one place. The artist designing the storyworld can specify *territorialities* that specify each character's proclivity to appear at each stage. Characters move among stages under the control of the engine, based on their territorialities and their need to encounter other characters in order to execute plans.

Each character knows (and can therefore react to) only those events that s/he witnesses or is told about. The engine provides a detailed *gossip* system for moving information through the cast. Characters can spike the gossip system with deliberate lies, which can then be transmitted through the cast in the same manner as actual events, and can also be traced back to their source. Characters can also reveal information about events that they were asked to keep secret, and can be tracked down after doing so.

In accordance with standard dramatic practice the Erasmatron engine permits characters to spy successfully upon each other with the flimsiest of camouflage. It supports a large number of props as well as various object-manipulation capabilities such as possession, trade, transportation, and concealment.

Assumption #2: Division of labor through indirection

It would not be difficult to devise a storytelling engine granting detailed artistic control over every aspect of the storytelling process; such an engine already exists in the form of any general-purpose programming language. To be of any utility, a storytelling engine must make assumptions, impose constraints, and perform mechanical tasks, thereby minimizing the amount of detailed specification required of the artist. This creates a dilemma, for artists demand detailed control of the storytelling process but need to have tedious details handled by the software. My solution to this dilemma was to grant the artist varying degrees of indirection of control. Some tasks, such as the specifics of how a character chooses between various options, require detailed specification from the artist; others, such as each character's gossiping behavior, require only the specification of the value of a single variable, *Loquacity*; the engine handles all the mechanics of gossip.

Such indirection is only obtained by means of metaphors readily grasped by the storybuilder; this constraint renders the use of such indirection rather opportunistic. Thus, I was fortunate to be able to use Loquacity as a variable indirectly controlling the propensity of actors to share gossip. Concepts such as roles, stages, and props, have been easy to integrate into the overall design, but many other elements of the storytelling engine are still much too close to their programming roots. In particular I have been unsuccessful in finding metaphors to address the mathematical constructs used to create inclination equations in the role scripts. These scripts still read too much like program code rather than theatrical instructions.

Assumption #3: Ease of editing

The delivered product of interactive storytelling, a storyworld, is unavoidably large; all small storyworlds are boring. Thus, building storyworlds is a big task demanding months or years of labor. This in turn implies that the creation of a storytelling engine alone is insufficient; we must also provide the artist with a powerful editor for building his/her storyworld. The Erasmatron provides such an editor.

Software designers all too often confuse potential with power. A general-purpose programming language such as C++ has vast expressive potential because one can code up anything with it. But the exercise of this potential is constrained by the ability of the user to articulate an expression in a reasonable period of time. There can be no doubt that C++ has the expressive potential to write a program capable of, say, writing profound novels. However, such a program would require billions and billions of lines of code and simply could not be written in any reasonable period of time. Thus, the true power of any software system arises just as much from its overall ease of use as from its intrinsic potential.

The most tedious task in the Erasmatron is the specification of algorithms for how the characters will choose among their options in any given role. Such algorithms must be expressed in a small programming language, akin to a custom macro or scripting language. Yet scripting languages in most products are ghastly monstrosities utterly inaccessible to the non-technical designer. I therefore went to great lengths to build a scripting language that would be easy for a non-technical designer to learn and use. As part of this effort, I developed a number of innovations that might be of general applicability.

Mouse-driven input

The first of these is not at all unique, but my experience demonstrates the value of an under-utilized technique. I replaced keystroke code entry with pure mouse data entry. The scripting language consists of some 500 code tokens, which are accessible through a set of menus. By stacking these menus as pop-up menus along the left edge of the window, I was able to make all code tokens accessible from a single layer of menu; thus, I avoided the messiness of deeply hierarchical menus. The user never types in program code; instead, s/he uses the mouse to select the intended object of alteration and then replaces it with another token from the set of menus. Expressions are expanded by replacing single-argument expressions with double- or triple-argument expressions. *Outsertion* enables the user to build an expression around an existing expression.

To make this system work, all tokens that require arguments provide those arguments, in unspecified form, when they are first inserted into an expression. The unspecified form of an argument is a suggestive label that is underlined to indicate that it remains unspecified and requires the attention of the designer. Thus, the function CountEvents will appear, when first applied, as follows:

– CountEvents(<u>Subject</u>, <u>Verb</u>, <u>DirObject</u>, <u>HowRecent</u>)

No setup

An extension of the concept of providing all arguments in advance is the provision of all requisite lines of script. Whenever a role is created or an option is added to a role, the lines of script that are necessary for that role or option are automatically inserted into the script. Thus, the user is never burdened with the task of declaring initial conditions; there are never any two-step operations in preparing a script.

The upshot of these devices is simple: the user can never make a syntactical error. Any script that can be written is certain to be syntactically correct. This safety feature, which should be universal to programming languages, is too seldom implemented.

Graceful error handling

Poison is another system for gracefully handling errors. Although the scripting system can obviate all syntactical errors, it is not possible to prevent all runtime errors, such as division by zero. When such errors arise, the engine responds gracefully. The error-generating code is flagged as poisoned, and the engine proceeds normally. However, when the time comes to evaluate results, the engine simply ignores results that have been poisoned. Thus, runtime errors do not throw the engine off; they result only in a refusal to consider the poisoned options. The story will proceed normally, hampered only by a lessened breadth of options. In short, runtime errors result in slightly less entertaining stories, but nothing worse.

No flow control

Flow of control considerations are not imposed on the artist. There are no if-statements and no looping structures. There is an implicit if-statement in the boolean expression at the beginning of each script, but there is no generalized branching control. There are several functions that provide built-in looping structures; we have found these adequate to the needs of interactive storytelling. The absence of flow control considerations greatly reduces the burden on the artist.

Assisting agents

The Erasmatron includes a variety of supporting analyses of the storyworld; these are presented through a variety of rather stupid agents that I call *lizards*. It seems to me that the term *wizards* is misleading; since a wizard is an immensely powerful person, capable of performing wondrous feats of magic, I should think that any software agent called a wizard should be endowed with vast powers. The wizards we have seen implemented in software strike me as quite stupid; I wonder what the poor unschooled users think of these supposedly powerful nothingburgers. Therefore, in the interest of clarity of expression (and perhaps some truth in advertising), I have labeled my software agents *lizards*.

Conclusions

Most of the hard thinking that went into the Erasmatron technology involved one or more of these three assumptions (primacy of interactivity, indirection of effort, and ease of editing). Once I had articulated and understood them, the design problems were much clarified. Sadly, I developed these ideas in parallel with the creation of the software; my many wrong turns led me, like a mouse in a maze, to these assumptions as my guiding principles. I am certain that other sets of assumptions can be used to create other interactive storytelling systems – but these three worked for me.

Chapter 12

Story grammars

Return of a theory

R. Raymond Lang
Xavier University of Louisiana, New Orleans

Introduction

This chapter describes a declarative model for simple narratives. The model characterizes event sequences that constitute a story when reported in natural language. Previous work in story generation has followed one of two tracks: (1) declarative, or isolating the structure of stories and then creating text conforming to that structure, and (2) procedural, or modeling and recreating the processes used by human authors. Researchers in the first track often were unable to implement their model; but implementations arising from the second track did not directly address what constitutes a story. By implementing a story grammar, we address both these issues and constructively demonstrate the viability of utilizing formal grammars to describe stories.

Background

Anthropology and linguistics intersect when attention focuses on the folklore pertaining to a culture. In the early nineteenth century, Wilhelm and Jakob Grimm published their collections of traditional domestic tales of the German people (Grimm 1987). Subsequently, Aleksandr Afanasev published his collection of Russian folk tales (Afanasev 1974, 1975), which Vladimir Propp used in the 1920s in his investigations into the morphology of the folktale (Propp 1968). Contemporary investigations into story structure reached a watershed in 1973, when B. N. Colby published a grammar for Eskimo folktales

(Colby 1973). Colby was the first to use formal grammars to describe linguistic phenomenon beyond single sentences.

A variety of questions motivated researchers to employ formal grammars as a way of describing stories. Some were interested in the cognitive mechanisms used by people to summarize and recall stories and proposed that story grammars were an integral part of human language ability. Others were attempting to discern the common structure of stories and turned to formal grammars as a knowledge representation device. Our own interest is closer to the latter; and we do not make any cognitive claims concerning what human beings do when creating or reading a story. We use a formal grammar to describe narratives; to this end, we have developed a set of structural components along with rules for their composition. Our model is (1) general enough to apply to folklore compilations of the sort described above, and (2) sufficiently detailed to rule out constructions of non-stories. Below, we briefly review previous work in story modeling.

Previous work in Declarative Story Modeling

Rumelhart develops a model for the organization that takes place in connected discourse but is absent in strings of sentences (Rumelhart 1975). In general, it is almost always necessary to infer (unstated) causal relationships to understand groups of sentences. These causal relationships relate sentences to each other. Rumelhart presents a grammar describing the inter-sentence bindings that arise in simple stories. The grammar is context-free and consists of syntactic rewrite rules each of which has a corresponding semantic interpretation rule. The primitives are meta-sentence components such as setting, episode, and event. Below are two rules from David Rumelhart's story grammar.

1. Attempt → Plan + Application
   ⇒ *MOTIVATE*(*Plan*, *Application*)
2. Application → (Preaction)* + Action + Consequence
   ⇒ *ALLOW*(*AND*(*Preaction*, *Preaction*, ...),
   {CAUSE | INITIATE | ALLOW} (Action, Consequence))

In Rumelhart's grammar (and those derived from his grammar), the relationship a component has to other components is expressed in "semantic" annotations accompanying the "syntactic" rules. Scare quotes distinguish the story grammar use of the terms "syntactic" and "semantic" from conventional use. In story grammars, the terms are intended to mean something like "structural" and "extra-structural," but in fact mean rather "captured by the grammar" and "not captured by the grammar." If the "syntactic" structure of a portion of a text makes a particular rule applicable, then the relationship of this component to others is gleaned from the annotation to the rule. Unfortunately, the "syntax" given in story grammars doesn't rule out many constructions; while the "semantic" annotations are inadequately defined. This deficiency leaves the grammars open to wishful parsing and generation, a serious flaw which proponents of story grammars were unable to overcome. A major part of this work is a rigorous, formal framework used in relating story components to one another (Goldman & Lang 1993; Lang 1997).

Following Rumelhart's "Notes on a Schema for Stories," other researchers expanded on Rumelhart's grammar (Bower 1976; Frisch & Perlis 1981; Johnson & Mandler 1980; Mandler & Johnson 1977; Stein & Glenn 1979) while others attacked the foundations of the possibility of a "grammar for stories" (Black & Bower 1980; Black & Wilensky 1979; Garnham 1983; Wilensky 1982). Eventually, the story grammars project was abandoned as unsuccessful, largely due to the crude state of formal techniques available at the time, but also due to the excessive demands made of story grammars as a cognitive mechanism.

Story Generation by Author Modeling

Around the same time as Rumelhart's seminal paper on schemas for stories, Meehan published his dissertation on story generation (Meehan 1976). His system, Tale-Spin, inspired work in story generation from the perspective of author modeling, that is, by modeling the cognitive processes of a human author of stories. Turner's Minstrel (Turner & Dyer 1985; Turner 1990, 1991a, b) and the system described by Okada and Endo (Okada & Endo 1992) are representative samples of author-modeling systems for stories.

Meehan's Tale-Spin is a simulation of a forest world, producing natural language output describing the interactions of characters pursuing goals such as eating and drinking in a context where duplicity and hostility occur along with honesty and friendliness. Although Tale-Spin provides access to the meanings (conceptual dependency forms, in this case) from which the natural language text is constructed, the model by which the meanings themselves are generated is left implicit; and the relationships among the components of a story are deeply entwined in the procedures which drive the simulation.

Michael Lebowitz develops a model of story telling based upon an extensible library of plot fragments (Lebowitz 1985). These plot fragments serve the goals of the author, which may be nonsensical from the point of view of the characters. For example, an author may have a goal to keep lovers apart; and, in pursuit of this goal, he will insert into a story elements that prevent lovers from meeting. It would be absurd for lovers themselves to seek obstacles to their meeting; but as a device for enhancing a story's dramatic interest, it makes perfect sense for the author to devise such obstacles. Lebowitz's Universe program generates plot outlines using an algorithm very similar to that used in Tale-Spin except that author goals rather than character goals drive the mechanism. The research issue addressed by Lebowitz treats the realization of an author's goals in a story.

Scott Turner and Michael Dyer describe Minstrel (1985), a story-telling program which generates believable and logically consistent stories that make a point. Turner describes further development of Minstrel in subsequent papers (Turner 1990, 1991a, b). Turner's primary interest is in modeling human creativity and human story-telling behavior, and he uses King Arthur-style tales as his domain. Although we are working in a domain bearing superficial resemblances to Turner's, our objective is a model that is independent of the process human authors undertake when writing a story.

A new grammar for stories

This section describes selected features from our formal model for simple narratives (Lang 1997). The model takes the form of a definite clause grammar, hereafter referred to as "the-grammar". The nonterminals are meta-components such as setting, episode, outcome, etc. The terminals are first-order predicate calculus schemas for the events, states, goals, and beliefs which, when instantiated and rendered into natural language, are the content of a simple narrative. The language described by the-grammar consists of lists of FOPC expressions. Each list is an ordered representation of the facts and events contained in some tale; but the list does not specify the relations among the various terms in it. The example below shows an event list representing a portion of "The Bad Wife." The list adequately captures the states and events in the story; but it does not represent the relationships among them. For example, nothing in the list indicates that the trick carried out by the peasant at time int(x7, x8) serves the goal held during time int(x10, x8) that his wife be in the pit. The information about the relationships among the elements of the event list is specified in the rules of the-grammar.

```
[holds(lives(peasant), int(x1, x2)),
holds(married_to(peasant, wife), int(x1, x2)),
holds(disobeys(wife, peasant), int(x1, x2)),
occurs(quarrel(peasant, wife), int(x2, x12)),
occurs(do(peasant, walk(in(woods))), int(x4, x5)),
occurs(finds(peasant, pit, under(bush)), int(x5, x6)),
goal(peasant, holds(loc(wife, in(pit)), int(x8, x9)), int(x10, x8)),
occurs(do(peasant, trick(wife)), int(x7, x8)),
holds(loc(wife, in(pit)), int(x8, x9)),
holds(alone(peasant), int(x8, x9)),
occurs(time_passes, int(x9, x20)),
...]
```

The story rule

We model a story as having two sub-components, a setting and an episode list, each having temporal intervals such that the setting interval must meet that of the episode list. *Starting rule for stories.*

```
story(story(Setting, Ep_list)) –>
    setting(Setting, S_time),
    episodes(Ep_list, E_time),
    {meets(S_time, E_time)}.
```

The left hand side of the rule states that a story is a labeled pair story(Setting, Ep_list). The right hand side states (1) that Setting and the temporal interval S_time must satisfy the rule for a setting; (2) Ep_list and the temporal interval E_time must satisfy the rule for episodes; and (3) the temporal intervals S_time and E_time must satisfy the constraint meets.

Rules for episodes

The episodes rule shown below defines this component as a non-empty list of components of the form ep(Ev,ER,A,O). The subcomponents of the episode are as follows:

Ev, an initiating event
ER, an emotional response on the part of the protagonist
A, an action response on the part of the protagonist
O, an outcome or state description which holds at the conclusion of the episode.

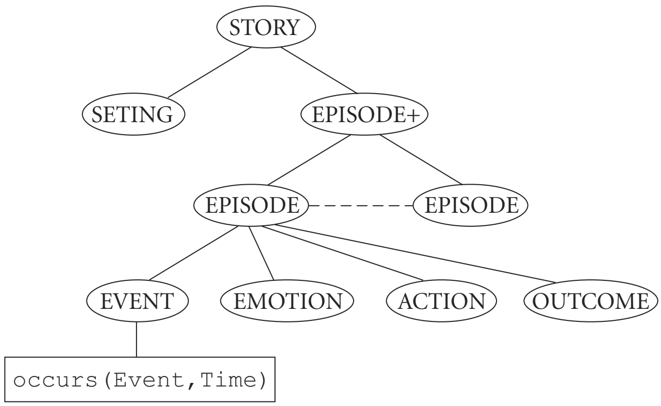

Figure 1. A partial parse tree for a simple narrative showing a leaf template.

The second argument of the episodes rule is the temporal interval over which all the episodes in the list take place. The temporal constraints in the episodes rule capture the relations among the intervals associated with the first episode, the remaining episodes, and the overall episodes list. The episode rule describes the relations among an episode's four components and the associated temporal intervals.

The left hand sides of most rules specify other components which the right hand side relates to the non-terminal defined by the rule. This allows the grammar to capture constraints within the rule itself rather than as a "semantic annotation." By adding arguments to the non-terminals in the rules, the grammar achieves the expressive power of a Type 0 grammar, as Black and Wilensky conclude a story grammar must (Black and Wilensky 1979). Below are two rewrite rules from the grammar for stories.

```
episodes(episodes([ep(Ev, ER, A, O)|Es]), int(Start, End)) –>
    episode(ep(Ev, ER, A, O), P, int(Start, Mid)),
    episode_rec(Es, P, int(Mid, End)).

episode(ep(ev(Ev, Ev_time), emot(Em, Resp_time)), A, O), Ep_time) –>
    story_event(Ev, Ev_time),
    wm_call([emot_reaction, Ev, Em]),
    emot_response(Ev, Em, Ev_time, Resp_time),
    wm_call([act_motiv, Em, A]),
    action_response(A, O, Act_time, Outcm_time),
    {starts(Ev_time, Ep_time),
    finishes(Outcm_time, Ep_time),
```

meets(Ev_time, Resp_time),
starts_before(Resp_time, Act_time)}.

Figure 1 shows a partial parse tree. The ovals depict non-terminals, and the rectangle depicts a terminal expression. The grammar specifies the structure of the tree and the kind of event list expression needed at the leaf nodes. The expressions at the leaves are instantiated by a separately defined world model.

The implementation

The ambiguity and poor specification of previous narrative models made them difficult or impossible to implement. This weakened the claim that these models of narratives were, in fact, computational in nature. We present a concrete implementation in support of our claim that our model, based on our theory of rational intention in autonomous agents (Goldman & Lang 1993; Lang 1997), is indeed a computational model describing a non-trivial class of narratives. The implementation, named Joseph (**J**ust **O**rdinary **S**tories **E**numerated by **P**rolog **H**ierarchically), produces randomly generated natural language narratives conforming to the-grammar.

Components of Joseph

The tasks of the Joseph story generation system are divided among the following components:

Story grammar: At the core of Joseph is the implementation of the-grammar. The story grammar defines structured series of story components.

Grammar interpreter: The grammar interpreter defines the search strategy of the generation process. We use depth-first, iteratively deepening search (Korf 1987) plus random choice to find a sequence of grammar rule rewrites which defines a valid story.

Temporal predicates: The events in a story must satisfy temporal relations specified in the-grammar. We implement Allen's temporal relations (Allen 1984) to enforce temporal constraints on story components.

World model: A story must have content as well as form. The story grammar produces abstract representations of stories; the grammar specifies terminals as schemas but does not specify the bindings of variables contained in those schemas. The potential instantiations of terminals are drawn from a set of actions that characters may perform and fluents that vary during the course of the tale. This set of actions, fluents, and characters is enumerated in a world model.

Natural language output unit: The story grammar and the world model define event lists: sequences of events which when rendered into natural language constitute a story. These event lists are accompanied by the parse tree describing the structure of the story. The event list and the parse tree are rendered into natural language text to produce the final output.

Figure 2 illustrates the interaction among these components. The grammar interpreter initiates the generation process by invoking the top level rule for a story. When the generation process reaches a terminal, the grammar rule specifies a *leaf schema* and requests that the world model instantiate it. A leaf schema determines the form of the terminal and specifies how the world model elements fit into the parse tree. When the grammar interpreter has produced a completely instantiated parse tree and event list, these two structures are mapped to surface text.

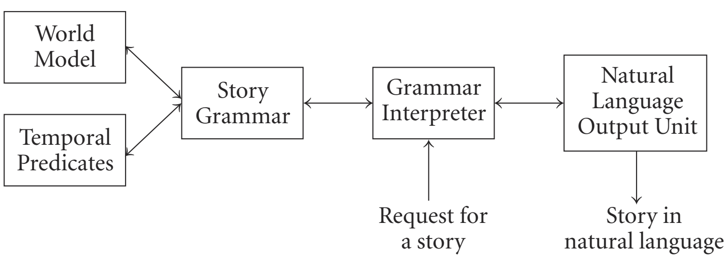

Figure 2. Schematic diagram of the relationship among the five Joseph components.

The world model

Coherence relations such as causality and goal-directedness are the "glue" connecting the story's states and events. By the *form* of a story, we mean these coherence relations among the constituent components. By a story's *content* we mean the component states and events themselves. Our model separates rules governing the form of a simple narrative from elements making up its content. We implement this distinction by packaging the grammar terminals into a separate world model. The model Joseph uses to generate stories supports creation of stories resembling Russian folk tales. The predicates in the model represent the relationships among the characters, events, actions, effects, emotions, and goals appearing in arbitrarily selected Russian folk tales.

During generation, the world model instantiates the grammar terminals. For example, two terminal components of an episode are an event and the protagonist's reaction to that event. However, the grammar specifies neither the event nor the reaction. To instantiate these components, the grammar invokes the world model, which specifies possible events and character reactions. The list of instantiated terminals (i.e. states and events), up to the point of the world model invocation, is available to the world model predicates so that these predicates can check preconditions of relations and bind variables to domain elements mentioned in the story.

A story grammar is independent of the world model. In other words, the story grammar does not require any particular means of providing an event, state, reaction, etc. The Joseph system's world model is implemented as sets of associations between events and their effects, emotions and the actions that arise out of them, goals and the actions which serve them, states and consequences of those states, etc. We refer to these associations collectively as the coherence predicates. Each association is expressed in general terms as the head of a prolog clause, the body of which describes the conditions that make the association compatible with the events and states of a partially generated story. When a world model predicate is encountered during generation, the grammar interpreter collects all such compatible associations and chooses one at random to instantiate the grammar terminal expression.

The coherence predicates do not specify individual domain elements such as a character, place, object, etc. These atomic domain elements are abstracted out of the coherence predicates and specified separately along with their attributes. This allows the coherence predicates to be expressed in general terms. For example, one coherence predicate associates a character's seeing a desirable thing with the character's adopting the goal of having that thing, without saying what the thing might be. Separating the coherence predicates from the domain elements and their attributes enables easier modification and maintenance of the world model and simplifies addition of both new coherence associations and new atomic domain elements.

Output samples

This section presents selected stories generated by this implementation. The simplest possible story in our model consists of a setting followed by a single episode composed of an event to which the protagonist reacts without forming any goals.

> once upon a time there lived a dog. one day it happened that farmer evicted cat. when this happened, dog felt pity for the cat. in response, dog sneaked food to the cat. farmer punished dog.

A slightly more complicated single-episode story has the protagonist adopt a goal and carry out action(s) in pursuit of that goal. Stories with goals are more complex because (1) the system must constrain the protagonist's actions to those intended to achieve the goal, and (2) the system must track the effects of these actions to determine if the goal is met.

> once upon a time there lived a cossack. one day it happened that imp possessed daughter of a boyar. when this happened, cossack felt love for the daughter of a boyar. in response, cossack made it his goal that he would be married to the daughter of a boyar. cossack exorcised the imp from the daughter of a boyar. cossack was married to daughter of a boyar.

Our implementation also produces multiple-episode stories. Episodes may be arranged in two ways: sequentially or nested. Episode nesting takes place when the world model instantiates an action's effect to an event rather than a state. In the tale below, the nested episode is emphasized. The action of the outer episode (taking a walk in the woods) does not have a state as its effect. Instead, this action triggers an event (finding a pit) which initiates a nested episode.

> once upon a time there lived a peasant. peasant was married to wife. one day it happened that peasant quarreled with the wife. when this happened, peasant felt distress. in response, peasant took a walk in the woods. *peasant found a pit when he looked under the bush. when this happened, peasant desired to punish wife. in response, peasant made it his goal that wife would be in the pit. peasant tricked wife. wife was in the pit.* peasant lived alone.

The final example illustrates goal failure. When a character adopts a goal, it must attempt to achieve it. The world model enumerates (1) the actions that a character may take toward a given goal and (2) the effects these actions have. The grammar tries to match the effects of the actions with the goal. When it succeeds, the goal has been met. If the world model does not provide an action having an effect that entails the goal or unifies with it, then the goal fails. Our theory of rational intention (Goldman & Lang 1993; Lang 1997) specifies conditions for an agent to give up a goal. The Joseph system incorporates these conditions implicitly.

> once upon a time there lived a peasant. peasant was hungry. one day it happened that the peasant met christ. when this happened, peasant felt awe. in response, peasant begged christ to provide food. christ told peasant to eat ram. when this happened, peasant felt obedient. in response, peasant made it his goal that ram would be eaten. peasant trapped ram. ram whacked peasant. peasant believed it impossible that ram would be eaten. peasant was hungry.

Results and future work

The model described in this chapter defines sequences of events which, when reported in natural language, constitute simple narratives. The model is implemented in SWI-Prolog. The implementation, called Joseph, generates event sequences in conformance with the-grammar. Although surface text realization is not the focus of this work, Joseph does include a simple surface text unit that renders event sequences into natural language. Joseph constructs event sequences from a pool of events reverse engineered from arbitrarily selected Russian folk tales. The characters, events, responses, and interactions contained in these tales are combined to generate new stories. We gauge the quality of the model according to the coherence and continuity of the resulting narratives.

The results are encouraging. Tales produced by the Joseph system vary widely in both content and structure. This is evidence that the model captures a broad range of the kind of narratives we are interested in. This chapter includes a few short output samples; however, typical Joseph-produced tales are several paragraphs long and contain multi-episode plot developments. For example, a generated tale could begin with the protagonist's spouse going on a trip in one episode, getting captured by a rival ruler in another episode, and being rescued by the protagonist in a third. Very often, these multi-episode developments are intermingled with other, unrelated episodes; this mimics folktales in which a character has chance encounters while on an overarching quest or mission.

The results also point to future directions for this work. Stories produced by Joseph exhibit varying degrees of coherence. Some stories exhibit strong coherence; others contain plot developments that require the reader to infer intermediate events not reported in the narrative. Expecting a reader to infer unreported events does not prevent a text from being a story. To the contrary, stories commonly omit some events. The issue is characterizing the kinds of events the reader can reasonably be expected to infer. Classifying "omittable" events is a function of a reader model as outlined by Bailey (Bailey 1999) or of a narrator model as described by Szilas (1999). We are also approaching the issue of coherence by incorporating explanation closure (Shubert 1990) into the search strategy (Stewart & Lang 1998; Gardere 2000).

The model, as presently implemented, captures various kinds of non-stories as well as the ethnic narratives under consideration. Some stories juxtapose characters in a way that conforms to common-sense notions associated with those characters (maidens are innocent, wolves are fierce and cruel, etc.); other stories place characters in counterintuitive roles. This is evidence that the model suffers from an inadequate representation of characters. We are presently exploring agent design models (Bailey 1999; Meech 1999; Frank & Stern 1998; Sengers 1998).

Our approach depends heavily on the world model and suffers from the problems that many knowledge-based systems share. The present implementation enumerates, prior to generation, the actions, goals, effects, etc. that may appear in a story. The present implementation is ontologically promiscuous in this regard. Furthermore, although the world model predicates have access to all states and events *reported* in a story up to the point the world model predicate is triggered, the implementation does not consider *unreported* states or events entailed by those that are reported.

Despite these shortcomings, the Joseph story generation system represents a significant achievement since it is the first such system constructed from an explicit, formal model for stories. The Joseph system constructively demonstrates the value of the insights and intuitions put forth by the proponents of story grammars in the 1970s and early 80s. Lakoff, Colby, Rumelhart, Mandler and Johnson, Stein and Glenn, Bower, Thorndyke, and Frisch and Perlis were correct to assert that formal grammars can be utilized to describe the features of narratives. Although Joseph's domain is ethnic folktales, the approach applies to stories in general. For example, initial efforts have been made to use this approach to generate algebra word problems (Matthews 1999; Thompson 2000). The viability of story grammars is good news since formal grammars are well-analyzed and well-understood tools for describing classes of structured objects such as narratives.

References

Afanasev, Aleksandr. N. (1974). *Russian folk tales*. New York: E.P. Dutton.

Afanasev, Aleksandr. N. (1975). *Russian fairy tales*. New York: Pantheon Books.

Allen, James F. (1984). Towards a general theory of action and time. *Artificial Intelligence*, *23* (2), 123–154.

Bailey, Paul (1999). Searching for storiness: story-generation from a reader's perspective. *AAAI Fall symposium on Narrative Intelligence* (pp. 157–163). Falmouth, Massachusetts: AAAI Press.

Black, John B. & Gordon H. Bower (1980). Story understanding as problem-solving. *Poetics, 9*, 223–250.

Black, John B. & Robert Wilensky (1979). An evaluation of story grammars. *Cognitive Science, 3*, 213–230.

Bower, Gordon H. (1976). Experiments on story understanding and recall. *Quarterly Journal of Experimental Psychology*, *28,* 511–534.

Colby, Benjamin N. (1973). A partial grammar of eskimo folktales. *American Anthropologist*, *75*, 645–662.

Frank, Adam & Andrew Stern (1998). Multiple character interaction between believable characters. *Proceedings of the 1998 computer game developers conference* (pp. 215–224). San Francisco, California.

Frisch, Alan M. & Donald Perlis (1981). A re-evaluation of story grammars. *Cognitive Science, 5*, 79–86.

Gardere, L. (2000). An application of backward reasoning to Narrative Intelligence. *Proceedings of the eleventh annual CCSC–south central conference* (pp. 50–57). Corpus Christi, Texas.

Garnham, Alan (1983). What's wrong with story grammars? *Cognition*, *15*, 145–154.

Grimm, Jacob L.C. & Wilhelm C. Grimm (1987). *The complete fairy tales of the brothers Grimm*. New York: Bantam.

Goldman, Robert & Raymond Lang (1993). Temporal intentions. *Second symposium on logical formalizations of commonsense reasoning*. Austin, Texas.

Johnson, Nancy S. & Jean M. Mandler (1980). A tale of two structures: underlying and surface forms in stories. *Poetics, 9*, 51–86.

Korf, Rich (1987). Search. In S. Shapiro (Ed.), *Encyclopedia of Artificial Intelligence*. New York: Wiley-Interscience.

Lang, R. Raymond (1997). *A formal model for simple narratives*. Doctoral dissertation, Tulane University.

Lebowitz, Michael (1985). Story telling and generalization. *Proceedings of the Seventh annual conference of the Cognitive Science Society* (pp. 100–109). Berkeley, California.

Mandler, Jean M. & Nancy S. Johnson (1977). Remembrance of things parsed: story structure and recall. *Cognitive Psychology, 9*, 111–151.

Matthews, Clyde (1999). Using formal grammars to encode expert problem-solving knowledge. *Proceedings of the southeast regional conference of the Association for Computing Machinery* (pp. 258–263). Mobile, Alabama.

Meech, Jon (1999). Narrative theories as contextual constraints for agent interaction. *AAAI Fall symposium on Narrative Intelligence* (pp. 38–43). Falmouth, MA: AAAI Press.

Meehan, James R. (1976), *The metanovel: Writing stories by computer*. Doctoral dissertation, Yale University.

Okada, Naoyuko & Tsutomu Endo (1992). Story generation based on dynamics of the mind. *Computational Intelligence, 8*, 123–160.

Propp, Vladimir (1968). *Morphology of the folktale*. Austin: University of Texas Press.

Rumelhart, Dave E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211–236). New York: Academic Press.

Sengers, Phoebe (1998). *Anti-boxology: agent design in cultural context*. Doctoral dissertation, Carnegie Mellon University.

Shubert, Len (1990). Monotonic solution of the frame problem in the situation calculus. In H. Kyburg Jr., R. P. Loui, & G. N. Carlson (Eds.), *Knowledge representation and defeasible reasoning* (pp. 23–65). Boston: Kluwer Academic Publishers.

Stein, Nancy L. & Christine G. Glenn (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing, 2* (pp. 53–120). Norwood, NJ: Ablex.

Stewart, Damon & Raymond Lang (1998). Explanation closure as a knowledge representation tool for automated story generation. *Proceedings of the southeast regional conference of the association for computing machinery* (pp. 332–337). Atlanta, Georgia.

Szilas, Nicolas (1999). Interactive drama on computer: beyond linear narrative. *AAAI Fall symposium on Narrative Intelligence* (pp. 150–156). Falmouth, Massachusetts: AAAI Press.

Thompson, T. (2000). Ontology development with CASPOR. *Proceedings of the eleventh annual CCSC–south central conference* (pp. 58–63). Corpus Christi, Texas.

Thorndyke, Perry W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology, 9*, 77–110.

Turner, Scott R. & Michael G. Dyer (1985). Thematic knowledge, episodic memory and analogy in MINSTREL, a story invention system. *Proceedings of the seventh annual conference of the Cognitive Science Society* (pp. 371–375). Berkeley, California.

Turner, Scott R. (1991a). A case-based model of creativity. *Proceedings of the eleventh annual conference of the Cognitive Science Society* (pp. 933–937). Chicago, Illinois.

Turner, Scott R. (1991b). A case-based model of creativity. *Workshop on creativity, AAAI-91*. Anaheim, California.

Turner, Scott R. (1990). MINSTREL: A model of computer storytelling. *Workshop on interactive fiction and synthetic realities, AAAI-90*. Boston, Massachusetts.

Wilensky, Robert (1982). Points: A theory of the structure of stories in memory. In W. G. Lehnert & M. H. Ringle (Eds.), *Strategies for Natural Language Processing* (pp. 345–374). Hillsdale, NJ: Lawrence Erlbaum.

Part III

Agents and Narrative

Chapter 13

Virtual Babyz

Believable agents with narrative intelligence

Andrew Stern
PF. Magic / Mindscape, San Francisco, CA

Introduction

The most popular contemporary mediums for experiencing stories are books, movies, television and theater. Yet as the number of people who own personal computers continues to grow, the potential for the computer to become a new medium for stories also grows. And because the computer is an interactive medium, artists must begin thinking about how to design and implement interactive stories, in which the user's participation develops and shapes the narrative structure itself.

While there already exists a thriving industry producing interactive entertainment, namely the videogame and computer game industry, there has been little success with creating powerful interactive narrative experiences in these games. Virtually all of today's computer games focus on some sort of action-oriented, strategy-oriented or puzzle-oriented interactivity as the core of the experience. Some incorporate a story-line to accompany the game, but these stories are mostly linear and unchangeable, often serving as a justification for solving yet another puzzle or fighting another opponent. Players have little or no control over the course of the narrative, and AI plays little or no role in developing the narrative (Stern 1999; Mateas 1999; Stern 1998). Games often have characters in them, such as in adventure or role-playing games, but with few exceptions they are not "believable" (Bates 1992), behaving one-dimensionally and predictably, with little potential for more than shallow interactivity. Perhaps most fundamentally, the intention of today's computer game is to play a game, with story holding a secondary, supportive role at best.

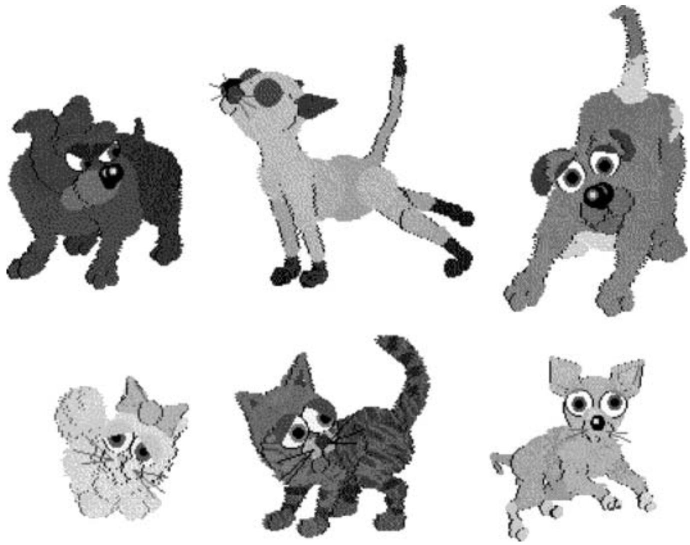

Figure 1. Virtual Petz.

Recognizing a dearth of meaningful interactive experiences with virtual characters, we began the Virtual Petz (Figure 1) products with the design goal to create the richest interactive "illusion of life" we could on a personal computer, within the framework of a non-goal-oriented play environment. Users "adopt" their virtual Dogz and Catz as puppies and kittens, and play with, raise and nurture them in the same manner that one would with real pets, with petting, toys, food, going places, behavior training, and so on. The Petz characters are directly interactive, with the appearance of rich personality and emotion, and the ability to express themselves in a performance-like way through action and behavior. To implement these socially intelligent agents we developed a behavior-based architecture with a model of personality and emotion, all tightly integrated with an expressive realtime-3D-rendered animation system and seamless user interface (Frank, Stern & Resner 1997).

In early versions of the Petz products, users interacted with one virtual character at a time. In later versions, as we put multiple characters on-screen together and allowed their behaviors to play off of one another (with the user as an ever-present interactive participant), we found to our surprise that the interplay between the variety of personalities gave rise to many narrative situations. Small "stories" seemed to emerge as these complex synthetic characters acted out their innate personalities. Without explicitly building narrative into the system, recognizable short-term narratives were occurring (Stern, Frank &

Resner 1998). One of our favorite examples of emergent narrative in Petz is as follows:

A tired old bulldog won't be able to sleep when a young playful kitten is bouncing around him. Because the dog has a nurturing relationship towards the kitten, he will tend to bring the kitten some food. Left alone for the moment, the bulldog begins to lie down for a nap. However the kitten, by her own nature and personality, has a short-attention span and quickly returns to jumping around the bulldog. The bulldog expresses his frustration at being tired, but because the kitten is gesturing for attention, he responds by bringing the kitten a toy. This cycle continues until the kitten gets tired and takes a nap, allowing the now disgruntled bulldog to finally get some rest.

Note that the bulldog was never explicitly programmed to "distract" the kitten so he could sleep - but if you look at these "naturally" unfolding events as a whole, it comes together as a vignette about a dog trying to distract a kitten so he can sleep. To be sure, the user's subjective experience is enhanced by the "Eliza effect" - the tendency for people to treat programs that respond to them as if they had more intelligence than they really do (Weizenbaum 1966). But it was clear to us that creating a broad base of richly interactive behaviors for virtual characters laid fertile ground for interactive narrative - much to the testament of our users, who posted hundreds of messages on our website bulletin board describing their experiences and relationships with their individual Petz (PF. Magic 1998; Stern 2000).

The success of the Virtual Petz (over 2 million copies sold worldwide between 1995 and 1998), as well as the success of other virtual character products such as Tamagotchi (Bandai 1996) (10+ million sold), Furby (Tiger Electronics 1998), and Creatures (Grand et al. 1997) is an indication that people are interested in more than the traditional computer game genres. Our next step was to create even more intelligent virtual characters, and purposefully endow them with some narrative intelligence to increase the potential for more explicit interactive narratives.

Setting the stage for interactive narrative

With interactive narrative as a key design goal for our next product, we set out to create new characters and environments using the following criteria:

– that we continue to use our tried-and-true direct interaction interface, where the user controls a hand-shaped cursor to directly touch and pick

Figure 2. Virtual Babyz

up characters and objects, and that the characters have this same direct interaction with objects and each other;
- that these new characters be made more intelligent (e.g., better able to manipulate and use objects), more expressive (e.g., facial expressions and simple language) and more communicative (e.g., able to understand simple spoken words via voice recognition), thereby making them more capable to perform in narratives;
- that the virtual environments these characters live in have many opportunities for story-like situations to occur, stocked with objects and props designed for playful mischief and humor;
- that the characters be familiar and recognizable enough that we can leverage off of the user's own knowledge and expectations for dramatic effect;
- and that we choose characters that we can successfully implement at the current state of animation and artificial intelligence technology, so as to stay believable.

Human cartoon baby characters fit all of these criteria quite nicely. Within our simplified cartoon-like visual style, Babyz can display a wide range of emotional facial expressions such as happiness, giggling, laughing, frowning, crying, throwing a tantrum, anger, curiosity, tiredness, and so on (Figure 2).

They can crawl around their virtual house and pick up objects, throw them, use them, eat them, carry them to different places. Through voice recognition Babyz can respond to simple words spoken by the user in the form of praise, discipline, and the names of objects. Babyz will "learn" to speak back these words in the form of "baby talk", allowing them to truly say what they (seem to) want and feel.

The Babyz live in a virtual house with all the traditional baby accoutrements such as cribs, highchairs, and changing tables. Many of the toys allow for mischief and fun, such as mushy food that can be thrown and splatted, rubber balls that tend to bounce around the room and knock things over, cookie jars placed on challengingly high countertops, goofy clothing and outfits, and so on.

The Babyz personalities are based on well-established cartoon archetypes, such as "the clever naughty kid", "the spoiled brat", and "the sweet little angel". By making the Babyz sound and act in these recognizable ways, it made it much easier for us to design understandable and entertaining narratives. (Interestingly, in the history of traditional animated cartoons there have been few baby characters, perhaps because cartoons are often a bit violent, for which babies are not as well suited as cats and mice or coyotes and roadrunners.)

Short-term narratives in Babyz: Poops and pranks

A narrative in Babyz is a sequence of actions and behaviors that follow some sort of recognizable continuity. There are a variety of short-term narratives that can occur, lasting anywhere from twenty seconds to several minutes each. Two examples are described later in this section: soiling a diaper (leading to a diaper change), and playing a mischievous prank on a fellow Baby.

The Babyz behavior architecture is designed to allow only one short-term narrative to be occurring at one time. However, if needed, a short-term narrative can be briefly diverted for short amounts of time to allow for unpredictable interruptions, such as user interaction (e.g., being briefly tickled, picked up, spoken to, or offered a toy), or in response to internal metabolism (e.g., the character may be hungry and decide to quickly eat some food that happens to be nearby). After the interruption is over, the narrative will attempt to resume where it left off. But if the narrative is distracted for more than ten or twenty seconds, it will probably abort and allow for a different narrative to begin.

Short-term narratives are implemented as high level behavior goals, each goal having multiple possible plans that can be executed in a non-linear order.

These goals are often spawned as *reactions* to user interaction or other events in the environment, or to the character's own internal metabolism. Note that many of the goals are non-narrative, such as being tickled or being picked up and carried; in a reactive situation, narrative and non-narrative goals compete for execution. At any decision point, each goal's filter function is queried to compute how important it is for that goal to execute under the current circumstances. Filter functions are custom code in which the programmer can specify under what conditions a goal should execute. In a reactive context, a filter function for a narrative goal is customized to respond strongly in situations in which objects its behavior requires are now available. Part of the craft of authoring behaviors is balancing the output of these filter functions; it is easy to accidentally code a behavior to happen far too often or too seldom for believability.

Narrative goals can also be spawned *deliberately* as a need to regularly express the character's particular personality and maintain an illusion of free will. This is achieved by regularly querying all goals' filter functions in a non-reactive context, using the character's personality attributes as a basis for evaluating its goal's importance; only the filter functions for narrative goals are coded to respond in this context. Additionally, a simple "story manager" is keeping track of how often certain interesting narrative goals have occurred over time, and will deliberately spawn such a goal if the user hasn't experienced one in a while.

All instantiated goals are sorted in order of urgency priority, with only one goal executing at any one time (the highest priority goal); all others are suspended (that is, waiting to start or to resume execution). At any time the current active goal can be interrupted and suspended if a higher priority goal is spawned. In fact goals can spawn other goals, allowing for multiple goals to queue up for eventual execution. Goals may delete themselves if their context conditions become invalid.

As a first example of a short-term narrative (and perhaps the quintessential baby behavior), the metabolism model is keeping track of how much time has passed since a Baby last ate, and may spawn a high-priority "soil diaper" goal. This goal always begins with the same plan, "poop in diaper", in which the Baby sits still wherever it may happen to be on-screen at that time, making sounds and facial expressions to the effect of dirtying its diaper. After this plan finishes, the goal lowers its own priority, and suspends itself. This allows other medium-priority independently spawned goals to execute, such as "play with toy", "crawl and explore", "build blocks" and so on. (Which new goals get chosen is influenced by, for example, what toys the user may be interacting with at the time.) Note that the Babyz's behavior during these new goals will be *customized* to reflect the fact that the Baby has a dirty diaper. For example, appropriate alternate locomotion animations, facial expressions, vocal sounds and diaper sounds will be chosen for use during the behavior. Eventually the "soil diaper" goal's filter function will force itself to resurface as the highest-priority goal and begin choosing from more crisis-oriented plans such as "act cranky", "cry", "itch bottom", or "babble" with baby-talk words like "poo poo" or "doo doo", each expressing more severe facial expressions and sounds during the behavior. The goal will continue choosing from this set of plans indefinitely until the user puts the Baby on the changing table, causing the goal to begin choosing from plans such as "happy diaper change", "resist diaper change", "act cranky" or "crawl and explore". Once the diaper is actually changed, the goal finishes and deletes itself, thereby ending the narrative.

Note that at any time during this narrative the user could interrupt and cause new higher priority goals to be spawned, such as "being tickled", "being picked up and carried", "react to toy shaken in my face", and so on. However these non-narrative goals are designed to end as quickly as possible if any narrative goals are waiting. Some user interactions will not disrupt an ongoing narrative at all, but instead influence its execution, such as saying "shhh" if the Baby is crying, or the influencing the choice of which toy to play with, as described earlier.

Another short-term narrative example is the goal "play a prank", perhaps deliberately spawned by a naughty Baby's need to regularly express its personality, in conjunction with the presence of another rival Baby nearby. Before choosing a plan, the goal first scans the environment for objects tagged as good prank toys, such as a glob of wet food or a bouncy ball. (The user can influence this choice by putting certain toys within reach.) The goal first chooses the "pickup toy" plan with the chosen toy. Depending on what toy it picked up, the Baby may choose plans such as "throw toy at other Baby" or "shoot toy at other Baby". Depending on the outcome of the plan, as well as the user's reaction (such as verbal praise or discipline), the goal may finish with a plan such as "point and laugh", "act angry", "act shameful". Note that if this "play a prank" goal is interrupted by the user, for example by tickling the Baby, this goal is designed to delete itself, since the nature of its narrative requires tight continuity, and has no suspension-of-disbelief consequences for aborting. In the case of the earlier "soil diaper" goal, it will never delete itself until the diaper actually gets changed.

Other examples of short-term narratives include “playing musical instruments” to improvise a song, and “reading of a picture book” (a sort of narrative-within-narrative).

Long-term narratives in Babyz: Rivalries, relationships and development

The Babyz characters (as well as their Petz predecessors) are designed to be regularly played with over the course of many weeks or months. Over time the Babyz will change and develop, eventually able to walk, talk, and understand a few spoken words. They may change how they feel about and behave towards the user and each other. Over the long-term the hope is that users will suspend their disbelief that these are artificial characters, and bond with their virtual Babyz, forming rewarding emotional relationships with them.

Babyz have an persistent fuzzy “association matrix” memory, where they keep track of how positively or negatively they feel towards the user and the other Babyz and objects they encounter. This memory is constantly being updated as they interact with their environment. When Babyz first meet, their initial attitude towards one another is established as a function of how compatible their particular personalities are, how well their first meeting goes (which can be influenced by the user), as well as some randomness. If two Babyz feel negatively towards each other, this can begin a long-term sibling rivalry, where they take turns playing pranks on each other over time. (The previous section on short-term narratives describes an example of a prank narrative.) At first a rivalry begins with simple mild pranks such as stealing a toy, or startling the other Baby by saying “boo”. But over time the pranks get more and more elaborate, such as throwing objects at each other, knocking over building blocks, messy food fights, and so on. The rivalry can continue indefinitely this way, or subside from user intervention such as keeping these Babyz apart to prevent them from fighting, or praising them in the moments when they are together and not fighting.

A long-term rivalry narrative is not implemented as a behavior goal, as short-term narratives are. The history of a rivalry is kept track of with a simple persistent episodic memory, which is queried by the story manager to decide what short-term prank narrative goals to spawn next, and when. It is the overall continuity of related short-term narratives executed regularly over time (days) that constitute the long-term narrative.

Another set of long-term narratives in Babyz is the development of skills. (Note that all skill behaviors in Babyz are pre-authored, with the user’s interaction "unlocking" them over time, to create the illusion that the Babyz are learning.) Over the course of several weeks, through the help of user encouragement, Babyz can advance from crawling on their stomachs to crawling on all fours, to taking their first steps, to walking. When first adopted Babyz only make simple cooing and gurgling sounds, but if stimulated by the user's voice, they will begin trying to pronounce single syllables, and eventually become able to say simple words in a baby-talk fashion. Over time Babyz can graduate from exclusively suckling on a milk bottle for nourishment to eating food from a spoon, and feeding themselves. At first Babyz will only timidly tap at a toy piano or drum, but with the user's encouragement they will begin playing simple songs, eventually in synchrony with the user and each other.

Like the rivalry long-term narrative, the overall continuity of related short-term narrative goals executed over time is what constitutes the long-term development narrative. Each of these development stages are kept track of by a simple set of persistent variables, to which the behavior and animation systems constantly refer for modifying how goals and plans are performed.

Working towards "good" interactive stories

The experience of interacting with Babyz and Petz was made purposefully open and unstructured, giving users the freedom to play and socialize in their own way and at their own pace, encouraging them to come up with their own interpretation of the characters' feelings and thoughts. This is unlike many video and computer games which tend to overly restrict what the player can do at any one time, requiring them to finish one "level" before advancing on to the next. But this amount of freedom is also unlike the structure of most stories, especially dramatic ones, which tend to carefully and tightly follow a narrative arc from inciting incident to crisis to climax. How do the narratives from a looser, more "character-centric" approach to story compare to those from a more rigid "plot-centric" one?

A goal for many of us in the field of interactive virtual characters and narrative intelligence is to eventually create interactive experiences as (or more!) powerful than those in traditional story mediums such as books, movies, TV, and theater. Our favorite non-interactive stories from the past can serve as examples of how "good" a story can theoretically be - that is, how engaging, moving and affective it can be. Of course we would like to achieve this kind of affective power in our virtual worlds, but achieving it with non-trivial interactivity is very difficult. Those carefully-crafted plot structures which contribute so much to making good stories so powerful are not very pliable. To add significant interactivity to narrative, we believe it is helpful (at first) to move away from traditional plot structures and towards a looser character-centric experience. (By "significant interactivity" we mean interactivity that causes continuous, meaningful and varied changes to the events of a story, not just a glorified "next" button.)

Aylett (Aylett 1999) points out that as a virtual world becomes closer to real life, it becomes more likely that narrative could emerge from the virtual world as it does from human life experience. For example, just as we tend to tell stories about the events that happen to ourselves on a given day, we could recognize stories in the events that occur in a virtual world. However to achieve this life-like effect, the characters and environments in the virtual world must be endowed with a sufficiently rich set of behaviors. Creating this richness is a fundamental challenge for the authors of virtual characters and environments.

To this end we put all of our effort into creating as many behaviors and behavior-alternates as we could for Petz and Babyz. As a rule we found that creating 6 or 7 alternates for each behavior seemed to reach a critical threshold for the illusion of life, perhaps surpassing the point where users can keep track of behavior repetition. In total, the Petz production team of 4 engineers and 4 animators worked for three years to author approximately 100 short-term narrative goals, constructed from 2000 low-level animation pieces. In one year of work for this first version of Babyz we authored approximately 50 short-term narrative goals, constructed from 1000 low-level animation pieces.

Based on customer feedback in the form of emails and bulletin board postings, and on our own observations when playing with the characters, we feel that the Babyz and Petz characters exhibit a reasonably convincing illusion of life, and allow for occasional short, simple emergent narratives to occur. But two shortcomings of the emergent narrative approach stand out, as Aylett observed. First there is a risk that narrative may not emerge at all, and second, even if it does emerge, it may be boring. In a virtual environment where the control of agents is decentralized and uncoordinated, just as in real life (ostensibly), there is no guarantee that a meaningful and coherent chain of events (i.e., narrative) will occur at all. And if it does occur it may only be a fragment of what we consider a "good story". In our own evaluation of Petz we certainly found many moments when the experience seemed fragmented and even boring. We attempted to address this shortcoming in Babyz with the addition of a few explicit long-term narrative behaviors as described earlier, which are deliberately spawned when it is detected that no emergent narratives have occurred recently. (However, Petz probably has an overall richer potential for narrative than Babyz, because Petz currently has twice the amount of behavior and animation content.)

We believe that given enough time to carefully craft behaviors, and the time to create a sufficient number of them, the character-centric approach makes it possible to create an interactive narrative experience that is more eventful and entertaining than mundane "real life" experience, but inevitably with less economy and efficiency than a tightly constructed story (e.g., drama). Experiencing such a virtual world would be akin to "hanging out" with someone who is a talented improvisational actor, who is always trying their best to make entertaining and exciting things happen. However, if it is the artist's goal to create interactive *drama*, we believe one cannot rely primarily on emergent narrative and must adopt a more centralized approach to controlling the characters. In future work we hope to integrate the looser character-centric approach with a more structured plot-centric approach, while attempting to keep the same level of freedom and interactivity that the character-centric approach offers (Mateas & Stern 2000).

Conclusion

This paper has described the overall design philosophy behind Babyz and gave examples of short-term and long-term narrative experiences it offers. During implementation we found it expedient to borrow techniques from several disciplines including artificial intelligence, artificial life, videogames, improvisational acting, and traditional cartoon animation. Our goal was to create a complete, polished, mass-appeal interactive entertainment product, as well as to make progress in the areas of lifelike computer characters and interactive narrative.

In *Hamlet on the Holodeck: The Future Of Narrative in Cyberspace*, Janet Murray suggests that interactive virtual characters "may mark the beginning of a new narrative format", taking on the task of redefining what it means to be human in the face of artificial intelligence (Murray 1997). By endowing our new set of virtual characters with some explicit narrative intelligence, as well as some simple natural language capability, we hope that we are one baby step closer to achieving that goal.

Acknowledgments

Babyz was made possible by a passionate team of designers, programmers, animators, artists, producers and testers at PF. Magic / Mindscape that include Adam Frank, Rob Fulop, Ben Resner, John Rines, Mike Filippoff, Andrew Webster, Jan Sleeper, Jared Sorenson, Dave Feldman, Neeraj Murarka, Bruce Sherrod, Bret Berry, Laurie Sirois, and Darren Atherton.

References

Aylett, Ruth (1999). Narrative in Virtual Environments: Towards emergent narrative. In M. Mateas & P. Sengers (Eds.), *Proceedings of the 1999 AAAI fall symposium on Narrative Intelligence*, FS-99-01 (pp. 83–86). Menlo Park: AAAI Press.

Bandai (1996). Tamagotchi keychain toy. Website URL: http://www.bandai.com.

Bates, Joseph (1992). The nature of characters in interactive worlds and the Oz Project. Technical Report CMU-CS-92-200, School of Computer Science. Pittsburgh: Carnegie Mellon University.

Frank, Adam, Andrew Stern, & Benjamin Resner (1997). Socially intelligent Virtual Petz. *Proceedings of the 1997 AAAI fall symposium on socially intelligent agents*, FS-97-02 (pp. 43–45). Menlo Park: AAAI Press.

Grand, Stephen, David Cliff, & Anil Malhotra (1997). Creatures: Artificial Life autonomous software agents for home entertainment. *Proceedings of the first intl. conference on autonomous agents* (pp. 22–29). Minneapolis: ACM Press.

Mateas, Michael (1999). Not your grandmother's game: AI-based art and entertainment. *Proceedings of the 1999 AAAI spring symposium on Artificial Intelligence and computer games*, SS-99-02 (pp. 64–68). Menlo Park: AAAI Press.

Mateas, Michael & Andrew Stern (2000). Towards integrating plot and character for interactive drama. *Proceedings of the 2000 AAAI fall symposium on socially intelligent agents: The human in the loop* (pp. 113–118). Menlo Park: AAAI Press.

Murray, Janet (1997). *Hamlet on the Holodeck: The future of narrative in cyberspace*. New York: The Free Press.

PF Magic / Mindscape (1995–1999). Virtual Petz website, Virtual Babyz website. Website URLs: http://www.petz.com, http://www.babyz.net.

Stern, Andrew, Adam Frank & Benjamin Resner (1998). Virtual Petz: A hybrid approach to creating autonomous, lifelike Dogz and Catz. *Proceedings of the second intl. conference on autonomous agents* (pp. 334–335). Menlo Park: AAAI Press.

Stern, Andrew (1998). Interactive fiction: The story is just beginning. *IEEE intelligent systems, 13* (5), 16–18.

Stern, Andrew (1999). AI beyond computer games. *Proceedings of the 1999 AAAI spring symposium on Artificial Intelligence and computer games*, SS-99-02 (pp. 77–80). Menlo Park: AAAI Press.

Stern, Andrew (2000). Creating emotional relationships with virtual characters. *Austrian Research Institute for Artificial Intelligence workshop on "Emotions in humans and in artifacts"*. Cambridge: MIT Press. Publication forthcoming.
Tiger Electronics (1998). Furby toy. Website URL: http://www.furby.com
Weizenbaum, Joseph (1966). Eliza. *Communications of the ACM*, *9*, 36–45.

Chapter 14

Web guide agents

Narrative context with character

Katherine Isbister and Patrick Doyle
Katherine interface, San Francisco; Stanford University, Stanford, CA

Introduction

According to (Schank & Abelson 1995), the bulk of human knowledge and memory is communicated and encoded in story form. Although there has been a long-standing interest in how interface agents might bring social and narrative perspective to information resources (Oren et al. 1990), there is currently little explicit use of adaptive storytelling in the repertoire of those agents. Most of the interface agents in use today are only embodied alternatives to traditional menu- or prompt-driven mechanisms for performing simple tasks such as searching for files or providing context-sensitive help. Although they possess bodies and offer some social context, these utilitarian agents ordinarily lack intelligent believability (Ball et al. 1997; Rist et al. 1997). There has been active research on pedagogical agents that do exhibit intelligent behavior, though they create narratives only in the sense that their pedagogical requirements impose a structured sequence of behaviors (e.g. (Rickel & Johnson 1997; Lester & Stone 1997)). Finally, there has been work on the creation of stories through the interaction between the user and agents that are characters in the story (Bates 1992; Hayes-Roth et al. 1997), though in these systems story is the sole purpose of the interaction.

Our work combines our interests in interface agents as characters and characters as participants in narrative. We wish to use narrative and social context explicitly as ways to help users organize familiar ideas, learn new material, and engage with content. In searching for a way to explore these ideas, we each settled on the idea of an interface agent as a “guide” to a digital location. A human guide’s task is to provide information about the nature, content and purpose of an environment. Good guides do this by *interpreting* the environment for a visitor. Interpreting is providing a social context for understanding, often bringing a place to life by using one's own perspective and the artifacts at hand, and usually involves telling the visitors situated stories – stories tied closely to the environment – to make the experience more vivid and emotionally engaging for them (Pond 1993). Thus the guide's task is explicitly a social and narrative one.

In this paper, we will discuss two independent projects that have built social, storytelling agents. Isbister's agent is tightly integrated in an online 3-D Website tour. Her project explores ways to make an agent effectively adapt its narrative to different groups who take the tour, as an expert human tour guide does. Doyle's agent, rather than being bound to a particular site, instead explores different Websites as a persistent companion to a user. One of the intelligent behaviors he is examining is narrative guidance through these sites, guidance assisted by an annotation mechanism that extends standard Web markup. By comparing these projects we have been able to identify some common qualities that a narrative agent should possess, and common issues their designers must address.

How character interactions improve the web experience

One can think of communication as comprising three elements: context, structure, and content. The Web is a vast and growing body of content, but has serious deficiencies in the other aspects of communication. Character interactions can help address these problems by providing a *social* context and a *narrative* structure.

Research suggests that people already unconsciously treat computers as social entities (Reeves & Nass 1996). This is a strong justification for social interfaces. In addition to the benefits of flexibility and error-tolerance, social interactions bring with them a well-understood context for communication. Providing a character with a personality and a social role (such as a museum tour guide) allows us to take advantage of peoples' social strategies for establishing context – their goals for interaction, the relationship between their interests and the character's, the range of appropriate responses, the relevant and valuable information in a dialog – that is, their social filters.

Characters can also provide a consistent *narrative* structure to the interaction. Not only can a character reinterpret the contents of a single page in a narrative style, but the persistence of the character means that the whole course of a Web experience can be recalled and reasoned about in narrative terms. Dialog with the character is a way to maintain persistence of memory, and to draw relationships between what was encountered, what is being examined, and what future options might be appropriate.

Our experiments examine both aspects of communication. Isbister's guide focuses on the social context for presenting an environment to the user; Doyle's guide constructs a narrative structure for the museum pages it visits. Both improve the user's communication with the Web.

The experiments

Each of us has independently embedded characters in the Web to examine the usefulness of our ideas. Both characters serve as guides and companions to users exploring an entertaining or educational environment.

In order to foster communication, both of us have chosen characters that are not at all photo-realistic. Simplified characters with exaggerated features and gestures are easier to "read" very quickly, and to interpret (Thomas & Johnson 1981). Cartoon characters take advantage of stereotypes about personal characteristics, as well as about how to express various emotions and attitudes (Blair 1994). They can act as social signifiers very quickly because of this, rather than requiring a long period in which one "gets to know" the character.

Cartoon-style characters also typically tend to display only very simple and obvious motivations and reactions to situations. In this way they are well-suited to the limited interactivity that the current Web environment could provide for engaging with them. By using simple characters, we set up expectations for simple interactions with them about their environment and about themselves.

Kyoto digital city tour guide

Isbister is currently part of a team that is building a digital version of Kyoto, Japan (http://www.digitalcity.gr.jp). This online city is meant to be an evolving resource both for outsiders and for Kyoto residents. One of Isbister's tasks is to ensure that the Kyoto digital city will be an inviting place for real people to use. As a visitor herself, she is focusing on making the site approachable and engaging for outsiders who might someday want to visit Kyoto. To help accomplish this, she is creating an agent-led group chat tour of the city. The tour will be a point of entry to the online resource and to Kyoto, ideally increasing visitor interest in and use of the digital city. Isbister hopes the tour will also encourage dialogue and relationships among those who participate, and increase exposure to Kyoto's history among those who are friends and family of people who participate in the tour.

In creating the agent, Isbister is focusing on storytelling strategies that will produce an engaging experience for tour takers. She derived a list of abilities for the agent by researching the behavior of actual tour guides in Kyoto. Isbister hopes the process of developing the agent's storytelling abilities will lead to a contribution to the narrative intelligence/agent research domain, specifically involving timing and duration strategies for situated storytelling, especially to groups.

Researching the tour guide role

To prepare, Isbister went on several guided tours of Kyoto, making notes about how tour guides did their work. She found that tour guides made use of illustrative stories frequently, supplementing the rich visual environment of the city with explanations of how Japanese people, both past and present, made use of these settings. Stories included things such as: descriptions of how a given site was constructed and its history of destruction and reconstruction; descriptions of peak historic events that happened at a given site; and descriptions of seasonal events and customary activities that occur at the site. Tour guides also reincorporated material from previous tours, describing what other visitors said and did when they visited the sites.

The tour guides would introduce the stories upon arrival at the site, and at specific points in the site that were directly relevant to the particular story. While visitors took things in visually, the guide would create a narrative context, providing visitors with stories they could share with fellow tour members, as well as with people back home.

The tour guide would time his or her storytelling to the visitors' absorption with the site at hand. That is, the guide would provide story context while the visitors were engaged, and move on when it became clear that the visitors were ready for a new venue. In addition, the guide would provide follow-up stories to those that were met with particular interest by visitors.

To summarize, Isbister observed the following qualities in human tour guide storytelling to groups:

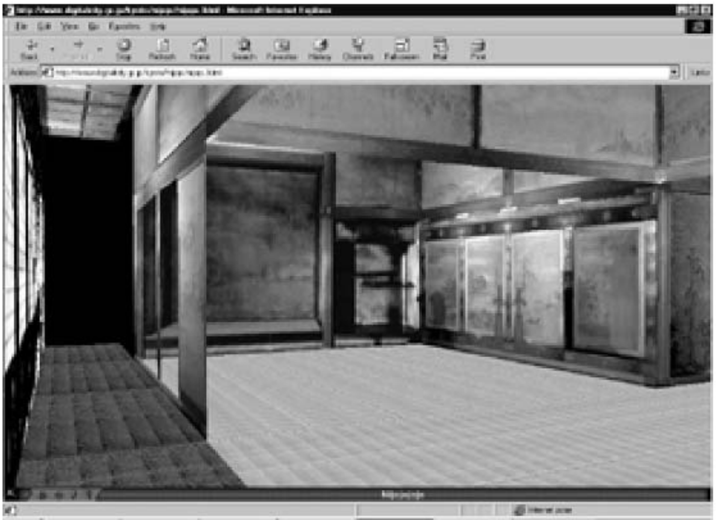

Figure 1. Screenshot of the Shogun's private quarters, one of the 3DML stops on the tour.

1. Stories were told about particular locations, while visitors looked at those locations.
2. Some stories included material that tour guides reincorporated from previous tours.
3. Stories that the guide selected seemed to be those that would lead to easy retelling by visitors to others.
4. Guides adjusted the timing of storytelling as well as choice of follow-up stories based on visitor interest level.

Tour and guide design

Isbister used the criteria above to guide the development of a prototype group tour guide, which leads a tour of Kyoto's Nijo Castle on the Web.

The Kyoto Tour Guide project has four main components: (1) 3-D explorable tour sites; (2) a database of gesture-choreographed stories that relate to these sites, which are performed by the Microsoft Agent; (3) a commercial chat server (I-chat); and (4) an agent, written in perl, that drives the tour and the performance of the Microsoft Agent.

(1) 3-D tour sites: The prototype tour currently consists of three 3-D sites. All of the sites are locations in Kyoto's Nijo Castle. Nijo Castle served as the

Kyoto base for the Tokugawa Shoguns during the Edo period in Japan. (see Figure 1).

(2) Story database: We have created a database of stories that are related to the sites on the tour. These stories have been crafted according to above criteria: anecdotes are selected that relate closely to the particular sites, and that reveal interesting and easily-retold information about Japan's history and culture.

Each story includes both dialogue and gestures that the Microsoft Agent character will perform. The gestures add emotional as well as informational value to the stories (Isbister & Ishida 1999). The agent's spoken delivery of the dialog allows visitors to devote more of their visual attention to the site that they are exploring. The stories are HTML files containing JavaScript commands that drive the MS Agent. The database currently includes three versions of each story – short, medium, and long – to be delivered by the guide depending upon the level of user interest and activity during the tour. The story database is implemented using PostgreSQL ver. 6.4.2, running on a Sun with the Solaris operating system. Currently, the database is a simple table that allows lookup and selection of stories based on length, title, and the tour stop to which they relate. In the future, we plan to categorize these stories by content type as well, to allow for adaptive story telling based on tour-takers' specific topic interests. We would also like to develop a way to rank the stories based on user feedback as to which are the most enjoyable and interesting, to help keep the tour fresh and reflective of the latest visitor interests.

(3) Chat server: The tour is hosted using a commercial chat server called I-chat. I-chat makes it easy to associate particular Web pages with chat rooms, and to push new pages to all chat room participants (see Figure 2). Tour takers log into our local I-chat server and are in the same chat room for the duration of the tour. I-chat's implementation also makes it easy to create an agent that can log into the chat environment in the same way that a user does. The perl agent that drives our guided tour is logged into the tour chat room, and can easily monitor and log user activity, in order to make the story selections that it will push as HTML pages to all users' Web browsers.

(4) Tour agent: The tour agent, written in perl, makes decisions about what story to tell at each tour stop. This agent is logged into the chat room, and is able to monitor the group's conversation. The current implementation tracks the quantity of conversation, and looks for positive and negative keywords that indicate how visitors feel at the moment (negative words include words such as "boring, dull, too long"; positive words include words such as "wow, cool, neat, interesting"). The current prototype agent selects stories using a very sim-

Figure 2. Prototype of the Kyoto digital city tour.

Table 1. Decision rule for prototype tour guide agent

	Valence of Conversation Contents	
Quantity of Talk	***Negative***	***Positive***
Low	medium length	long length
High	short length	medium length

ple decision rule (see Table 1). We plan to adjust the agent's decision-making mechanism after examining its initial performance with tour-takers.

To make sure the tour stops long enough (but not too long) at each tour location, the agent asks visitors (through dialog delivered by the Microsoft Agent) to provide an explicit verbal cue that they are ready to go on. The agent moves on to the next tour stop when it gets feedback from the majority of tour-takers that they want to move forward. We believe this explicit request for feedback serves two purposes: it allows the tour-takers to modulate the stop length far more subtly than the agent could, and it gives them a feeling that the tour (and guide) are adjusting to them – that they have an influence on events, and needn't feel hurried or bored.

There is also a post-tour questionnaire that asks visitors who have taken the tour to give feedback about the stories and the experience itself. As mentioned above, we hope to develop a way to feed this information back into the ranking and selection of stories by the tour guide agent.

An annotation-sensitive tour guide

Doyle has arrived at this problem from a different direction. Given the growing size and popularity of the Internet, his concern has been to find mechanisms that enable interface agents of varying degrees of sophistication to act intelligently and believably across a wide array of sites. His approach has been to add *annotations* to these sites. Annotations are declarations and procedures embedded in the environment and made accessible to wandering agents. They provide useful information about available content and activities, as well as assisting agents in the choice and timing of their actions on these sites.

Previous explorations have taken place in text-based environments called MUDs, where agents have used annotations to solve puzzles, play games, and enhance their emotional intelligence with respect to the environment (Doyle & Hayes-Roth 1998a, b). While the World Wide Web does not offer as sophisticated or flexible a platform for agent behavior, its ubiquity and obvious amenability to simple forms of annotation make it a natural choice for exploring this approach.

Web sites are organized collections of documents, but frequently the organization is opaque to the visitor. Usually this is because Web browsers have no good affordances for revealing complex structure of sites. We have not yet gone beyond the metaphors of the desktop or the card stack. The introduction of a character, however, provides the designer with another kind of interface to the user, one that can construct a narrative presentation that may be more natural and easy to comprehend. If people understand the world in terms of stories, our tendency will be to create or associate stories with the text and images we encounter as we navigate the site. This suggests that the site's designer may be able to communicate more effectively by actively offering narratives to site visitors, rather than requiring them to create their own.

Doyle's work on annotations for enabling agents and enhancing sites seemed a natural mechanism for exploring this idea. By embedding details of the site's content as well as actual stories on the site in a form that the agents can make use of, we can produce an adaptive guide that has both knowledge of the user and an understanding of the site. This guide can entertain or edify according to the user's interests while at the same time furthering the annotation designer's goals for the site.

Figure 3. Merlyn in the Briar Rose Room.

The agent operates in a split-frame window of a Web browser. The upper frame, occupying most of the window, contains the current site's content. The lower frame holds JavaScript code for interfacing between the agent, the visitor, and the site, and displays buttons the visitor can use to command the agent. The agent's architecture (its intelligence, behaviors, and annotation handling) is written in Java. Hooks to the Microsoft Agent API are used to provide the animated character and speech synthesis. The lower frame persists as the user navigates the Web, so while the agent is an extra layer of interface beyond the standard browser, it does not prevent the user from visiting any Web sites, whether or not they are annotated.

Web annotations are represented in the Extensible Markup Language (XML). Every page on an annotated site has a corresponding XML file containing its annotations. Whenever a user visits a new page on the site, the guide agent requests the annotation file from the server, parses it, and adds that information to its local store. If there are no annotations available, or if they are useful only to other kinds of annotation-sensitive agents with different abilities or goals, the agent can still fall back on its built-in repertoire of behaviors. The agent does not require that the annotations be there, but it will make use of them if they are.

A virtual art museum

Our current testbed on the Web is an art museum. The museum consists of a set of galleries, each of which contains rooms filled with artworks. There are presently two galleries: the *Nativity Exhibit* houses medieval and Renaissance religious art revolving around the birth of Christ, while the *Pre-Raphaelite Exhibit* features 19th century artworks in the style of that movement.

Within these galleries, annotations provide a visiting agent with details about their artworks, including their historical contexts, the lives of their artists, their relationships to other artworks in the gallery, and details about specific characteristics of the art. Any agent familiar with the artwork ontology can immediately extract this information for whatever purpose.

The Web agent used in the virtual art museum is a character called Merlyn, aptly named after the forgetful old wizard of T.H. White's *The Once and Future King*. Figure 3 shows Merlyn in the Briar Rose Room of the Pre-Raphaelite gallery, in the process of telling a story to the user.

Merlyn's purpose is to explore the art museum together with a child. To do this he uses the annotations in several ways. First, he can describe the paintings he and the child see as they travel through the museum; he can provide information about when they were made, by whom and how. This information is available both on demand and through his autonomous lecturing behaviors. Since he retains these annotations as he travels, he will also be able to refer back to them if they relate to current topics ("Remember the other picture of Tristan we saw?")

He can also use the annotations to play games with the user; one simple example Merlyn can play is "I Spy," the children's game in which participants take turns guessing what object one child is thinking of.

The agent's third major function is storytelling. Merlyn uses his internal database or the annotations in the museum as his sources. One of the implemented tales in the Pre-Raphaelite gallery, for example, is the story of Sleeping Beauty, which is told using Edward Burne-Jones' four "Briar Rose" paintings as illustrations. This is the story Merlyn is reading to the user in Figure 3. He will autonomously offer to tell stories when he encounters new ones, and the user can direct him to stop or move forward or backward through the story through the buttons on the control panel. Merlyn uses the combination of an internal timer and a sensor that monitors the user's actions and will offer to proceed if the user appears to be bored or finished with the page.

In the art museum ontology, a story is a kind of *tour*, and consists of a sequence of artworks on Web pages, each of which has one or more pieces of

```
<TOURS>
 <TOUR NAME="Sleeping Beauty"TYPE="story">
 <TOPIC>the legend of Sleeping Beauty</TOPIC>
  <TEXT TARGET="ALL">
    Once upon a time, there lived a
    beautiful princess in a great castle…
    . . .
  </TEXT>
  <STOP>
   <NAME>Briar Rose Room South Wall</NAME>
   <URL>gallery/briar/briar4.html</URL>
  </STOP>
  <NEXT-STOP>
   <NAME>Briar Rose Room North Wall</NAME>
   <URL>gallery/briar/briar2.html</URL>
  </NEXT-STOP>
 </TOUR>
</TOURS>
```

Figure 4. Fragment of an XML Tour Annotation.

dialog associated with it (Figure 4). Each artwork may be associated with one or more tours, and within each tour different pieces of dialog may be associated with different categories of reader (so the rendition of the Sleeping Beauty story might be different for a child and adult). A tour can be any sequential narrative, and not simply a story; in addition to hearing about Sleeping Beauty, a visitor might also opt to hear the history behind the painting of the Briar Rose series. In either case the active process of narration, we believe, results in a more attentive, more engaged audience, and there is reason to believe it improves recall of the material as well (Lester et al. 1997).

Unlike other annotations that are highly structured text fragments, tour dialog is stored in the annotation files in full English text. While Merlyn can use those fragments to build natural-language descriptions of the art and artists, the tour text quickly moves outside the range of the ontology. Instead, he takes the English text and alters it through simple syntactic manipulations, such as pauses, stutters, or interjections. This allows him to adapt it, albeit in a limited way, to his personality, his emotional state, and to the style of interaction he is using with his audience. In this case we are sacrificing flexibility for the sake of a well-scripted tale. Ultimately we would prefer an agent and a markup language sophisticated enough that we could build English text from a highly structured description of a story, but the magnitude of that problem is well beyond the scope of our work.

Since the agent stores the annotations it encounters, it retains a memory of every story it sees, as well as where these stories came from. He can easily track requests for stories to be retold (or requests to stop telling others), so he can infer from their topics and keywords what other known and annotated locations might be of interest to his interactors. This sort of simple adaptation to the user's preferences is a necessary first step in making these embodied agents relevant and entertaining for their audiences, which is itself necessary if they are ever going to have real value as a new kind of interface.

Challenges

While interactive tour guides offer the possibility of increased user attentiveness, retention, and enjoyment (Lester et al. 1997), these benefits require we make our agents believable to their users, both as living entities and as expert guides. Our research suggests that the following four traits are critical for creating believable and compelling guides. The first three points have emerged from Doyle's work on annotation-based tour guide storytelling; the last has emerged from Isbister's investigation of human tour guides leading group tours.

– *Intelligent reincorporation*. Reincorporation in this context means the reintroduction of ideas or entities that have been seen earlier in the tour, with some reference to how they relate to the topic under discussion. In improvisational theater (Johnstone 1992) it is well understood that reincorporation is a key to building a story satisfying to the audience; a sequence of unrelated events does not make a story. Similarly, we suspect that a key to creating an effective tour narrative is reincorporation of earlier material. This requires that the agent not only track what material has been seen but also when connections are either pedagogically or dramatically effective.
– *Empathy with the content*. (Elliott et al. 1997; Elliott 1998) has argued that understanding of and reasoning about narrative is strongly associated with reasoning about emotion. If we think of stories as descriptions of sequences of events that have emotional associations for the listener, then we can enhance the significance of our tour content to the user by infusing emotion into the presentation. Beyond incorporating emotions in the presented material, we can build our agents to react emotionally to what they present. This not only heightens the significance of these narratives, it also enhances the believability of the tour guide. As demonstrated in (Persson et al., in preparation), clear emotional responses by an interactive character to the material at hand can be instrumental in creating a satisfying relationship between user and character(s).

- *Presentation through personality*. The most obvious flaw of an intelligent character that relies upon pre-written text is that the text must either be designed for that particular personality, rendering it dangerous for other personalities to use, or so devoid of character that the recitation seems stilted and unbelievable. Ultimately, one would desire an annotation language and an agent powerful enough that the concepts could be explained, reinterpreted, and formed into dialog by the agent. Unfortunately, this is still a hard unsolved problem. Syntactic sugar is a simple approximation that frequently provides good results, as users come to associate idiosyncratic verbal behaviors with the agent's personality. Nonetheless, a more intelligent mechanism for integrating content and presentation will ultimately be required for widespread use.
- *Artful timing/delivery.* As noted in the section on human tour guides, successful storytelling involves knowing when to begin a story, how long it should last, and whether one should elaborate with related stories. Making good decisions about timing and delivery requires the ability to detect user interest, and react appropriately to it. Human beings use many subtle cues to indicate low or high engagement, and detection of user interest is currently very primitive. This will continue to be an important area of research, both for development of characters and for development of satisfactory and subtle interactions with interfaces, in general. We found the mix of user and character control of timing in Agneta and Frida (see Chapter 15) a very interesting and valuable approach to this problem.

Conclusions

Communication requires context, structure, and content. By adding interactive characters to the Web, we can address its structural and contextual limitations. As an example of our ideas, we have created guides and companions that explicitly use narrative to create a social context and to convey the content of Web sites. By telling stories, guides put what might be otherwise dry or overwhelming information into structures that give it meaning and social value. Using reincorporation, emotion and personality, and artful timing, human guides can present information about tour sites in appealing and engaging ways. We anticipate that the lessons we are learning designing interface agents according to these principles will allow us to produce sophisticated storytelling agents, and will be useful to others designing characters to interact in social settings and tasks.

Acknowledgments

Patrick Doyle's work has been done as part of the Virtual Theater Project at Stanford University, under the direction of Barbara Hayes-Roth. He was supported by NSF Grant IRI-9424261, gifts from Microsoft and Intel, and an Intel Foundation Fellowship. Katherine Isbister would like to thank NTT Open Lab and Omron Software researchers who contributed to this project, as well as the students at the Stanford Japan Center.

References

Ball, Gene, Dan Ling, David Kurlander, John Miller, David Pugh, Tim Skelly, Andy Stankosky, David Thiel, Maarten Van Dantzich, & Trace Wax (1997). Lifelike computer characters: The Persona Project at Microsoft Research. In J. Bradshaw (Ed.), *Software agents* (pp. 191–222). Menlo Park, CA: AAAI Press.

Bates, Joseph (1993). The nature of character in interactive worlds and the Oz project. In C. Loeffler (Ed.), *Virtual Realities: Anthology of industry and culture*. New York: Van Nostrand Rheinhold.

Blair, Preston (1994). *Cartoon animation*. Tustin, CA: Walter Foster.

Doyle, Patrick, & Barbara Hayes-Roth (1998a). Guided exploration of virtual worlds. In F. Sudweeks, M. McLaughlin, & S. Rafaeli (Eds.), *Network and netplay: Virtual groups on the internet* (pp. 243–263). New York: MIT Press.

Doyle, Patrick & Barbara Hayes-Roth (1998b). Agents in annotated worlds. In K. Sycara & M. Wooldridge (Eds.), *Proceedings of the second international conference on autonomous agents* (pp. 173–180). New York: ACM Press.

Elliott, Clark, James Lester, & Jeff Rickel (1997). Lifelike pedagogical agents and affective computing: an exploratory synthesis. In M. Wooldridge & M. Veloso (Eds.), *Artificial Intelligence today* [Lecture Notes in Artificial Intelligence, 1600] (pp. 195–212). Berlin: Springer-Verlag.

Elliott, Clark, Jacek Brzezinski, Sanjay Sheth, & Robert Salvatoriello (1998). Story-morphing in the Affective Reasoning paradigm: Generating stories automatically for use with "emotionally intelligent" multimedia agents. In K. Sycara & M. Wooldridge (Eds.), *Proceedings of the second international conference on autonomous agents* (pp. 181–188). New York: ACM Press.

Hayes-Roth, Barbara, Robert van Gent, & Daniel Huber (1997). Acting in character. In R. Trappl and P. Petta (Eds.), *Creating personalities for synthetic actors* (pp. 92–112). Berlin: Springer-Verlag.

Isbister, Katherine & Toru Ishida (1999). Designing for social interaction in cyberspace. In *IPSJ (Information processing society of japan) Magazine*, *40* (6), 569–574.

Johnstone, Keith (1992). *IMPRO: Improvisation and the theater*. New York: Routledge.

Lester, James, Sharolyn Converse, Susan Kahler, S. Todd Barlow, Brian Stone, & Ravinder Bhogal (1997). The persona effect: affective impact of animated pedagogical agents. In S. Pemberton (Ed.), *Proceedings of CHI '97* (pp. 359–366). New York: ACM Press.

Lester, James & Brian Stone (1997). Increasing believability in animated pedagogical agents. In L. Johnson (Ed.), *Proceedings of the first international conference on autonomous agents* (pp. 16–21). New York: ACM Press.

Oren, Tim, Gitta Salomon, Kristee Kreitman, & Abbe Don (1990). Guides: Characterizing the interface. In B. Laurel (Ed.), *The art of human-computer interface design* (pp. 367-381). Reading, MA: Addison-Wesley.

Persson, Per, Kristina Höök, & Marie Sjolinder (This volume). Agenta & Frida: Merging web and narrative?

Pond, Katherine (1993). *The professional guide: Dynamics of tour guiding*. New York: Van Nostrand Reinhold.

Rickel, Jeff & Lewis Johnson (1997). Integrating pedagogical capabilities in a virtual environment agent. In L. Johnson (Ed.), *Proceedings of the first international conference on autonomous agents* (pp. 30–38). New York: ACM Press.

Rist, Thomas, Elisabeth André, & Jochen Müller (1997). Adding animated presentation agents to the interface. In A. Puerta & E. Edmonds (Eds.), *Proceedings of the 1997 international conference on intelligent user interfaces* (pp. 21–28). New York: ACM Press.

Schank, Roger & Robert Abelson (1995). Knowledge and memory: the real story. In R. Wyer (Ed.), *Advances in social cognition*, vol. VII (pp. 1–86). Hillsdale, NJ: Lawrence Erlbaum Associates.

Thomas, Frank & Ollie Johnson (1981). *The illusion of life: Disney animation*. New York: Hyperion Books.

Chapter 15

Agneta & Frida

Merging web and narrative?

Per Persson, Kristina Höök and Marie Sjölinder
Swedish Institute of Computer Science, Kista

Introduction

One of the basic presumptions within narrative theory in general, and discourse psychology in particular, is the notion of a reader constantly striving for *coherence* in his or her understanding and experience of a given text (Graesser et al. 1994; Gernsbacher & Givón 1995; Chatman 1978: 31; Bordwell 1985: 38).[1] Coherence is sought for on many levels of discourse processing. On lower levels, word recognition, grammatical processing and local cohesion are central, and in visual discourse (like narrative films), object, event and situation recognition are equally important. Higher levels of coherence-making include inferring and constructing temporal relations between segments ('this is a flashback') as well as spatial relations ('this event takes place far away from the previous scene'). In addition, readers create and constantly update models about characters' mental life (e.g. intentions, beliefs, emotions) and it seems like most of the causality of a story is present on this level ('X is angry with his wife *because* he believed she was unfaithful to him'; 'Y was killed *because* he knew too much'). On the basis of such inferences, readers form moral judgments of actions and characters ('Z is treating her badly'; 'X is a mean person').

During the course of the story readers also create predictive inferences ('the prince will save her'). At a global level, the gist, theme or morale of the story will be extracted ('love is stronger than death') and the narrative will be interpreted ('Kafka's *The Metamorphosis* is really about how we try to find ourselves in an increasingly ungraspable society'). Finally, readers will make aesthetic judgments about the narrative ('this film really sucks'), speculate about the attitudes and objectives of the author ('the director of this film must be politically aware'; 'this book wants to teach us something'). Oftentimes there will also be a critical evaluation of the theme ('yes I understand that this film tries to tell us that we should seize the day, but I happen to disagree with that philosophy').

With some variation in strength, there is probably emotion and affect present on all of these levels of coherence. Moral judgment of the character behavior, for instance, contains a strong element of emotion and is probably the most fundamental aspect of what we in everyday terms call *identification* and *sympathy* (cf. Smith 1995).

It is important to point out that these inferences and relationships are not primarily textual phenomena, but *mental* ones. They are not 'in the text', but exist rather as readers' constructions during and after the course of the story. They build up to an experience of holism - the feeling that the text 'keeps together' and form a more or less tight structure in which things relate to each other (cf. Trabasso, Suh & Payton 1995). Coherence is, on this account, *accomplished* by the reader through a huge battery of tacit, and hence non-conscious, everyday assumptions, knowledge and prejudices about the perceptual, physical and socio-cultural world. Sometimes, the text supplies the reader or spectator with concrete information, but most of the time the text presents nothing but *cues*, requiring a huge and well-structured system of background knowledge in order to become meaningful. These processes of 'gap-filling' and supplying the context in which a specific text segment becomes graspable are still poorly understood (Graesser et al. 1994: 374). They seem to involve biological, psychological as well as socio-cultural assumptions, making it difficult to maintain the separation between textual structures and the activities of the reader / interpreter (Persson 2003).

The reason why readers construct similar models of coherence is because they share tacit assumptions. Readers differ in their understanding and interpretation of a given narrative, due to the fact that such assumptions differ. The amount of relations a reader manages to establish in a given narrative should also affect the memory of that text. The greater coherence, the better memory.

We wanted to investigate if we could make use of the active construction processes in the reception of narratives in a situation of web browsing. While narrative coherence often is quite tight, the experience of web browsing, on the other hand, is quite fragmented (and not based on *characters* in the way narratives are). We wanted to explore the possibility of merging a web browsing experience with a narrative one.

The Agneta & Frida system

From the web, we collected about 40 actual sites about film production, representing small production companies and organizations as well as local film production collectives. The sites presented information about present productions, production financing, manuscript sales, marketing of films, actors, co-workers, profile of the company, premiers, etc. Some sites contained an extensive body of documents, whereas others were quite simple in structure. Some were professionally designed; others were of a 'home-page' character. All sites were in English, except one, which was authored in Swedish. We created an index of the sites and removed all outgoing links.

To this collected information space we added Agneta and Frida. These two animated female characters – mother and daughter – would sit on the desktop, watching the browser more or less like watching television. With the help of a professional graphic designer and the voices of two actors, we created a library of short audio-visual animations or 'films' that would be triggered by two sets of cues.

First, most comments and behaviors of Agneta & Frida were related to specific locations in the information space. Downloading a page, clicking a link, dragging the mouse over an image or playing a soundtrack would execute general everyday speculations to what something on a site could mean, what the purpose of the site was, or if the design was likeable or not (see Figure 1). Error messages and browser malfunctions would be critically remarked upon. We programmed all of these comments beforehand. There was thus no real 'intelligence' in the system. Often the comments alluded to Agneta and Frida's everyday life and thus provided the user with their 'back story' (e.g., 'that looks like uncle Harry!' [laugh]). Since we envisioned Agneta and Frida to be computer illiterates, their remarks about computer technology and its male dominance were fairly critical and often ironic in tone. In fact, one of the overall purposes of the system was to provide a non-transparent and self-reflexive interface, making apparent the mechanisms by which it works and the socio-economical factors lying behind its origin. Agneta and Frida miss no chance to make visible and make fun of patriarchy, capitalism and techno-geeks.

In this way, we scripted the comments to reflect Agneta and Frida's personalities, ideologies, morals and sense of humor. Since both of them are strong-willed individuals, they occasionally even got into verbal disputes. None of the comments were scripted to be 'helpful', but to invoke laughter and sometimes reflection on the information browsed. Although a user might come back to the same location, none of these behaviors were executed more than once.

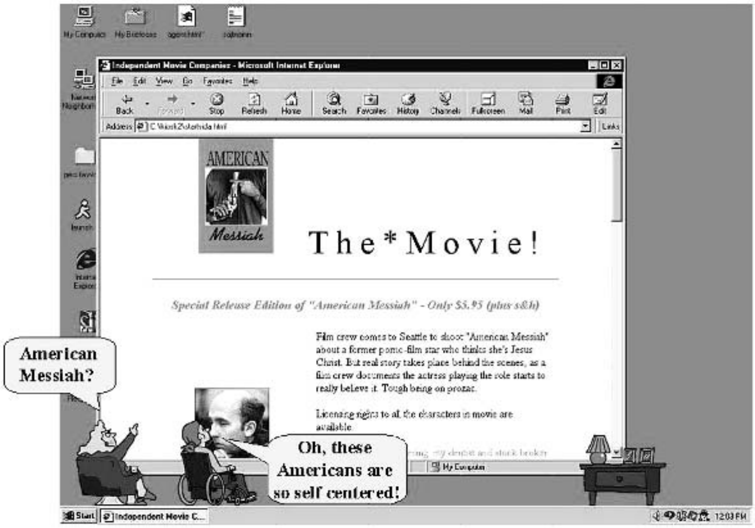

Figure 1. Agneta and Frida reacting to the site of a film production company.

A second set of behaviors / comments was of a more general nature, unrelated to content or user's activity in the information space. This included blinking, picking noses, going to the toilet / kitchen, drinking coffee or general gossiping about uncle Harry and Miss Andersson (the owner of a repulsive poodle that often enters Agneta and Frida's back yard, and about which they occasionally fantasize killing). These would be triggered at certain intervals when there were no other behaviors running. This set of behaviors was included to create more lively characters, with lives independent of the happenings in the browsing session. We wanted to avoid the impression that the behaviors were only automatic reflexes of user's actions. Some behaviors in this category reoccurred now and again (e.g. blinking, yawning and sighing).

In both behavior types, Agneta and Frida spoke English with exaggerated Swedish accents (the Swedish Chef of *The Muppet Show* acted as our leading inspiration here). As we aimed for users between 15 and 30 years of age, the web page contents, the commentaries as well as the jokes of Agneta & Frida were all scripted with that age group in mind.

The user could regulate the intensity of Agneta and Frida's behaviors. Through a pop up menu, the user could set the level of activity (0-5), which in different degrees disconnected some behavior and defined the time interval between the remaining ones. The way to access this menu and the other features of the system were described or hinted at by Agneta or Frida when they were clicked upon.

The Agneta and Frida system explores the possibilities of multiple characters. Having two characters instead of just one made pre-scripted *dialogue* possible. This enabled us to introduce personality, humor, negotiation and self-reflection more naturally than through a single character. In fact, few systems exploit this design choice with the exception of André & Rist (2000).

We implemented these ideas using JavaScript and Microsoft Agent Tool.

The study

While the entire study covered aspects of narrative, navigation, exploration and believability (Höök et al. 2000), this paper will concentrate on the narrative aspect: Did the system succeed in merging the Agneta & Frida narrative with the web browsing experience?

How to measure narrative experience?

One way to assert 'narrativity' is to investigate whether a reader or user actually draws inferences, constructs mental models and makes interpretations of the type discussed in the introduction of this paper. Discourse psychology research about narratives has developed methods to investigate some of these processes, especially on the cognitive side of the matter (Graesser et al. 1994). We choose to develop somewhat other methods.

Our first hypothesis was related to how the user would speak about their experience. If users merged the Agneta & Frida narrative with the information browsing on the sites, we reasoned, they would talk about the total experience as one entity. Thus, after their interaction with the system, we asked the subjects to describe the system as if they were telling a friend about it. Hopefully, subjects would talk about Agneta & Frida as a natural part of the system and relate Agneta and Frida's comments and views to the interaction with the sites. For instance, would Agneta and Frida's comments, jokes and narratives perhaps encourage the user to construct coherence between nodes on the bases of events, situations and psychology of the characters? If so, Agneta & Frida would provide a *social context* to the web information, which would be persistent, helping the user to 'digest' the information. This ambition we seem to share with Isbister & Doyle (Chapter 14).

While the subjects' description of the system would reflect the more conscious aspects of the experience, we employed another method to capture more deep-level dimensions. Inspired by Maglio and Matlock (1999) and Lakoff and Johnson (1980, 1999), we performed a *metaphor analysis* of the interviews. From Maglio & Matlock's study, we knew that web browsing is often perceived as a spatial activity: the user is viewed as an agent moving through the space of sites and web pages. Maglio and Matlock found this by examining the metaphors used when subjects described their surfing through web pages: 'I browse/surf the web'; 'I go to pages'; 'I enter/leave pages'; 'pages contain information'; 'the web is an information space in which I look for things'.

We decided to follow the method used by Maglio and Matlock, focusing on narrative versus spatial verbs and adverbs in the interviews. Spatial verbs and adverbs were characterized by movement (e.g. 'going through'). As for narrativity, we looked for words containing temporal dimensions (e.g. '. . .and then. . .') or intentional/psychological words (e.g. 'giving up'; 'bored'; 'anxious'). In contrast to spatial experiences, we hypothesized, narratives are temporal and causal chains of events and this would have to be reflected in the use of verbs and adverbs.

Finally, we measured *disturbance* and *recall*. If the user was able to integrate the narrative of Agneta and Frida with the web content, we hypothesized, subjects would be less disturbed by the two characters than by a case in which the Agneta & Frida story ran 'in parallel' to the web content. In the latter case, the comments and activities of the characters would be experienced as intrusive. As for recall, we assumed that the emotional reactions caused by the remarks from Agneta & Frida – e.g. laughs, frustration, moral judgment and agreeableness – would enhance the recall of the information remarked upon. We assumed that Agneta and Frida would encourage the user to construct a narrative context and associative links between information in the site, which would improve memory. Thus, we expected the Agneta and Frida subjects to perform better on a post-usage recall test, than would subjects without Agneta & Frida.

Subjects

The 38 subjects were recruited to be between 20 and 30 years. Eighteen subjects used Agneta & Frida (the 'withA&F' group) and 20 subjects explored the web sites without the characters present (the 'withoutA&F' group).

The subjects of the withA&F group were in the range $19 < 26.2 < 41$ years old (seven women and eleven men). Ten had a technical background; the rest had other professions. All but one had a university degree. The subjects in the withoutA&F group were in the range 22 < 26.6 < 32 years old (twelve women and eight men). Fourteen had technical background. All but one had a university degree. All had a good understanding of English.

Subjects were signed up through the experimenters' friends and colleagues and were given movie tickets in return for participation.

Tasks and procedure

The withA&F group was first asked about their age, gender, occupation, education, command of English, as well as their experience of computers and the web.

Subjects in the withA&F group got the following instruction (in Swedish):

> Imagine this situation: you are at home one evening and you have nothing in particular to do and you are not especially tired. You have a fast and efficient computer at home with a good and fast connection to the Internet. A friend has suggested some cool web sites on the net that you might want to have a look at. Check out the web link the same way you would do if you were sitting at home. I'll be in the other room and you can come and get me when you are finished.

The system was started and Agneta & Frida would appear on the screen. The experimental leader would at this point leave the room to allow subjects to feel free to do what they wanted and stay as long as they wanted. Subjects' interactions with the system and facial expressions were video-recorded.

Afterwards they were interviewed about their experience based on the question: "If you met a friend downtown and were asked to describe what this system was all about, what would you say (in as many details as you can)?" Interviews were taped on audiocassette.

They filled in a questionnaire with three sets of questions:[2]

a. Estimated time spent and estimated number of pages visited.
b. A number of Likert-scale questions on whether they perceived Agneta & Frida as believable characters and whether they found them entertaining or disturbing. In total, 14 questions were asked in various ways in order to get at users' experience of the system.
c. Finally, they marked which of a set of 20 statements they could agree with. There were statements such as "I have the same kind of humor as Agneta & Frida", or "I do not like animated characters".

Once the questionnaire was completed, they were shown 38 screen-shots of web pages from the sites and were asked if they had seen them and whether they remembered any jokes Agneta & Frida had made about that web page. The screen-shots were randomly selected from the site. As there were 40 different film sites with a number of web pages for each site, these 38 screen-shots only captured a small set of the entire web space. In average, the subjects had actually seen 17.2 of the 38 pages.

Finally, subjects in the withA&F group were asked to freely comment the system.

In total, it took each subject in average 2 hours to complete the steps. Afterwards we explained the study and the design rationale behind Agneta & Frida.

The withoutA&F group went through the same steps as the withA&F group, except all questions related to Agneta & Frida were removed from the questionnaire, and their facial expressions were not recorded.

Results

Merging the Agneta & Frida story with the web browsing?

The interviews made it clear that subjects had understood both the basic story of Agneta and Frida, as well as appropriated the gist of the web information. On the one hand, subjects inferred that Agneta & Frida were mother and daughter, that they were poor, that Frida was unmarried, and that Frida was more sarcastic and often more knowledgeable than her mother. Here is an example of a subject providing the back-story of the characters (translated from Swedish):

> Yes, they are relatives and they do not really get along that well but they are stuck, they cannot do anything else. If they had the possibility to do something else, they would. Then... Frida is the daughter, right? She started to look at me, she turned around and looked at me when I did something that she did not like, or, I do not know the intention of that... [...] so they seem to be tired of what they are doing.

As many as 50% of the subjects agreed with the statement "I know someone who is like one of the characters."

From the interviews, it was equally obvious that the subjects made sense of the contents of the web pages. They understood that it was a heterogeneous collection of links to independent movie production companies, some of which were quite poor and amateurish. Subjects interested in film production were fascinated by how low-budget production companies manage to survive and release new titles. Thus, occasionally, web information was turned into a narrative.

Nevertheless, were the two 'narratives' weaved into one or did they run separately side-by-side? Did the story of Agneta & Frida provide the narrative context in which the narrative of film production was understood?

On the one hand, the interviews indicated that this had not been the case. Out of the 18 interviews, 55% completely separated their description of Agneta & Frida and the contents of the web pages. Typically, subjects would first talk about the web information and then they would talk about the lives of Agneta & Frida:

> There were lots of names of different film production companies and so I went in and looked at different . . . what they did, what projects they were currently involved in and suchlike. [—] Yes, Agneta and Frida, they... it was a mother and her daughter who was sitting in a wheelchair and they were around all the time and commented on the pictures. (Subject 18)

Only four subjects described situations in which the web page content was connected to what Agneta & Frida had said.

A qualitative analysis of the interviews, however, sketched a somewhat more complex picture. Rather than incorporating the Agneta & Frida story with the web content story, the presence of the two characters tended to be incorporated into the *browsing behavior* and subjects' *interactions* with the web pages.

> I can't remember the name, but there was some site I was at and looked at and then I probably sat there for a quite long time and read, I think, and then, like that, they turned around towards me and then 'But hey, what is going on here.' Now I can't remember which one of them is which, but she said that 'this little gray thing here on the side, that's the mouse – we want to see some action here!'. They make lots of fun comments. (Subject 15)

Some subjects would even change their navigation because of Agneta and Frida's comments:

> I cannot claim that they really helped but. . . they are sometimes quite amusing and sometimes funny, as when I was at *Alcatraz* and then, just as I was about to leave [the site] they commented on people hanging in ropes and then I became interested so I went back to have a look at what it was, who was hanging in which ropes, and quite right, they were hanging in ropes. So that was kind of amusing. (Subject 12)

> One does react on what they say, if they whine then you think that 'aha, this is a boring page' and so then you quickly move away from it. (Subject 13)

Agneta and Frida's comments also encouraged some users to negotiate with web information and web design in ways they would probably not have done otherwise:

> Yes, one of the pages... contained two iron knuckles, that I remember from the [...] first reaction I got when I entered the page. These two – Agneta and Frida – commented that dark blue text on top of black background is a 'no no', and that, I thought, was not the worst problem with the web page. I mean, dark blue text on top of a black background, that is at least possible to read, but the text on the rest of the page was white text with iron knuckles behind and that was completely unreadable. (Subject 7)

Metaphor analysis

Here is an example of a spatial coding of one of the interviews (translated from Swedish):

> *I have surfed on* the Internet and I *came into* a site that dealt with independent movies and there were all sorts of weird places. Oh, and among other things something from Canada, but unfortunately one had to download so much *there* so I gave up that page and *went away from there*. Eh... then there were some different eh film companies and such.

An example of narrative coding of the same text:

> I have surfed on the Internet and I came into a site that *dealt with* independent movies and there were all sorts of weird places. Oh, and among other things something from Canada, but unfortunately one had to download so much in there so *I gave up* that page and went away from there. Eh... *then* there were some different eh film companies and such.

The metaphor analysis revealed that the withA&F group tended to talk about their experience in terms of narrative verbs and adverbs (68% narrative), while the withoutA&F group used more spatial verbs and adverbs (only 45% narrative) – cf. table 1. The difference between the conditions is statistically significant (Mann-Whitney: $p>0.95$).[3] This seemed to indicate that users actually merged the narrative and the spatial into one.

Being disturbed

None of our subjects were indifferent to Agneta & Frida. They raised strong emotions, both positive and negative:

> The animated characters were extremely disturbing and distracted me completely with the effect that I could not concentrate on the browsing and lost

Table 1. Number of spatial versus narrative words (verbs and adverbs) for the with A&F and the without A&F groups.

	No of words in total	No of spatial words	No of narrative words	No of narrative and spatial words	Ratio of spatial versus narrative number of words
With A&F	4088	108	234	342	32% vs. 68%
Without A&F	6402	204	252	456	45% vs. 55%
Total	**10490**	**312**	**486**	**798**	

> interest in the contents of the web pages. The surfing quickly turned into fear of what Agneta & Frida would say next.
>
> The ladies enlighten the atmosphere. Nice with some company so one does not become completely stiff and dry.

33% of the subjects found that Agneta & Frida often stole attention from the information in the web pages, which may indicate that as many as 2/3 in fact managed to incorporate them into their web browsing experience. 77% of the subjects agreed with the statement "It felt good to have two ladies to browse with". Some users on the other hand, got quite irritated and felt that Agneta & Frida got in their way. The question "Characters beneath the browser disturb me when I search for information" rendered the result that 22% were often disturbed, 33% were sometimes disturbed, while 44% were never or almost never disturbed.

Web content recall

There was no difference between the two groups in terms of how much they remembered of the web pages. Out of the 38 randomly selected test pages, the withA&F group remembered 88% of the pages they had seen, while the withoutA&F remembered 89%. Subjects were able to accurately recall the comments Agneta & Frida had made at particular pages. It seems like Agneta & Frida failed to create the context needed to better tie the different sites in the space together into one coherent narrative experience.

Discussion

Subjects in our study did not gracefully merge Agneta and Frida and the web content into one narrative whole. Sometimes they enjoyed the contents of the web pages, sometimes they were amused by the comments by Agneta and Frida, and at some points web browsing and interaction were integrated into the story of the two characters. Although the study taught us important aspects of interaction with embodied believable agents, it did not generate the result we had hoped for. One reason for this may be the fact that we had too high expectations of the creativity of the user in weaving the two narratives together. As mentioned, many of the comments and behaviors of Agneta and Frida were not really related to the web information on the pages, but of a more general nature. One lesson to be learned form this study is that the agents' comments must be tightly connected to the information displayed if disturbance effects are to be avoided. On the face of it, it seems like the systems presented by Isbister & Doyle (Chapter 14) appropriated these features to a greater degree than our system did. This, however, is also related to the objective of the system (e.g. supporting serious wayfinding, guiding or explorative play), and in this respect it seems like the Agneta & Frida system has a somewhat different function than the systems of Isbister and Doyle.

The study also made it clear that evaluation of systems with believable agents needs to take into consideration a larger context of cultural user expectations. Humor, for instance, which was a rather central aspect in the system, is based on personal preferences and socio-cultural dimensions. The question whether Agneta and Frida are likable or not for a given subject is to large degree a function of those contextual parameters, which an evaluation study must take into account. In retrospect, we should have asked more about the preferences of the users, for instance their general attitude towards embodied agents in interfaces.

While the results did not come out as we had hoped, we still believe that the methods developed and deployed could be valuable in evaluating narrativity in interactive systems.

Acknowledgements

The work reported here was done as part of the EC- and SITI-funded PERSONA project. Jarmo Laaksolahti assisted in the implementation. Kristina Tullgren and Jenny Holm conducted the analysis of narrative and spatial verbs and adverbs as part of their Master's work. Jussi Karlgren provided some of the ideas for metaphor analysis. Anna-Lena Ereback helped us with the practical user study work. In particular, we thank the anonymous participants in our study.

Notes

1. The media by which narratives are received may vary (e.g., text, sound, moving imagery and computer games) and each has its own term for denoting the 'receiving end' (e.g., reader, listener, spectator, player and user). We will be using 'text' and 'reader' in the general sense of the terms, to denote narrative structure and reception of that structure independent of media.

2. The full questionnaire (in Swedish) can be found at [http://www.sics.se/~kia/questionnaire.htm].

3. Mann-Whitney was used since usage of spatial versus narrative verbs and adverbs cannot be assumed to be normally distributed. The comparison was made by comparing the number of spatial expressions and the number of narrative expressions, normalized by the number of analyzed verbs and adverbs.

References

André, Elisabeth & Thomas Rist (2000). Presenting through performing: On the use of multiple lifelike characters in knowledge-based presentation systems, *Proceedings of intelligent user interfaces conference, IUI'2000* (pp. 1–8). New Orleans, Louisiana.

Bordwell, David (1985). *Narration in the fiction film*. London: Methuen.

Chatman, Seymour (1978). *Story and discourse. Narrative structure in fiction and film*. Ithaca: Cornell University Press.

Gernsbacher, Morton & T. Givón (Eds.) (1995). *Coherence in spontaneous text*. Amsterdam: John Benjamins.

Graesser, Arthur, Murray Singer & Tom Trabasso (1994). Constructing inferences during narrative text comprehension, *Psychological Review*, *101* (3), 371–395.

Höök, Kristina, Per Persson, & Marie Sjölinder. (2000) Evaluating users' experience of a character-enhanced information space. *AI Communications, 13* (3), 195–212.

Lakoff, George & Mark Johnson (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.

Lakoff, George & Mark Johnson (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic Books.

Maglio, Paul & Teenie Matlock (1999). The conceptual structure of information space. In A. Munro, K. Höök & D. Benyon (Eds.), *Social navigation of information space* (pp. 155–173). London: Springer-Verlag.

Persson, Per (forthcoming). *Understanding Cinema* Cambridge: Cambridge UP.

Smith, Murray (1995). *Engaging characters: Fiction, emotion, and the cinema.* Oxford: Clarendon Press.

Trabasso, Tom, S. Suh & P. Payton (1995). Explanatory coherence in understanding and talking about events. In M. Gernsbacher & T. Givón (Eds.), *Coherence in spontaneous text* (pp. 189–214). Amsterdam: John Benjamins

Chapter 16

Schizophrenia and narrative in artificial agents

Phoebe Sengers
Cornell University, Ithaca, NY

The premise of this work is that there is something deeply missing from AI, or, more specifically, from the currently dominant ways of building artificial agents. This uncomfortable intuition has been with me for a long time, perhaps from my start as an AI researcher, although for most of that time I was not able to articulate it clearly. Artificial agents seem to be lacking a primeval awareness, a coherence of action over time, something one might, for lack of a better metaphor, term 'soul.'

Roboticist Rodney Brooks expresses this worry eloquently:

> Perhaps it is the case that all the approaches to building intelligent systems are just completely off-base, and are doomed to fail.... [C]ertainly it is the case that all biological systems.... [b]ehave in a way which just simply seems *life-like* in a way that our robots never do.
>
> Perhaps we have all missed some organizing principle of biological systems, or some general truth about them. Perhaps there is a way of looking at biological systems which will illuminate an inherent necessity in some aspect of the interactions of their parts that is completely missing from our artificial systems.... [P]erhaps at this point we simply do not *get it*, and... there is some fundamental change necessary in our thinking... [P]erhaps we are currently missing the *juice* of life. (Brooks 1997: 299–300)

Here, I argue that the 'juice' we are missing is *narrative*. The divide-and-conquer methodologies currently used to design artificial agents results in fragmented, depersonalized behavior, which mimics the fragmentation and depersonalization of schizophrenia in institutional psychiatry. Anti-psychiatry and narrative psychology suggest that the fundamental problem for both schizophrenic patients and agents is that observers have difficulty understanding them narratively. This motivates a narrative agent architecture, the Expressivator, which structures agent behavior to support narrative, thereby creating agents that are intentionally comprehensible. The methodology in this chapter integrates the narrative traditions of cultural studies with the technical traditions of Artificial Intelligence, thereby presenting itself, too, as a form of Narrative Intelligence.

The problem

Building complex, integrated artificial agents is one of the dreams of AI. Classically, complex agents are constructed by identifying functional components – natural language processing, vision, planning, etc. – designing and building each separately, then integrating them into an agent. More recently, some practitioners have argued that the various components of an agent strongly constrain one another, and that the complex functionalities classical AI could come up with could not easily be coordinated into a whole system. They offer other construction methodologies instead. In particular, behavior-based AI proposes that the agent should be split up, not into disparate cognitive functionalities, but into behaviors, each of which integrates all of the agent's functions for a particular behavior in which the agent engages. Examples of such behaviors include foraging, sleeping, and hunting.

Even such systems, however, have not been entirely successful in building agents that integrate a wide range of behaviors. Rodney Brooks, for example, has stated that one of the challenges of the field is to find a way to build an agent that can integrate many behaviors, where he defines many to be more than a dozen (Brooks 1990). Programmers can create robust, subtle, effective, and expressive behaviors, but the agent's overall behavior tends to gradually fall apart as more and more behaviors are combined. For small numbers of behaviors, this disintegration can be managed by the programmer, but as more and more behaviors are combined their interactions become so complex that they become at least time-consuming and at worst impossible to manage.

In both cases, divide-and-conquer methodologies lead to integration problems. With classical agents, who are split up by functionality, there are often problems with a functional underintegration. This underintegration manifests itself in various kinds of inconsistency between the different functions, such as not being able to use knowledge for one function that is available for another. For example, the agent may speak a word it cannot understand or visibly register aspects of the world that do not affect its subsequent behavior.

In behavior-based agents, underintegration manifests itself on the behavioral level. These agents generally have a set of black-boxed behaviors. Following the action-selection paradigm, agents continuously redecide which behavior is most appropriate. As a consequence, they tend to jump around from behavior to behavior according to which one is currently the best (a similar observation is made by (Steels 1994)).

What this means is that the overall character of behavior of the agent ends up being deficient; generally speaking, its behavior consists of short dalliances in individual, shallow high-level behaviors with abrupt changes between behaviors. It is this overall defective nature of agent behavior, caused by under-integration of behavioral units, that I term *schizophrenia* and propose to address here.

Schizophrenia is a loaded term. I use it here to draw attention to important connections between current approaches to agent-building and the experience of being schizophrenic in institutional psychiatry. In next two sections, I draw out those connections, then show how an alternative approach to psychiatric schizophrenia can motivate changes in AI practice. These changes form the basis for narrative agent architecture.

Schizophrenia

Schizophrenia's connection to AI is grounded in one of its more baffling symptoms – the *sentimente d'automatisme*, or subjective experience of being a machine (Janet, 1889). This feeling is the flip side of AI's hoped-for machinic experience of being subjective, and is described by one patient this way: " 'I am unable to give an account of what I really do, everything is mechanical in me and is done unconsciously. I am nothing but a machine' " (an anonymous schizophrenic patient; cited in (Ronell 1989: 118)). R. D. Laing describes how some schizophrenic patients experience or fear experiencing themselves as things, as its, instead of as people (Laing 1960). Schizophrenia is, for some, a frightening feeling of being drained of life, of being reduced to a robot or automaton.

This feeling of mechanicity is correlated with a fragmentation of the affected patient's being; sometimes, a schizophrenic patient's very subjectivity seems to be split apart.

> In listening to Julie, it was often as though one were doing group psychotherapy with the one patient. Thus I was confronted with a babble or jumble of

> quite disparate attitudes, feelings, expressions of impulse. The patient's intonations, gestures, mannerisms, changed their character from moment to moment. One may begin to recognize patches of speech, or fragments of behaviour cropping up at different times, which seem to belong together by reason of similarities of the intonation, the vocabulary, syntax, the preoccupations in the utterance or to cohere as behaviour by reason of certain stereotyped gestures or mannerisms. It seemed therefore that one was in the presence of various fragments, or incomplete elements, of different 'personalities' in operation at the one time. Her 'word-salad' seemed to be the result of a number of quasi-autonomous partial systems striving to give expression to themselves out of the same mouth at the same time. (Laing 1960: 195–196)

Laing goes on to describe Julie's existence in ways that are eerily similar to the problems with autonomous agents we discussed in the last section: "Julie's being as a chronic schizophrenic was... characterized by lack of unity and by division into what might variously be called partial 'assemblies', complexes, partial systems, or 'internal objects'. Each of these partial systems had recognizable features and distinctive ways of its own" (197). Like the parts of behavior-based agents, each subsystem exists independently, with its own perception and action. Subsystems communicate, in Brooks' phraseology, 'through the world,' not by being integrated as a unified whole:

> Each partial system seemed to have within it its own focus or center of awareness: it had its own very limited memory schemata and limited ways of structuring percepts; its own quasi-autonomous drives or component drives; its own tendency to preserve its autonomy, and special dangers which threatened its autonomy. She would refer to these diverse aspects as 'he', or 'she', or address them as 'you'. That is, instead of having a reflective awareness of those aspects of herself, 'she' would *perceive* the operation of a partial system as though it was not of 'her', but belonged outside. (198)

In this sense, there is a direct link between schizophrenia and behavior-based methodology – and symptomatology.

Depersonalization

While we can presume that artificial systems do not particularly care about being fragmented, for schizophrenic patients this feeling of coming apart, of losing life, of being reduced to a machine, is intensely painful. It is therefore ironic that, as a number of critics have argued, psychiatric institutions themselves reinforce this feeling of mechanicity and lack of autonomous self. For example, Erving Goffman, in his ground-breaking anthropological study *Asylums* (Goffman 1961), argues that a major feature of psychiatric institutions is the "programming" of each inmate "into an object that can be fed into the administrative machinery of the establishment, to be worked on smoothly by routine operations." (16)

One of the signs of this mechanization is the reduction of patient to symptomatology. Patients are constantly monitored, their behavior continuously being examined for and interpreted as signs of illness. The patient's actions only function insofar as they are informational – they only *act* as ciphers, which it is then the responsibility and right of the doctor to decode. Rather than being taken seriously as such, a patient's words are used to place the patient in the narrative of the doctor's diagnosis. "When you spoke, they judged your words as a delusion to confirm their concepts" (Robear 1991: 19).

Understood symptomatically, the patient's subjective experience is ignored. Susan Baur describes this limitation of the institutional approach to mental illness:

> I... believe that the medical model of mental illness excludes too much of the patient. Using this model, only parts of the patient are considered, and even when these parts are assembled by a multidisciplinary team into a manikin of a schizophrenic or of a manic-depressive, the spirit that animates the real person gets lost. Especially in chronic cases where mental illness and the desperately clever adaptations it inspires have become central to an individual's personality, the patient's own story and explanations – his delusions and imaginary worlds – must be included (Baur 1991: 105–106).

The patient is formalized, reduced to a set of somewhat arbitrarily connected symptoms. The patient is no longer a living, unique, complex individual, but fragmented into a pile of signs: "she is autistic," "she shows signs of depersonalization," "she lacks affect."

This fragmentation into symptoms, psychiatrist R. D. Laing argues, actually *reinforces*, rather than treats, schizophrenia. When mechanistic explanations reduce the patient to a bundle of pathological processes, the patient as human is rendered incomprehensible. Laing argues that institutional psychiatric practice cannot fully understand schizophrenia because it actually *mimics* schizophrenic ways of thinking, depersonalizing and fragmenting patients.

> The most serious objection to the technical vocabulary currently used to describe psychiatric patients is that it consists of words which split man up verbally in a way which is analogous to the existential splits we have to describe here.... [W]e are [then] condemned to start our study of schizoid and

> schizophrenic people with a verbal and conceptual splitting that matches the split up of the totality of the schizoid being-in-the-world. Moreover, the secondary verbal and conceptual task of reintegrating the various bits and pieces will parallel the despairing efforts of the schizophrenic to put his disintegrated self and world together again. (Laing 1960: 19–20)

By studying schizophrenics in isolation and in parts, psychiatry threatens to itself become schizophrenic, and schizophrenics incomprehensible.

This problem of conceptual splitting parallels closely the problem of AI, suggesting that mechanistic explanations of the sort necessary to build agents are also responsible for their de-intentionalized appearance. The symptomatology of institutional psychiatry is reflected in behavioral black-boxing in behavior-based AI. In the next section, we will explore alternatives to this fragmentation in psychiatry, searching for clues for dealing with the problem of schizophrenia in AI.

Anti-psychiatry and narrative psychology

In the '60's and '70's, Laing and other sympathetic colleagues, termed 'anti-psychiatrists' for their opposition to mainstream psychiatry, suggested that the schizophrenizing aspects of institutional psychiatry can be avoided by changing our viewpoint on patients: instead of thinking of schizophrenics as self-contained clusters of symptoms, we should try to understand them phenomenologically, as complex humans whose behavior is meaningful. The schizophrenizing clinical approach reifies the patient's behavior into a cluster of pathological symptoms, with no apparent relation to each other or the patient's broader life experience. "[S]he had auditory hallucinations and was depersonalized; showed signs of catatonia; exhibited affective impoverishment and autistic withdrawal. Occasionally she was held to be 'impulsive.' " (Laing & Esterson 1970: 32) The phenomenological approach, on the other hand, tries to understand the patient's experience of herself as a person:

> [S]he experienced herself as a machine, rather than as a person: she lacked a sense of her motives, agency and intentions belonging together: she was very confused about her autonomous identity. She felt it necessary to move and speak with studious and scrupulous correctness. She sometimes felt that her thoughts were controlled by others, and she said that not she but her 'voices' often did her thinking. (Laing & Esterson 1970: 32)

Anti-psychiatrists believe that statistics and symptomatology, the foundations of institutional psychiatry, are misleading because they reduce the patient to a mass of unrelated signs. Instead of leading to a greater understanding of the patient, the patient's subjective experiences are lost under a pile of unconnected data.

> It is just possible to have a thorough knowledge of what has been discovered about the hereditary or familial incidence of manic-depressive psychosis or schizophrenia, to have a facility in recognizing schizoid 'ego distortion' and schizophrenic ego defects, plus the various 'disorders' of thought, memory, perceptions, etc., to know, in fact, just about everything that can be known about the psychopathology of schizophrenia or of schizophrenia as a disease without being able to understand one single schizophrenic. Such data are all ways of *not* understanding him. (Laing 1960: 33)

These insights are underscored by the perspective of narrative psychology, an area of study developed by Jerome Bruner (Bruner 1986) (Bruner 1990) (see Chapter 3) which focuses on how people interpret specifically intentional behavior. Narrative psychology shows that, whereas people tend to understand inanimate objects in terms of cause-effect rules and by using logical reasoning, intentional behavior is made comprehensible by structuring it into narrative or 'stories.' Narrative psychology suggests that this process of creating narrative is the fundamental difference between the way people understand intentional beings and mechanical artefacts.

That is to say, if I want to understand and build an inanimate object, I may decompose it, try to understand what different pieces are for, replicate how they work, and figure out the rules underlying its behavior. On the other hand, if I want to understand a person's behavior, I am interested in such things as what motivates him or her, the reasons he or she engages in particular activity, and how his or her behavior reflects on his or her whole personality.

This contrast between narrative explanations that explore the meaning of living activity and atomistic explanations that allow for the understanding and construction of mechanical artifacts echoes the criticisms of anti-psychiatry. Anti-psychiatrists, after all, complain that the difficulty with institutional psychiatry is that it reduces the patient to a pile of data, thereby making a machine of a living person. The anti-psychiatric solution of interpretation uses narrative understanding to 'repersonalize' patients: structuring and relating the 'data' of a patient's life into the semi-coherent story of a meaningful, though painful, existence; focusing on the patient not as an instance of a disease but as a particular individual and how that person feels about his or her life experience; and relating the doctor's narrative to its background conditions and the life context in which it is created and understood. It is only through this process of narrative interpretation that anti-psychiatry feels the psychiatrist can fully respect and understand the patient's subjective experience as a human being.

In AI, this distinction between mechanism and intentional being becomes problematic. AI agents should ideally be understandable both as well-specified physical objects and as sentient creatures. In order to understand intentional behavior, users attempt to construct narrative explanations of what the presumed intentional being is doing; but this approach conflicts with the mechanistic explanations designers themselves need to use in order to identify, structure, and replicate behavior. The resulting abrupt behavioral breaks create the (often correct) impression that there is no relationship between the agent's behaviors; rather than focusing on understanding the agent as a whole, the user is left to wonder how individually recognizable behaviors are related to each other and the agent's personality. Behaviors are designed in isolation and interleaved according to opportunity – but users, like it or not, attempt to interpret behaviors in sequence and in relationship to each other. The result of this mismatch between agent design and agent interpretation is confusion and frustration on the part of the user and the destruction of apparent agent intentionality.

At this point, there seems to be a basic and unsolvable mismatch between fragmentation and intentionality. But narrative psychology suggests that the fundamental problem with current agent-building techniques is not simply recognizable fragmentation in and of itself, but rather that fragmented agents do not provide proper support for narrative interpretation. From this follows the major insight of this chapter: *if humans understand intentional behavior by organizing it into narrative, then our agents will be more 'intentionally comprehensible' if they provide narrative cues*. That is to say, rather than simply presenting intelligent actions, agents should give visible cues that support users in their ongoing mission to generate narrative explanation of an agent's activity. We can do this by organizing our agents so that their behavior provides the visible markers of narrative.

Narrative agent architecture

What does it mean for agents to support narrative comprehension? The properties of narrative are complex (see Chapter 3); elsewhere I have discussed in detail how they can apply to AI (Sengers 1998) (Sengers 2000). For the sake of brevity, I will here limit discussion to the following properties:

– *context-sensitivity and negotiability:* In behavior-based systems, the 'meaning' of a behavior is thought of as always the same: the name the designer gives the internally-defined behavior. But in narrative comprehension, meaning is not a matter of identifying already-given symbols, but comes out of a complex process of negotiation between the interpreter and the events being interpreted. The meaning of the same event can change radically based on the context in which it occurs, as well as on the background, assumptions, knowledge, and perspective of the interpreter. In order to design narratively expressive agents, designers must respect (rather than attempt to override) the context- and audience-dependency of narrative comprehension.
– *intentional state entailment*: In most behavior-based systems, the reason a behavior is run is implicit in its action-selection mechanism. The behavior is then necessarily communicated to the user on a "just the facts, ma'am" basis: it is usually easy to see *what* an agent is doing, but hard to tell *why*. But in narrative, the reasons or motivations behind actions are just as important as – if not more so than – what is done. People do not want to know just the events that occur in the narrative, but also the motivations, thoughts, and feelings behind them. Supporting narrative comprehension means communicating clearly not just what the agent does, but its reason for doing it.
– *diachronicity*: Behavior-based agents jump from behavior to behavior according to what is currently optimal. Each of these behaviors is designed independently, with minimal interaction. But a fundamental property of narrative is its diachronicity; it relates events over time. In a narrative, events do not happen randomly and independently; they are connected to and affect one another. Narrative support in a behavior-based agent requires normally independent behaviors to be able to influence each other, to present a coherent picture of narrative development to the user over time.

These properties are the motivation for the Expressivator, an agent architecture that focuses on the narrative expression of agent behavior. The Expressivator is an extension of Bryan Loyall's Hap (Loyall & Bates 1991; Loyall 1997), a behavior-based language designed for believable agents. The Expressivator has been tested in The Industrial Graveyard, a virtual environment in which the Patient, a discarded lamp character implemented with the Expressivator, attempts to eke out a miserable existence while being bullied about by the Overseer, an agent implemented in Hap.

Generally, the Expressivator supports narrative comprehension using the following heuristic:

> Behaviors should be *as simple as possible*. The agent's life comes from thinking out the *connections* between behaviors and *displaying* them to the user.

Simpler behaviors are essential because *complex processing is lost on the user*. Most of the time, the user has a hard time picking up on the subtle differences in behavior which bring such pleasure to the heart of the computer programmer. But the properties of narrative interpretation mean that simpler behaviors are also *enough*. Because the user is very good at interpretation, *minimal behavioral cues suffice*.

More specifically, the Expressivator provides systematic support for narrative comprehensibility through the following mechanisms:

– *context-sensitivity and negotiability*: Rather than building an agent from conventional context- and communication-independent actions and behaviors, a designer builds agents from context-dependent *signs* and *signifiers* which are to be communicated to the user.
– *intentional state entailment*: *Transitions* are added between signifiers to explain why the agent's observed behavior is changing.
– *diachronicity*: Signifiers can use *meta-level controls* to influence one another, presenting a coherent behavioral picture over time.

Signs, signifiers, and sign management

Typically, behavior-based agents are designed for correctness, not for user comprehensibility. The first step the Expressivator takes in creating narratively understandable agents is to open the architecture up for communication. Agent design is based, not on the functions the agent must fulfill, but on its intended, context-dependent interpretation by the user. In the Expressivator, signs and signifiers support the construction of clearly communicated behavior; sign management allows the agent itself to keep track of what has been communicated, so it can tailor subsequent behavioral communication to the user's current interpretation.

Signs and signifiers

Current behavior-based approaches are based on an internal, problem-solving approach, and generally divide an agent into activities in which the agent likes to or needs to engage. Typical behavior-based systems divide an agent into three parts: (1) physical actions in which the agent engages, (2) low-level behaviors, which are the agent's simple activities, and (3) high-level behaviors, which combine low-level behaviors into high-level activities using more complex reasoning. Because these activities are implemented according to what makes sense from the agent's internal point of view, there is no necessary correlation between the agent's behaviors and the behaviors we would like the user to see in our agent.

But if the agent is to be narratively comprehensible, it may make more sense to design the agent according to the desired user interpretation, i.e. making the internal behaviors exactly those behaviors we want to communicate to the user. Then, communicating what the agent does reduces to the problem of making sure that each of these behaviors is properly communicated. For this reason, the Expressivator structures an agent not into physical actions and problem-solving behaviors, but into signs and signifiers, or units of action that are likely to be meaningful to the user. This structure involves three levels, roughly corresponding to those of generic behavior-based AI: (1) *signs*, which are small sets of physical actions that are likely to be interpreted in a particular way by the user; (2) *low-level signifiers*, which combine signs, physical actions, and mental actions to communicate particular immediate physical activities to the user; and (3) *high-level signifiers*, which combine low-level signifiers to communicate the agent's high-level activities.

There are several differences between these structural units and the default behavior-based ones. Unlike physical actions and behaviors, signs and signifiers focus on *what the user is likely to interpret*, rather than what the agent is 'actually' (i.e. internally) doing. In addition, signs and signifiers are *context-dependent*; the same physical movements may lead to different signs or signifiers, depending on the context in which the actions are interpreted. Most importantly, signs and signifiers carry an *explicit commitment* to communication; they require the agent designer to think about how the agent should be interpreted and to provide visual cues to support that interpretation.

Signs and signifiers are not simply design constructs; they also have technical manifestations. Formally, a sign is a token the system produces after having engaged in physical behavior that is likely to be interpreted in a particular way. This token consists of an arbitrary label and an optional set of arguments. The label, such as "noticed possible insult", is meaningful to the designer, and represents how the designer expects that physical behavior to be interpreted. The arguments (such as "would-be insulter is Wilma") give more information about the sign. This token is stored by the sign-management system described below, so that the agent can use it to influence its subsequent behavioral decisions. A low-level signifier is a behavior that is annotated with the special form *(with low_level_signifying...);* a high-level signifier is similarly annotated *(with high_level_signifying....).* Signifiers can also generate tokens for the sign-management system, as described below.

Sign management

Once a designer has structured an agent according to what it needs to communicate, agents can reason about what has been communicated in order to fine-tune presentation of subsequent signs and signifiers. That is, by noting which signifiers have been communicated, agents can reason about the user's likely current interpretation of their actions and use this as a basis for deciding how to communicate subsequent activity.

The most obvious way for the agent to keep track of what the user thinks is for it simply to notice which signs and signifiers are currently running. After all, signifiers represent what is being communicated to the user. But it turns out in practice that this is not correct *because the user's interpretation of signs and signifiers lags behind the agent's engagement in them*. For example, if the agent is currently running a "head-banging" signifier, the user will need to see the agent smack its head a few times before realizing that the agent is doing it.

The sign-management system deals with this problem by having the agent *post* signs and signifiers when it believes the user must have seen them. A behavior can post a sign each time it has engaged in some physical actions that express that sign, using the *post_sign* language mechanism. Similarly, once signs have been posted that express a low-level signifier, behaviors use *post_low_level* to post that that low-level signifier has been successfully expressed. Once the right low-level signifiers have been posted to express a high-level signifier, *post_high_level* is used to post that high-level signifier.

Each of these commands causes a token to be stored in the agent's memory listing the current sign, low-level signifier, or high-level signifier, respectively, along with a time stamp. Once signs and signifiers have been posted, other behaviors can check to see what has been posted recently before they decide what to do. The result is that the signs and signifiers the agent has expressed can be used just like environmental stimuli and internal drives to affect subsequent behavioral presentation, tuning the agent's behavior to the user's interpretation.

Transitions

The second requirement of narrative comprehensibility is that the user be able to tell *why* the agent is doing what it is doing. In behavior-based terms, every time an agent selects a particular behavior, it should express to the user the reason it is changing from the old behavior to the new one. This is difficult to do in most behavior-based systems because behaviors are designed and run independently; when a behavior is chosen, it has no idea who it succeeds, let alone why.

In the Expressivator, behavioral *transitions* are used to express the agent's reasoning. Transitions are special behaviors which act to 'glue' two signifying behaviors together. When a transition notices that it is time to switch between two signifiers, it takes over from the old signifier. Instead of switching abruptly to the new signifier, it takes a moment to express to the user the reason for the behavioral change.

Transitions are implemented in two parts, each of which is a full-fledged behavior: (1) *transition triggers*, that determine when it is appropriate to switch to another behavior for a particular reason, and (2) *transition demons*, that implement the transition sequence that expresses that reason to the user. Transition triggers run in the background, generally checking which behaviors are running (e.g. exploring the world), and combining this information with sensory input about current conditions (e.g. the Overseer is approaching). When its conditions are fulfilled, the transition trigger adds a special token to memory, noting the behavior which should terminate, the behavior which should replace it, and a label which represents the reason for the replacement (e.g. *afraid_of_overseer*).

Transition demons monitor memory, waiting for a transition for a particular reason to be triggered. They then choose an appropriate behavioral expression for the reason for change, according to the current likely user interpretation and conditions in the virtual environment. Expressing the reasoning behind behavioral change often requires changes to subsequent behaviors; for example, if the Patient starts doing some odious task because it is forced to by the Overseer, it should include some annoyed glances at the Overseer as part of the task-fulfilling behavior. Transitions are able to express these kinds of interbehavioral influences using the meta-level controls described below.

Meta-level controls

The third requirement of narrative comprehensibility is that behaviors should be structured into a coherent sequence. Instead of jumping around between apparently independent actions, the agent's activities should express some common threads. But these relationships between behaviors are difficult to express in most behavior-based systems because they treat individual behaviors as distinct entities which do not have access to each other. Conflicts and influences between behaviors are not handled by behaviors themselves but by underlying mechanisms within the architecture. Because the mechanisms that handle relationships between behaviors are part of the implicit architecture of the agent, they are not directly expressible to the user.

The Expressivator deals with this problem by giving behaviors *meta-level controls*, special powers to sense and influence each other. Because meta-level controls are explicitly intended for communication and coordination between behaviors, they are in some sense a violation of the behavior-based principle of minimal behavioral interaction. Nevertheless, meta-level controls are so useful for coordinating behavior that several have already found a home in behavior-based architectures. An example is Hamsterdam's meta-level commands, which allow non-active behaviors to suggest actions for the currently dominant behavior to do on the side (Blumberg 1996). In the Expressivator, behaviors can (1) *query* which other behaviors have recently happened or are currently active; (2) *delete* other behaviors; (3) *add* new behaviors, not as subbehaviors, but at the top level of the agent; (4) *add new sub-behaviors* to *other* behaviors; (5) *change the internal variables* that affect the way in which other behaviors are processed; (6) *turn off* a behavior's ability to send motor commands, and (7) *move running subbehaviors* from one behavior to another.

The most important function for these meta-level controls in the Expressivator is to allow for the implementation of transitions. Transitions, at a minimum, need to be able to find out when an old behavior needs to be terminated, delete the old behavior, engage in some action, and then start a new behavior. This means that transition behaviors need to have all the abilities of a regular behavior, and a few more: (1) they need to be able to know what other behaviors are running; (2) they need to be able to delete an old behavior; and (3) they need to be able to begin a new behavior. Ideally, they should also be able to alter the new behavior's processing to reflect how it relates to what the agent was doing before. In the Expressivator, transitions can do all these things with meta-level controls.

More generally, meta-level controls make the relationships between behaviors explicit, as much a part of the agent design as behaviors themselves. They allow behaviors to affect one another directly when necessary, rather than making interbehavioral effects subtle side-effects of the agent design. Meta-level controls give agent builders more power to expose the inner workings of agents by letting them access and then express aspects of behavior processing that other systems leave implicit.

Putting it all together

Narrative psychology suggests that narrative comprehension is context-sensitive, focuses on agent motivation, and seeks connections between events over time. The Expressivator supports comprehensibility by expressing the agent's actions with signs and signifiers, the reasons for agent activity with transitions, and the coherent threads through activities with meta-level controls.

These architectural mechanisms are described separately, but used together in the agent design process, changing the way in which agents are designed. In a typical behavior-based system, an agent is defined in 3 major steps: (1) deciding on the high-level behaviors in which the agent will engage; (2) implementing each high-level behavior, generally in terms of a number of low-level behaviors and some miscellaneous behavior to knit them together; (3) using environmental triggers, conflicts, and other design strategies to know when each behavior is appropriate for the creature to engage in. With the Expressivator, the choice and expression of these structural 'units' for the agent is not enough; in order to support the user's comprehension, the designer must also give careful consideration to expressing the reasons for and connections between those units. These connections are designed and implemented with transitions, which alter the signifiers they connect into a narrative sequence. In practice, transitions are the keystone of the architecture, combining signifiers in meaningful ways through the use of meta-level controls.

Results

The best way to see how the Expressivator changes the quality of agent behavior is to look at how its transitions work in detail. Here, I will go over one point where the agent switches behaviors, and explain how transitions make this switch more narratively comprehensible. One example does not proof

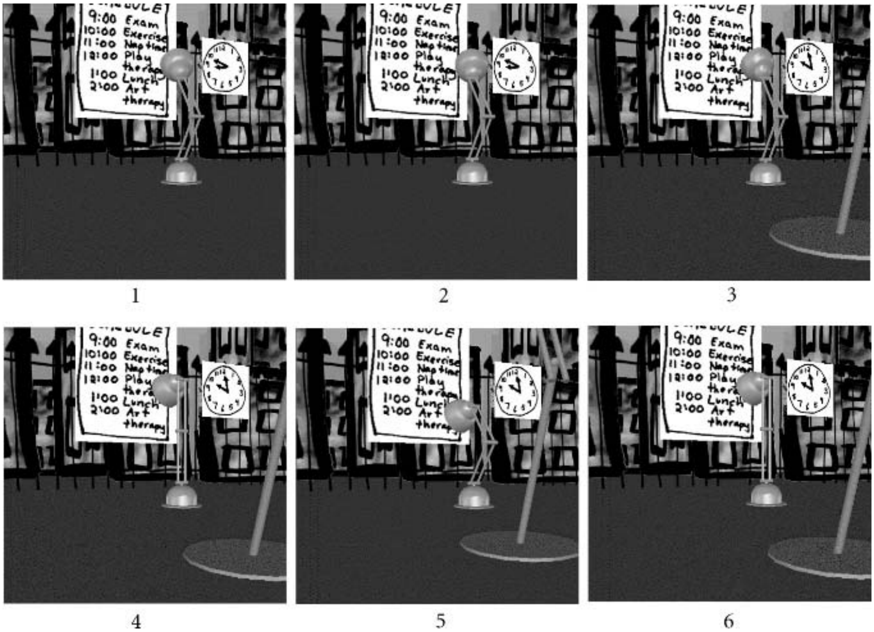

Figure 1. Response without transitions.

make, but it does take up a lot of space; the sceptical reader can find more in (Sengers 1998).

As our excerpt begins, the Patient notices the schedule of daily activities which is posted on the fence, and goes over to read the schedule. The Overseer, noticing that the Patient is at the schedule and that the user is watching the Patient, goes over to the schedule, changes the time to 10:00, and forces the Patient to engage in the activity for that hour: exercising.

The goal of this part of the plot is to communicate to the user the daily regime into which the Patient is strapped. The Patient does not have autonomy over its actions; it can be forced by the Overseer to engage in activities completely independently of its desires. The specific behavioral change from reading the schedule to exercising, then, should show the user that the agent changes its activity because (1) it notices the Overseer, (2) the Overseer enforces the scheduled activities; (3) the activity that is currently scheduled is exercising.

Without transitions, the Patient's response to the Overseer is basically stimulus-response (Figure 1). The Patient starts out reading the schedule. As soon as the Patient senses the Overseer, it immediately starts exercising. This reaction is both correct and instantaneous; the Patient is doing an excellent job of problem-solving and rapidly selecting optimal behavior.

But this behavioral sequence is also perplexing; the chain of logic that connects the Overseer's presence and the various environmental props to the Patient's actions is not displayed to the user, being jumped over in the instantaneous change from one behavior to another.

With transitions, attempts are made to make the reasons behind the behavioral change clearer (Figure 2). Again, the behavior starts with the Patient reading the schedule. This time, when the Overseer approaches, the Patient just glances at the Overseer and returns to reading. Since the Patient normally has a strongly fearful reaction to the Overseer (and by this time the Overseer's enthusiasm for punishing the Patient has already generally aroused sympathy in the user's mind), the user has a good chance of understanding that this simple glance without further reaction means that the Patient has not really processed that the Overseer is standing behind it.

Suddenly, the Patient becomes startled and quickly looks back at the Overseer again. Now, the user can get the impression that the Patient has registered the Overseer's presence. Whatever happens next must be a reaction to that presence. Next, the Patient checks the time and the schedule of activities to determine that it is time to exercise. Then the Patient whirls to face the Overseer and frantically and energetically begins exercising, tapering off in enthusiasm as the Overseer departs. This transition narrativizes the agent's behavior in the following ways:

- the agent design is predicated on the user's context-dependent interpretation, e.g. that the user will interpret the agent's short glance at the Overseer differently now than earlier in the story;
- the transition communicates that the change in behavior is connected to several factors: the presence of the Overseer, the clock, and the schedule. This is in contrast with the transition-less sequence, in which there is no clear connection between any of the environmental factors and the Patient's behavioral change;
- the subsequent exercising behavior is altered to fit into a narrative sequence by making it more frantic in response to the agent's panic during the transition.

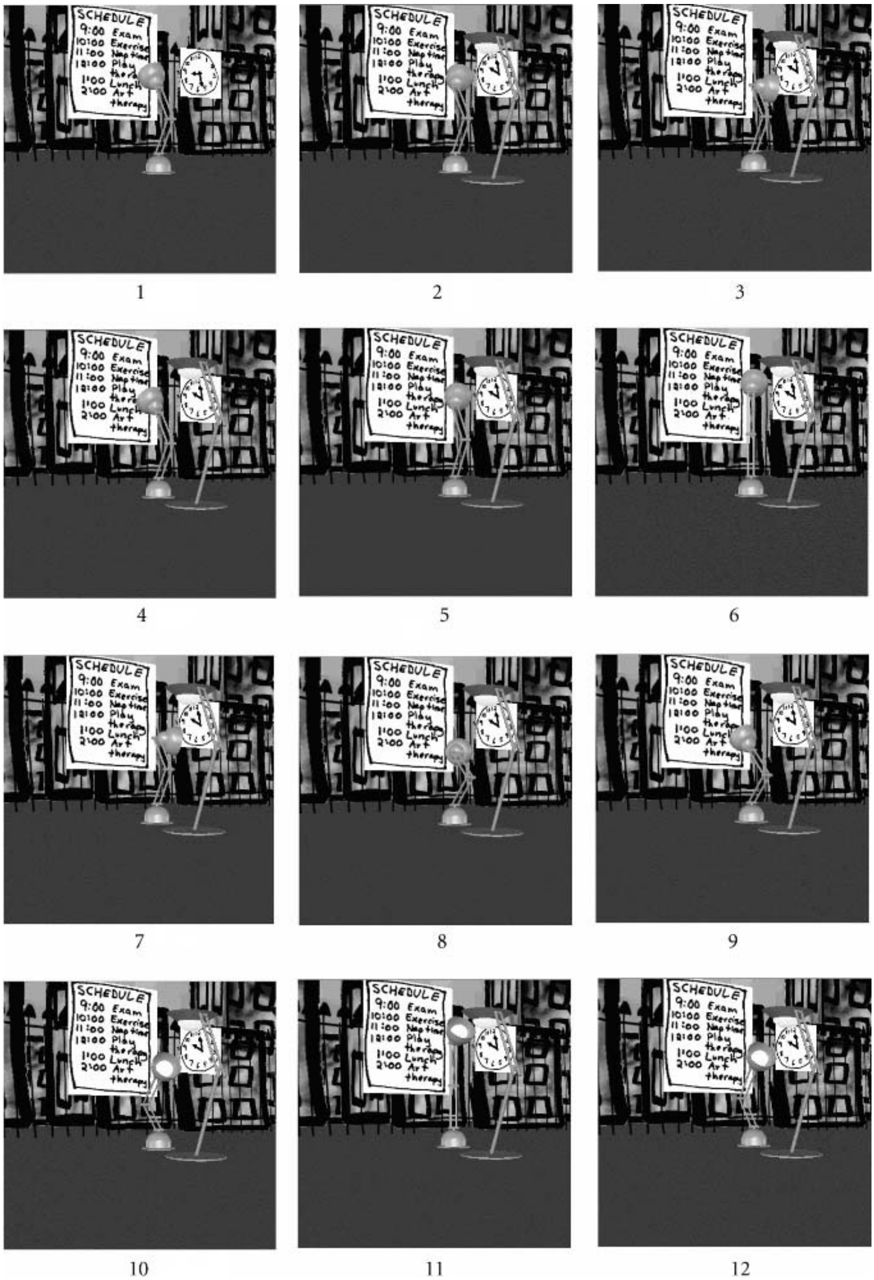

Figure 2. Response with transition.

Conclusion

In this paper, I have argued that there is a fundamental lack in autonomous agents' behavior, which reduces their apparent intentionality. By being constructed in a fragmented manner, agents suffer a kind of schizophrenia, a schizophrenia which can be addressed, in analogy to anti-psychiatry, by making agents narratively understandable. In order to do this, I have built an agent architecture which combines (1) redefinition of behaviors as signifiers and their reorganization in terms of audience interpretation, (2) the use of transitions to explain agent motivation, structuring user-recognized behaviors into narrative sequences, and (3) the use of meta-level controls to strategically undermine fragmentation of the agent's behaviors. Preliminary results are encouraging, but further work, preferably involving the development of support for graphical presentation, will be necessary in order to fully evaluate the implications of and possibilities for the architecture.

More generally, if black-box behaviorism involves thinking of human life mechanically, reducing it to a matter of cause-effect, while narrative allows for the full elucidation of meaningful intentional existence, then it seems likely that narrative – and by extension the humanities, for whom narrative is a modus operandi – can address meaningful human life in a way that an atomizing science simply cannot. If humans comprehend intentional behavior by structuring it into narrative, then AI must respect and address that way of knowing in order to create artifacts that stimulate interpretation as meaningful, living beings. This suggests that the schizophrenia we see in autonomous agents is the symptomatology of an overzealous commitment to mechanistic explanation in AI, a commitment which is not necessarily unhelpful (since it forms the foundation for building mechanical artifacts), but needs to be balanced by an equal commitment to narrative as the wellspring of intentionality.

References

Baur, Susan (1991). *The dinosaur man: Tales of madness and enchantment from the back ward.* New York: Edward Burlingame Books.

Blumberg, Bruce (1996). *Old tricks, new dogs: Ethology and interactive creatures*. PhD thesis, MIT Media Lab, Cambridge, MA.

Brooks, Rodney (1990). Elephants don't play chess. In Pattie Maes (Ed.), *Designing autonomous agents*. Cambridge, MA: MIT Press.

Brooks, Rodney (1997). From earwigs to humans. *Robotics and Autonomous Systems, 20* (2–4), 291–304.

Bruner, Jerome (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Bruner, Jerome (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Dennett, Daniel (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Goffman, Erving (1961). *Asylums: Essays on the social situation of mental patients and other inmates*. Garden City, NY: Anchor Books.

Janet, Pierre (1889). Félix Alcan (Ed.) *L'automatisme psychologique: Assai de psychologie expérimentale sur les formes inférieures de l'activité humaine*. Paris: Ancienne Librairie Germer Baillière et Cie.

Laing, R.D. (1960). *The divided self: An existential study in sanity and madness*. Middlesex: Penguin Books.

Laing, R.D. & A. Esterson. (1970). *Sanity, madness, and the family*. Middlesex: Penguin Books.

Loyall, A. Bryan (1997). *Believable agents: Building interactive personalities*. PhD thesis, Carnegie Mellon University, Pittsburgh. Technical Report CMU-CS-97-123.

Loyall, A. Bryan & Joseph Bates (1991). *Hap: A reactive, adaptive architecture for agents*. Technical Report CMU-CS-91-147. Carnegie Mellon University.

Robear, James Walter Jr. (1991). Reality check. In J. G. Oakes (Ed.), *In the realms of the unreal: "Insane" writings* (pp. 18–19). New York: Four Walls Eight Windows.

Ronell, Avital (1989). *The telephone book: Technology – schizophrenia – electric speech*. Lincoln: University of Nebraska Press.

Sengers, Phoebe (1998). *Anti-boxology: Agent design in cultural context*. Ph.D. thesis, Carnegie Mellon University Department of Computer Science and Program in Literary and Cultural Theory, Pittsburgh, PA.

Sengers, Phoebe (2000). Narrative intelligence. In Kerstin Dautenhahn (Ed.), *Human cognition and social agent technology*. Amsterdam: John Benjamins.

Steels, Luc (1994). The Artificial Life roots of Artificial Intelligence. *Artificial Life, 1* (1–2), 75–110.

Part IV

Analyzing the Stories We Tell

Chapter 17

Writing and representation

Philip E. Agre
University of California, Los Angeles, CA

Introduction

The notion of representation obviously labors under a long philosophical history (Judovitz 1988; Rorty 1979; Silvers 1989), not to mention the history of art (Hagen 1986; Wallis 1984), literature (Auerbach 1953; Brodsky 1987; Krieger 1987), and historiography (Hartog 1988; White 1973). These days, though, it also labors under an appreciable technical history, handed down through a practice of building computer systems that construct, maintain, and manipulate "representations" (Brachman & Levesque 1985; Haugeland 1981). And the philosophical and computational issues interact. I often find that philosophy helps to interpret the difficulties that arise in my technical practice. And I want to believe that technical practice can help philosophy. In writing the stories that follow, I have explored some places where technical questions align with philosophical answers. I don't yet know how to convert these answers back into technical practice.

I disagree with two widespread ideas about representation, namely "semantics" and "world models". The main tradition of semantics holds – presupposes – that a representation has a "meaning" or a "content" independent of the identity, location, attitudes, or activities of any particular agent (but cf. Barwise & Perry 1983). This meaning or content is often understood as a systematic, objective relationship between the representation itself and states of affairs obtaining in the world. A world model is a component of a physically realized computational system, an object whose internal structure stands in a systematic, objective, analogical relationship to states of affairs presumed to obtain in the world. Some computational process maintains the model as the world changes; reasoning about the world involves inspecting and manipulating the model. I will argue that both semantics and world models overlook central features of representations and their use.

On the positive side, I offer two suggestions about how people use symbolic representations. The first is that *people interpret symbolic representations in making sense of particular situations*. Interpretation is a situated activity. (Whatever the form of the representation, whether written or spoken or displayed on a video monitor, I shall speak of its configuration of symbols as a "text". The "representation" is the actual material object: the sheet of paper, the speech signal, or the video image.)

My second suggestion is that *what a given text is talking about is a fresh problem in every next setting* (Amerine & Bilmes 1988; Zimmerman 1970). The work of relating a text to a concrete setting – looking around, poking into things, trying out alternative interpretations, watching someone else, getting help – will generally be both "mental" and "physical", though it is best not to distinguish. Relating a text to a concrete setting takes work because the text might be relevant to the situation in a great variety of ways. The text has a great deal of "play", so that much of one's interpretive effort must wait until the time comes. This is the opposite of extracting a "meaning" from a text as soon as it arrives. The point is not that interpretation is wholly unconstrained by the text; rather, interpretation is constrained jointly by the text and by the circumstances in which it is interpreted. The only way to explain the point is through examples.

1. I'm trying to find a friend's house in a heavily wooded mountainous rural area. I had received directions to the house by e-mail and printed them out so I could carry them in the car. I knew the main road, but I knew nothing about the residential streets leaving it. The last two paragraphs read:

> About a mile up from the intersection, look on the left for Elk Tree Road – it's a dirt road with a little bus stop at the end. Follow Elk Tree *past* the first left (Elk Tree WAY) to the mailboxes and take the middle of the three-way fork on the right, Upper Elk Tree ROAD.
>
> My place is the first one on the left, #27. Park on the left shoulder near my red Honda, and come down the steps and up the stairs to my front deck. (If you went to the main door, you'd get my landlords, not me.)

I had the sense to check my odometer at the intersection so I'd know when "about a mile up" was coming. Even so I somehow missed Elk Tree Road the first time. The bus stop is obvious enough if you know where to look for it, but it's a little way up the dirt road and obscured by foliage. The most difficult part, though, concerned the “three-way fork on the right”. When I got to the mailboxes (dozens of them, in fact) I had to find some way to interpret the scenery as “the three-way fork on the right”. Unfortunately, I could only count two roads on the right. After much backing up and looking around, I decided that a large driveway roughly straight ahead was the third road. Setting off along the “middle” road, I looked for the first place on the left. The road snaked down a hillside with many houses on the right. Finally I came to a house on the left. I couldn’t find any street numbers (no surprise out here) but there was definitely a red Honda parked outside. I parked, got out, looked for the place that “come down the steps and up the stairs to my front deck” was talking about. I found such a place but it led me to the side door of a cluttered garage. A great deal of searching and asking around followed. Even though something was clearly wrong, that red Honda made me reticent to abandon any of my previous interpretations.

To make a long story short, the middle branch of the three-way fork immediately branched into Lower Elk Tree Road steeply down to the left and Upper Elk Tree Road steeply up on the right. A large hand-painted sign for Upper Elk Tree Road clearly marks the split, but I was already looking for the first place on the left. Writing this now, I realize that it makes perfect sense for *Lower* Elk Tree Road to branch steeply down and *Upper* Elk Tree Road to branch steeply up, but at the time I didn’t give the right branch any thought at all since I simply did not see the branch *as* a branch. If you’re standing there looking at the branch *as* a branch then it’s hard to imagine how it looked to me, but it never occurred to me that I had another choice of roads to make. The sign must have occupied a reasonable portion of my visual field, but I was already looking for a house, not a sign, and to the left, not the right. In short, I saw what I assumed the instructions were telling me to see. Told about a “three-way fork to the right” and unable immediately to see such a thing in that place, I *worked*, successfully, to see it. Told about “the first [place] on the left”, I unquestioningly saw what the instructions implicitly told me was there, namely the road on which my friend’s house was located.

2. A friend recently taught me to fold origami paper cranes. In walking me through the various steps, she often had to explain by some combination of words, pointing, and demonstrations where and how to fold next. The intermediate forms don’t look much like cranes, and the paper keeps taking on unexpected new identities as you fold it, even after you’ve gotten reasonably proficient. These forms can be hard to talk about because they’re importantly asymmetrical in nonobvious ways. Even demonstrations are only of limited use if you can't see the asymmetries for yourself.

In the course of her explanations my friend said things like "put your finger in the pocket", "fold it back to make a boat", and "make the legs skinnier". Making each of these metaphors refer to parts and aspects of the folded paper always took considerable effort, even though it was always wholly evident in retrospect. Much of my friend's job was to get me to look at my partly-folded origami crane in the right way, so that certain parts and aspects would stand out as units for me. She was teaching me the skill of seeing my paper *as* having a pocket, a boat, or legs. Although I got better at this skill, it never stopped taking work. The work only became more routine.

3. The final example comes from my experience teaching people to program computers. If you're comfortable in front of a computer terminal, it's easy to teach the wrong things. You've got all kinds of theory, but theory doesn't help someone who hasn't yet gotten the idea of being "in" the editor. So I sit the student at the keyboard and tell them very concretely what to look at and what to type. As they get comfortable, my instructions grow more abstract. For example, I might say "Type open paren then ...", only to see them type the letters "o p e n" and "p a r e n". In that case I have to point out where the parenthesis keys are. Later, though, I can say things like, "Let's define a function called ...". When they're learning to read code, I have to point out that there are conventions about indentation that result in common types of code having characteristic shapes. And I don't explain abstraction hierarchies until I explain that two hunks of code that look alike are often good candidates for a common abstraction.

As these examples illustrate, my prototype of representation is natural language, whether as spoken utterances, written texts, or mental thoughts. In each case, figuring out what in the situation the text was talking about took work: creative improvisation, reference to artifacts, and interactions with others. The work consisted in relating natural language to concrete situations: identifying the things the words were mentioning, seeing materials under metaphorical descriptions, and heuristically associating visual patterns with verbalized technical abstractions. And the work required to make sense of "take the middle of the three-way fork" or "put your finger in the pocket" or "define a function" might differ greatly in different settings or under different conditions.

I want to concentrate on representations written on paper. The image of standing in a kitchen or on a street with a written text in your hand is a good reminder that relating texts to circumstances requires work and that this work requires understanding in some measure what you're doing. You have to understand what you're doing since the text certainly doesn't (cf. Searle 1981). Obviously representations often influence your actions, but you don't understand what you're doing in virtue of owning them. This point is supposed to apply equally to all forms of representation, not just writing. The idea that understanding does not reside in representations is difficult and consequential. The paper's later stories will explore this idea in the context of "internal language".

Five fairly independent notes follow.

The homunculus and the orbiculus is an assault on the notion of a world model, placing it in a philosophical tradition of trying to explain the human ability to act competently in the world by pretending that the relationship between person and world is reproduced inside the person's head. This sort of explanation is seductive because it plays to the principal strength of current computational technology: building abstractions inside of computers that are almost entirely cut off from the outside world. But it doesn't work well in practice.

Writing as bad and good metaphor for representation contrasts two ways in which written texts might be regarded as prototypes of representation. The first, "bad" way focuses however tacitly on certain physical properties of written texts. This bad understanding of writing is central to the main tradition of computing. The second, "good", way concentrates on more abstract aspects of written texts: both the text and your surroundings are outside of you, and so it takes work to see the world as being what the text is talking about.

A story about photocopier supplies also concerns written instructions. A secretary is justifiably annoyed because somebody has put laser-printer dry toner in the photocopier. What happened and why? At issue is the selective use people must make of the representational materials that surround them. I defend the culprit, arguing that the phrase "dry toner" on a bottle is not "ambiguous" in a way that anybody could be expected to notice, even if evidence serving to identify and resolve the ambiguity is readily available. One would like photocopier users to be "careful", but it's hard to formulate the demand in an actionable way, given that such a problem could hide behind any of the vast number of unarticulated assumptions that form the background of any such activity.

A story about some instructions at a performance in an art gallery is another story about instructions, this time a single sentence that one person hollered to a group of others. These instructions did not function well because the recipients could not see, or even imagine, what in the setting the instructions could be talking about.

A story about my routines for reading the Sunday Globe is a story about an instruction I issued to myself in the course of reading the newspaper one Sunday morning. The pattern of activity in which the instruction participated had long since become routine. Nonetheless, the way in which the instruction went wrong reveals that it was something like an natural language imperative and not, for example, a computer program. In this case, a change in the environment led me to make a different sense of an ambiguous phrase than I used to.

The stories are not supposed to prove any general propositions. Instead, they invite you to be aware of similar phenomena in your own experience of everyday representation-use. Computational investigations and awareness of everyday life can influence one another. Parallel pursuit of these two kinds of inquiry will, I believe, lead to deeper understandings of why our life is the way it is and why machines can take certain forms and not others.

The homunculus and the orbiculus

In the old days, philosophers accused one another of believing in someone called a *homunculus* – from Latin, roughly "little person". For example, one philosopher's account of perception might involve the mental construction of an entity that "resembled" the thing-perceived. Another philosopher would object that this entity did nothing to explain perception since it required a mental person, the homunculus, to look at it. Computational ideas appeal to these philosophers because they can imagine "discharging" the homunculus by, for example, decomposing it into a hierarchy of ever-dumber subsystems (Dennett 1978: 124).

But the argument about homunculi distracts from a deeper issue. If the homunculus repeats in miniature certain acts of its host, where does it conduct these acts? The little person lives in a little world – the host's surroundings reconstructed in his or her head. This little world deserves a Latin word of its own. Let us call it the *orbiculus*. One way to say "world" is *orbis terrarum*, roughly "earthly sphere". But *orbis*, I am told, extends metaphorically in the same ways as "world" in English: one might speak of the world of a peasant or a movie director, meaning roughly their existential world, "the world they live in" (more literally, their sphere). So the orbiculus is your world copied into your head.

AI is full of orbiculi. A "world model" is precisely an orbiculus; it's a model of the world inside your head. Or consider the slogan of vision as "inverse optics": visual processing takes a retinal image and reconstructs the world that produced it (Hurlbert & Poggio 1988). You'll also find an orbiculus almost anywhere you see an AI person talk about "reasoning about X". This X might be solid objects, time-extended processes, problem-solving situations, communicative interactions, or any of a hundred other things. "Reasoning about" X suggests a purely internal cognitive process, as opposed to more active phrases like "using" or "participating in" X. AI research on "reasoning about X" requires representations of X. These representations need to encode all the salient details of X so that computational processes can efficiently recover and manipulate them. In practice, the algorithms performing these abstract manipulations tend to require a choice between restrictive assumptions and computational intractability (see Brachman & Levesque 1984; Hopcroft & Krafft 1987).

If you prefer the phrase "using X" over "reasoning about X", an AI person will ask you: "But we can reason about things that aren't right in front of us, can't we?" AI's version of mentalism offers seductive answers to many questions, and this is one of them. According to mentalism, reasoning about a derailleur proceeds in the same way regardless of whether the derailleur is in front of you or across town. Regardless of where the derailleur is located, you reason about it by building and consulting an orbicular derailleur-model. If having the derailleur present helps you, it is only by helping you build your model.

This is, of course, contrary to common experience. As we all know, the first several times you try to reason about a derailleur (1) it has to be sitting right in front of you and (2) you have to be able to look around it, poke at it, and take it apart (Chapman & Agre 1986). I've disassembled and reassembled several derailleurs. Yet without a derailleur in front of me, or at least a good diagram, I cannot explain how a derailleur changes gears, or even list the parts involved. Experts can, but not an amateur like me. Why aren't several disassemblies and reassemblies of derailleurs enough to build a mental model of them?

Maybe the computational complexity of reasoning with realistic world models is trying to tell us something. Maybe what you learn when you gain experience with a derailleur or a city or a recipe is more specific. Perhaps it is more biased to the specific things you've had to remember in the course of the activity. Perhaps it is more closely tied to your goals at particular moments of the activity. Perhaps it is more organized around the experience of the individual situations that arise in the course of the activity. These are difficult ideas. The question is complicated and messy and poorly worked out. But that's to be expected. Expecting it to be easy is a sign of addiction to the easy answers of the orbiculi.

Writing as bad and good metaphor for representation

Within the technologically informed human sciences, cognition is almost universally understood to involve the mental manipulation of assemblages of symbols called representations. These representations represent the individual's world – they are the orbiculus. The vast majority of this research assumes symbolic representations to have certain properties. They are:

– object-like (neither events nor processes)
– passive (not possessing any sort of agency themselves)
– static (not apt to undergo any reconfiguration, decay, or effacement, except through an outside process or a destructive act of some agent)
– structured (composed of discrete, indivisible elements whose arrangement is significant in some fashion)
– visible (can be inspected without thereby being modified), and
– portable (capable of being transported to anyone or anything that might use them without thereby being altered or degraded).

Although the cognitivist understands symbolic representations as abstract mental entities, all of these properties are shared by written texts (Latour 1986). Words like "structured", "inspected", "modified", "transported", and "altered" are metaphors that liken abstractions inside of computers to physical materials such as paper. Observe also that most of these properties are deficient or absent for spoken utterances, which evaporate as quickly as they are issued (Derrida 1976: 20) and are only decomposed into discrete elements through complex effort. Thus we can speak of a writing metaphor for representation.

This conception of representation-as-writing is topical for several reasons. The connectionist movement has lent urgency to the seeming conflict between symbolic manipulation and the relatively simple, uniform, statically and locally connected, highly parallel hardware of the human brain (Fodor & Pylyshyn 1988; Hutchins 1986; Rumelhart et al. 1986). Anthropologists such as Goody (1986), Ong (1982), Harris (1980, 1987), and Latour (1986) have challenged views of cognition that make universal principles out of psychological and social phenomena found only in literate cultures.

This section has three purposes. First I argue that symbolic representation in artificial intelligence is, historically, modeled on written texts, as opposed to (say) photographs or spoken utterances. Then I describe how writing is a bad metaphor for symbolic representation. These arguments implicate prevalent technical methods. Finally I describe how writing is a *good* metaphor for symbolic representation. These arguments suggest new technical directions.

Representation as writing

Roy Harris (1987), among others, has argued that ideas about representation in philosophy and linguistics have been biased by writing. He observes that these fields have emphasized those aspects of human utterances that appear in a conventional written representation. One might read in a textbook a sentence such as, "Suppose that John says to Mary, 'Please close the window'." and this sentence will be taken to specify some hypothetical event. We do not normally wonder, and only rarely are we told, about several aspects of John's action:

– his tone of voice
– his articulation of the various phonemes
– the shape of his intonation
– the timing of the various elements within the utterance
– the timing of his utterance relative to other actions and events
– whether he and Mary have a history of interactions over this window
– his position relative to Mary and the window
– his posture
– his gestures
– his facial expression
– the direction of his gaze
– whether and when he has caught Mary's gaze

(For the horrors of trying to make written notations of these things, see Atkinson and Heritage (1984) or Levinson (1983) for an introduction to Jefferson's notation system used in conversation analysis.)

Given that these aspects of speech regularly affect the import of utterances, a written sentence must be considered a poor representation of a spoken utterance. But most philosophical and linguistic analyses have proceeded on the basis of this idealized representation, a tradition that AI has carried on. The point is not that these fields talk about writing; only that they concentrate on the aspects of representation that writing normally captures. As a result, theories will naturally tend to lean on distinctions that writing captures, and not on the many distinctions it doesn't.

Among the many routes by which the writing metaphor entered AI practice, one moment stands out: Newell and Simon's (1963, 1972) invention of symbolic programming. Most of Newell and Simon's domains, especially in their earlier work, have been domains like cryptarithmetic in which the "world" consists of a sheet of scratch paper. Newell's production system models do not contain separate mechanisms for the scratch paper and for the agent's "short-term memory". Newell and Simon invented symbolic programming in order to implement the sorts of structures and operations that their models specified. List structures, like scratch paper, and like the symbolic structures of all subsequent AI programming languages, have many properties of writing and few properties of speech. People invented writing because there's nothing in their heads that's anything like paper.

Writing in the head

AI research is often caught in a pattern whereby mechanisms that seem extremely "expressive", "powerful", and "general" refuse to scale up. Let's return to the properties that AI has ascribed to symbolic representations – object-like, passive, static, structured, visible, and portable – and consider how they lead to difficulties of scaling and implementation.

Symbolic programming languages endow their datastructures with all six properties of writing. They implement these properties using pointers that metaphorically make objects visible to processes. Pointers connect the components of structures. One transports a structure by "passing" a pointer to it. Structures only change when processes change them. Pointers do not obey any locality beyond that of their own connectivity. Thus they are eminently reconfigurable.

Pointers cause two sorts of difficulties. First, they require their implementation medium to be infinitely reconfigurable (Chapman 1991: 35–41). They fight both against the locality of physical space and against the inertia of physical machinery and its interconnections. On serial machines we observe this difficulty in the complexities of dynamic storage management. On parallel machines we observe it in the complexities of shared-memory management. Pointers also cause algorithmic complexity. Just as you can write symbols on paper in any order, a pointer can point at anything. As a result, algorithms for the manipulation of symbolic pointer-structures often suffer from the combinatorial arbitrariness of the objects they reason with.

If writing is a good metaphor for symbolic representation, then, it is not because we have things in our heads that are object-like, passive, static, structured, visible, and portable. These properties of writing don't help us to understand human use of symbolic representation *in general* because they are the properties of written texts that are *specific to written texts*. Far from picking out the essence of symbolic representation, they dwell on the physical activity of using a written text: inscribing, gathering, comparing, storing, and destroying. The invention of writing was important precisely because it permitted these useful forms of activity.

Writing as representation

How, then, can writing serve as a model of symbolic representation? Imagine that you're using a recipe – that is, a recipe on paper – to help you cook dinner. Or perhaps you're using some directions – again, on paper – to help you get to a party. The paper has a paradoxical position. Even though it's a physical object with a definite size, mass, and location, it plays its role – at least qua representation – entirely through your interpretation. And even though it seems to offer opinions about the particulars of the situation, it only does so because you figure out what in your surroundings it's talking about.

The paper in your hand is both part of the material situation and doubly removed from it. It underdetermines the sense you make of it because it is separate from you – after all, someone else in the same situation would probably do something different. And it underdetermines what in the situation it picks out because you are separate from your surroundings – after all, it would probably be useful in other situations as well. A similar argument then applies, not just to notes written on paper, but to all symbolic representations: their meaning must be completed in the act of use (Ingarden 1973). The world does not come innately parceled out into the categories we find mentioned in written texts. Instead, people use representations to help them make sense of particular situations. What a given text is talking about is a fresh problem in every next setting. (For those who care about such things, this is what Jacques Derrida means by the word "writing". For introductions to Derrida's philosophy see Culler (1982) and Norris (1982). I have also been influenced by Garfinkel's (1984 [1967]) ethnomethodological ideas about the indexicality of representation use. For an introduction see Heritage (1984).)

This idea has many consequences for computation. The relationship between internal processes and internal symbolic representations is qualitatively the same as the relationship between a person and a sheet of paper. The idea of cognition as a process operating on symbolic representations is therefore of less help than we had hoped. We can put symbolic representations inside our robots, but the hard problems remain.

What other forms might a theory of symbolic representation take? Vygotsky (1978 [1934]; for an introduction see Wertsch 1985) suggests that cognitive skills involving mental representation arise from the everyday use of physical representations, many of which are embedded in patterned social relationships. These skills are diverse because the everyday forms of representation use are diverse: speaking and hearing give rise to internal speech, making and looking at pictures give rise to visualization, and so forth. And the "internal" and "external" uses of representations tend to blur together in practice. But the internal processes differ from the external processes because the insides of our heads aren't like the outsides. Instead, they are shaped by the kind of machinery we have in our heads. For example, internalized speech is far more consequential than internalized writing because our brains are better suited for reproducing speech than writing. If this is true, then many properties of speech that philosophy and linguistics have marginalized, such as tone of voice and intonation (and their pragmatic import in particular contexts of use), will have to be readmitted to our theoretical center stage.

A story about photocopier supplies

Back when I worked at counter jobs, I discovered that people are oblivious to signs. You could put a big red sign with stars and arrows

> The machine will be back up at midnight

anywhere you liked and people would still walk up to the counter and ask "when will the machine be back up?".

Here is an example of this effect.

> Date: Tue, 17 May 1988 17:05 EDT
> From: J...

To: All-AI
Subject: Xerox copiers and Lazer Printers

Due to someone's IGNORANCE, CARELESSNESS, or LACK OF PATIENCE, SOMEONE PUT DRY IMAGER FOR THE LAZER MACHINE INTO THE XEROX MACHINE on the 8th floor. These supplies, although both dry imager, ARE NOT INTERCHANGEABLE!!! It says on the box it comes in (and on the bottle itself) which machine it is for.

We were warned by the Xerox people before that if this happened again, they may discontinue servicing our machines – not to mention the cost of having it corrected (or maybe having to get a replacement).

If there is a problem with the xerox or lazer machine on the 7th floor and you do not know how to correct it or have a question, see D... She is in charge of the overall care of those machines as I am the machines on the 8th floor (I can be located at ...).

Because of one person's lack of resourcefulness (he/she could of went down to the 7th floor or seen me or D), we are all suffering! If it is after hours and you are not sure what to do, it is better to do nothing than ruin a machine.

–J

In the old days, when the photocopier needed dry toner, one looked around, saw the bottle marked "dry toner", opened the machine, found the reservoir marked "dry toner", and put the contents of the bottle into the reservoir. Sometimes someone would put the toner someplace else, or they'd put something else into the toner reservoir, but these things didn't happen often.

By the late 1980s, however, many rooms with photocopiers in them also had laser printers in them. (Laser printers did not yet have disposable print cartridges with their own toner supplies.) And in the copier/printer room on the 8th floor of the MIT AI Lab, the two types of dry toner were incompatible. So what happened when the copier ran out of dry toner? Exactly the same thing as before. Except that the bottle marked "dry toner" one happens to come across first might or might not be the correct dry toner.

Does this happen because people are too lazy to check whether it's laser or copier toner? After all, as J points out, "It says on the box it comes in (and on the bottle itself) which machine it is for". Do the people just go irresponsibly ahead putting laser toner in the copier figuring it's probably OK? Do they proceed despite a conscious uncertainty?

Although we must appreciate J's situation, I think all these hypotheses are unnecessary. Put yourself in the place of someone to whom the photocopier is asking for dry toner, and suppose that you had not yet known that the copier and printer employed two incompatible types of dry toner that could be confused for one another. You made dozens of separate moves in answering the copier's call to be resupplied with toner, and any of those moves could be mistaken in dozens of different ways. The actual problem, namely that the bottle marked "dry toner" was not actually the correct substance, is pretty obscure, as if somebody parked a car nearly identical to yours a couple spaces down. In the case of the wrong car, some discrepancy would probably force itself upon your attention before you got too far. Yes, evidence of the mistake was readily available, but were you really supposed to list all the things you might be doing wrong and go looking for evidence to rule out each one? That would be impossible.

The problem, in short, is only obvious in retrospect. Having been warned that toner comes in two types, one should probably start to check. But nobody is born with that knowledge. Someone could become perfectly proficient at replacing toner in copiers and still run afoul of this difficulty, simply because types of toner had never become an issue for them. The arrival of the laser printer would invalidate one of the innumerable implicit background assumptions of their toner-changing routine, but it is hard to articulate the general policy that could have informed them of this. Read the fine print on every label every time? But the world is full of representations; how do you decide when to stop reading them all and start doing something? The phrase "dry toner", to us, having been informed of the problem, is ambiguous: it could mean "laser dry toner" or "copier dry toner". But that ambiguity is only consciously and morally an ambiguity for *us*, the well-informed. And for all we know, "dry toner" could also be ambiguous in an unlimited variety of other ways.

In the end, all we can do is stop moralizing and fix the problem. Make the toner bottles so different that it's physically impossible to install the wrong stuff. Train everyone. Or lock up the toner.

A story about some instructions at a performance in an art gallery

One day I went to the MIT Media Lab to see a performance. The performance took place in a windowless room that's about thirty feet square with a high ceiling. When we arrived the doors hadn't yet opened, so the audience milled about outside. Finally the time came and the ticket-seller wandered into the lobby and yelled something like, "OK the doors are open". Then as people were drifting toward the door she moved into the middle of the crowd and yelled a complicated set of instructions that went something like this:

> You can sit on the chairs or you can go to the back wall and look through the windows or you can go up on the balcony, but don't lean against the side wall.

None of us could see the inside of the performance space as she was saying all this, so we had no way of knowing what she meant by windows, side wall, back wall, balcony, etc. This room has no windows or balconies, and so the yeller must have been referring to structures that were built specially for the performance. One could feel the crowd being uncomfortable, many of them turning to their neighbors in an attempt to get clarification. I found myself trying to visualize the scene, but I had no idea how to place even the "side wall", much less the balcony and the windows. Both the impossibility of visualizing the scene and the effort spent trying to visualize it seemed to make the instructions unusually hard to remember, as if they were nonsense syllables, and several people could be heard repeating parts of them over to themselves or to their neighbors.

I found this amusing, and adopting a gently ironic imitation of the register and diction of the yeller I said something like, "you can stand between the monsters but don't sit on the toadstools". Not many people got the joke, I'm afraid, and especially not the yeller. I found this interesting in itself. The yeller was obviously familiar with the room, and she presumably had no problem visualizing what she was talking about. She evinced no awareness that others might be having a problem.

Once inside, there was an audible rush to attach the words to parts of the room, which was dark and full of peculiar wooden structures. The "side wall" is immediately there on your right, verifiable by the readily visible chairs along it, and the "balcony" could be found along the back wall with a little scanning. The "windows" weren't at all obvious; they were windows in the wall supporting the balcony, behind which one could stand. I suggested we go for the balcony; although no stairs were immediately visible it was obvious where they should be and others were already headed that way ahead of us.

In this story, the peculiar relationship between the instructions and the physical setting disrupted an aspect of language understanding that normally goes unremarked. The situation resembles the mnemonist's method of loci: if someone tells a story that happens in a familiar space – or a space for which you have a cultural model, such as a story set in a generic Western kitchen –

the objects and actions in the story get "placed" in a way that can be either quasi-visual or kinesthetic or both. If you can't put the elements of the story in their places then you won't be able to hang onto them, just as we had trouble hanging onto the balcony and windows in the ticket seller's instructions.

This story connects to a larger theme about language. I want to believe that an utterance has no meaning outside the particular concrete setting where it's used. But then how can we can talk about things that are distant or hypothetical? Consider the examples that linguists ask you to evaluate out of context, like those sentences where you're supposed to indicate whether the pronoun can refer to (a) John or (b) Bill. Often I've found an interpretation not-OK until I work out a hypothetical context that makes it OK.

I suspect that such exercises get consistent results only because of cultural conventions about the default contexts. The willingness to evaluate decontextualized sentences *at all* is culturally specific. The cognitive anthropologist A. R. Luria (1976), for example, found his informants refusing to answer syllogistic reasoning tests until they knew the particulars of each sentence – the equivalents of the John and Bill who remain so comfortable as ciphers in my own culture. These people haven't learned the language game of decontextualized grammar and syllogism quizzes, and don't care to.

Heath (1983) describes how this capacity for decontextualization arises. She found that middle-class parents use rituals like bedtime-story-reading to introduce children to decontextualized letters, words, and forms of speech. Children who don't get this training have a harder time relating to decontextualized school exercises. Decontextualization is a complex and culturally specific skill laid on top of the more natural ability to relate language to familiar contexts. This has led to two quite different phenomena – one of them more fundamental and universal than the other – being run together and confused, with the later, more articulated phenomenon (the ability to perform certain tricks with decontextualized representations) getting all the credit.

A story about my routines for reading the Sunday Globe

This is a story about the indeterminacy of plans. The plan was something I said to myself as part of a settled routine. And I have a theory about the role of plans in settled routines. A routine might start out being mediated by an imperative utterance, such as a command you subarticulate to yourself. As the routine settles down, I hypothesize, it will still exhibit all the underdetermination, ambiguity, and indexicality of the original utterance, long after you've lost all awareness of any English being involved. Just because the activity has gotten "compiled" – as computer people would say – doesn't mean that the connection between the plan and the concrete situation of using it becomes any less problematic. Why? Perhaps you never stop saying the plan-text to yourself; perhaps as you routinize your actions you routinize your interpretive process as well.

Here, then, is the story. The Boston Globe recently began an expanded arts section in their Sunday edition. Called "Arts Etc", it gathers all the Sunday movie, arts, book reviews, arts schedules and advertising, and high-brow cultural commentary. This section doesn't have clearly delineated departments except for the final few pages, which are marked off for book reviews. The book review department, in fact, is wholly unchanged from the pre-Arts-Etc Sunday Globe. The first of the book review pages has its own banner and distinctive format, and all of the longer book reviews begin on that page and continue inside, where there are also shorter reviews and lists of best sellers and so forth.

Now, when reading the newspaper I will often come across the continuation of an article that looks interesting even though I hadn't noticed it when I was reading the page on which it began. So I'll have to back up to the earlier page to read the beginning of the article.

Last Sunday, then, I was reading a book review in the Globe. In particular, I was in the interior of the book review section, having followed an article from the book review section's first page (which, let us say, was page C15), when I came upon a headline about an author I was interested in reading about. Focusing on this headline, I found that it was a continuation. Whereupon, oddly, I turned to page C1 – i.e., the front page of the whole arts section – and not to page C15 – i.e., the first page of the book review section. I *knew* that all the book reviews began on C15, not C1, but I turned to C1 anyway. When I got C1 in front of me, it was not at all what I expected; momentarily confused, I figured out that I should turn to C15 instead.

Saying "C1" and "C15" is of course misleading. I don't think I knew that it was section "C" or page 15. My mistake, I think, turned on my never having reflected on the odd relationship between the two pages: the book review section was a clear "part" of the arts section, but the arts section didn't have any other clear "parts". Both page C1 and page C15 were the "front" of something – the arts section and the book review section, respectively. I had long been familiar with the Sunday Globe book review section's format, and its design and layout did little to make it look like a part of the superordinate arts section.

Here's what I think happened. When I went to turn to the beginning of the review I wanted to read, I turned to "the front", perhaps even "the front of the section". I'm not sure what I mean by double-quoting those two English phrases, but I want to mean something fairly literal. That is, I think I made my mistake because I was saying a phrase to myself in English, the phrase was ambiguous, and I interpreted it wrongly.

I've been reading newspapers for at least fifteen years and Boston Sunday Globe book reviews for almost ten years. Stumbling upon the continuation of an article and wanting to find its beginning is a routine I've been through many hundreds of times. And at least a couple dozen of these episodes must have occurred during my reading of the Sunday Globe's book reviews.

Why did I know to turn to the "front" of the section? Not all articles in the rest of the paper start at fronts of sections. I think at some point I noticed, and articulated to myself, that the book reviews start at the front of the book review section. So this was not the first time I've said "go to the front of the section" to myself in my head in such a situation. This internal uttering-to-myself and the actions I typically take in consequence of it must certainly have worn deep grooves in my brain by now. You might think that it was so thoroughly "compiled" that it no longer resembled English. Yet still it was capable of this very language-like underspecification of the situation. I still had to figure out what the English phrase was referring to in this specific concrete situation, and even though this figuring occurred perfectly automatically, smoothly, and routinely, it was still problematic. I could still get it wrong.

So "the front of the section" was ambiguous in this situation. But all of this still doesn't explain why, on the particular moment in question, it led me to turn to C1 rather than C15. Before the Globe reformatted and publicized its "Arts Etc", I had never given any particular thought to the idea of the Sunday Globe having an "arts section". In fact I clearly recall the first Sunday of the new section: the front page of the paper – i.e., A1 – had an ad for it, and despite the silly name I decided to give it a fair try. In fact it contained a useful article about Soviet dissident films, a topic that interested me. The matter of "the Globe's new arts section" – in exactly those words – had thus been on my mind. I don't want to conclude that the arts-section interpretation was "stronger" than the book-review-section interpretation, but whatever is operating as we constantly use background information to "fill in the details" of utterances when determining their relevance to particular concrete situations was operating here as well.

What I'm trying to figure out

Implicit in these stories are some ideas about representation and action.

1. Writing model of representation. By “writing”, I mean that the representation is something apart from you. It is a resource in situated action (Suchman 1987). You have to make sense of it in each next situation. There are no complete, systematic, guaranteed rules for this making-sense. And when you *do* manage to figure out what in some situation a representation is talking about, there is no way to finish listing what about the situation enabled you to do this.
2. Dependency model of routine evolution. All forms of activity are snapshots in the evolution of routines. The routines themselves are intertwined with the patterns of society and with the layout of particular places. The model says: when you think a new thought in some situation, you connect it back to its premises (Stallman & Sussman 1977). When you believe those same premises again in some future situation, you automatically call up the conclusion.

I’m trying to relate these two ideas. I think most concrete activity requires you to interpret representations, that is, “making sense of a representation in each next situation by figuring out what in the situation it is talking about”. So, for example, “turn left” will indicate different actions in different situations.

This sounds like a lot of work. If it were really a fresh challenge to make sense of “turn left” or “open the bottom drawer” on every moment, how would we ever do anything? But I don’t think it’s that bad. What you have in practice is a patchwork of routinized methods. In some past situation someone said “take the next left”, and you took certain specific concrete actions: you looked around, you searched for particular shapes and colors, and maybe you walked around and looked some more. You had your own reasons for doing all these things, and all of your actions and their reasons got connected, so now they’re ready and waiting to happen again in new situations. In subsequent situations some of the reasons might not have applied, so instead you took other actions, and these themselves led to new connections.

Perhaps after enough of this you develop a sufficient repertoire of routines to apply “turn left” to almost all of the left-turn situations you encounter in the average day. You’ve developed habits of interpretation. So whenever you tell yourself – or someone else tells you – “turn left”, you’ll be able to do it right away, “automatically”, without any hesitation or difficult figuring-out.

These routines, like all routines, will evolve (Agre 1985a, b). Very often the evolution of a routine involving a representation will permit you to undertake the activity without the representation being present in physical form. For example, using a recipe ten times might let you make the dish without the recipe.

One precondition for this effect seems to be that you understand the reasons for the recipe's instructions, but this turns out to be a complicated idea. In any event, there's a sense in which the representation never goes away. Even when you're routinely deciding to turn left or add salt, you're still – in some sense I wish I understood – saying the utterance "turn left" or "salt to taste" to yourself, and you're still interpreting it just like any other natural-language utterance that you need to relate to a concrete situation.

Acknowledgements

Conversations with Marty Hiller and Jeff Shrager helped me think about what people do with instructions and directions. David Chapman and Beth Preston carefully read a draft of this paper. Michael Mateas helped with the final revisions.

Patricia Morison helped with Latin etymology.

The "Upper Elk Tree Road" directions in the introduction and the computer message in "A story about photocopier supplies" are reproduced with the permission of their authors. I have altered them slightly to suppress identities and remove some comments on other matters.

Author's notes

This is a heavily abridged and revised version of a paper that I circulated informally in 1989. I have not tried to bring it up to date. Some parts of the argument are developed further in (Agre 1997). Another subsequent work that treats these issues is (Hutchins 1995).

References

Agre, Philip E. (1985a). The structures of everyday life. Working Paper 267, MIT Artificial Intelligence Laboratory.

Agre, Philip E. (1985b). Routines. AI Memo 828, MIT Artificial Intelligence Laboratory.

Agre, Philip, E. (1997). *Computation and human experience*. Cambridge: Cambridge University Press.

Amerine, Ronald. & Jack Bilmes (1988). Following instructions. *Human studies*, *11*, 327–339.

Atkinson, J. Maxwell & John Heritage (Eds.) (1984). *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press.

Auerbach, Erich (1953). W.R. Trask (Trans.) *Mimesis: The representation of reality in western literature*. Princeton: Princeton University Press.

Barwise, Jon & John Perry (1983). *Situations and attitudes*. Cambridge: MIT Press.

Brachman, Ronald J. & Hector J. Levesque (1984). The tractability of subsumption in frame-based description languages. *Proceedings of the national conference on Artificial Intelligence* (pp. 34–37). Austin, TX.

Brachman, Ronald J. & Hector J. Levesque (Eds.) (1985). *Readings in knowledge representation*. Los Altos, CA: Morgan Kaufmann.

Brodsky, Claudia J. (1987). *The imposition of form: Studies in narrative representation and knowledge*. Princeton: Princeton University Press.

Chapman, David (1991). *Vision, instruction, and action*. Cambridge: MIT Press.

Chapman, David & Philip E. Agre (1986). Abstract reasoning as emergent from concrete activity. In M. P. Georgeff & A. L. Lansky (Eds.), *Reasoning about actions and plans* (pp. 411–424). Los Altos, CA: Morgan-Kaufmann.

Culler, Jonathan (1982). *On deconstruction:Ttheory and criticism after structuralism*. Ithaca: Cornell University Press.

Dennett, Daniel (1978). *Brainstorms: Philosophical essays on mind and psychology*. Montgomery, VT: Bradford.

Derrida, Jacques (1976). G. C. Spivak (Trans.) *Of grammatology*. Baltimore: Johns Hopkins University Press.

Fodor, Jerry & Zenon Pylyshyn (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition, 28*, 3–72.

Garfinkel, Harold (1984). *Studies in ethnomethodology*. Oxford: Polity Press. Orginally published in 1967.

Goody, Jack (1986). *The logic of writing and the organization of society*. Cambridge: Cambridge University Press.

Hagen, Margaret (1986). *Varieties of realism: Geometries of representational art*. Cambridge: Cambridge University Press.

Harris, Roy (1980). *The language-makers*. Ithaca: Cornell University Press.

Harris, Roy (1987). *The language machine*. Ithaca: Cornell University Press.

Hartog, Francois (1988). J. Lloyd (Trans.) *The mirror of Herodotus: The representation of the other in the writing of history*. Berkeley: University of California Press.

Haugeland, John (1981). *Mind design: Philosophy, psychology, artificial inteligence.* Cambridge: MIT Press.

Heath, Shirley Brice (1983). *Ways with words: Language, life, and work in communities and classrooms*. Cambridge: Cambridge University Press.

Heritage, John (1984). *Garfinkel and ethnomethodology*. Cambridge: Polity Press.

Hopcroft, John E. & Dean B. Krafft (1987). The challenge of robotics for computer science. In J. T. Schwartz & C.-K. Yap (Eds.), *Algorithmic and geometric aspects of robotics, Volume 1* (pp. 7–42). Hillsdale, NJ: Erlbaum.

Hurlbert, Anya & Tomaso Poggio (1988). Making machines (and Artificial Intelligence) see. *Daedalus, 117*, 213–239.

Hutchins, Edwin (1986). Mediation and automatization. *Quarterly newsletter of the laboratory of comparative human cognition*, *8(2)*, 47-58.

Hutchins, Edwin. (1995). *Cognition in the wild*. Cambridge: MIT Press.

Ingarden, Roman (1973). R. A. Crowley & K. R. Olson (Trans.) *The cognition of the literary work of art*. Evanston: Northwestern University Press.

Judovitz, Dalia (1988). *Subjectivity and representation in Descartes: The origins of modernity*. Cambridge: Cambridge University Press.

Krieger, Murray (Ed.) (1987). *The aims of representation: Subject, text, history*. New York: Columbia University Press.

Latour, Bruno (1986). Visualisation and cognition: Thinking with eyes and hands. *Knowledge and society*, *5* (6), 1–40.

Levinson, Stephen C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

Luria, Alexander R. (1976). M. Lopez-Morillas & L. Solotaroff (Trans.), M. Cole (Ed.), *Cognitive development: Its cultural and social foundations*. Cambridge: Harvard University Press.

Newell, Allen & Herbert A. Simon (1963). GPS, a program that simulates human thought. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought* (pp. 279–293). New York: McGraw-Hill.

Newell, Allen & Herbert Simon (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Norris, Christopher (1982). *Deconstruction: Theory and practice*. London: Methuen.

Ong, Walter J. (1982). *Orality and literacy: The technologizing of the word*. London: Methuen.

Rorty, Richard (1979). *Philosophy and the mirror of nature*. Princeton: Princeton University Press.

Rumelhart, David E., Paul Smolensky, James L. McLelland, & Geoffrey E. Hinton (1986). Schemata and sequential thought processes in PDP models. In J. L. McLelland & D. E. Rumelhart, (Eds.), *Parallel sequential processing: Exploration in the microstructure of cognition, Volume 2: Psychological and biological models* (pp. 7–57). Cambridge: MIT Press.

Searle, John R. (1981). Minds, brains, and programs. In J. Haugeland (Ed.), *Mind design: Philosophy, psychology, Artificial Intelligence* (pp. 282–306). Cambridge: MIT Press.

Silvers, Stuart (1989). *Rerepresentation: Readings in the philosophy of mental representation*. Dordrecht: Kluwer.

Stallman, Richard M. & Gerald Jay Sussman (1977). Forward reasoning and dependency-directed backtracking in a system for computer-aided circuit analysis. *Artificial Intelligence*, *9* (2), 135–196.

Suchman, Lucy A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.

Vygotsky, L. S. (1978). M. Cole, V. John-Steiner, S. Scribner & E. Souberman (Eds.) *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press. Originally published in Russian in 1934.

Wallis, Brian (Ed.) (1984). *Art after modernism: Rethinking representation*. New York: New Museum of Contemporary Art.

Wertsch, James W. (1985). *Vygotsky and the social formation of mind*. Cambridge: Harvard University Press.

White, Hayden (1973). *Metahistory: The historical imagination in nineteenth-century europe*. Baltimore: Johns Hopkins University Press.

Zimmerman, Don H. (1970). The practicalities of rule use. In J. D. Douglas (Ed.), *Understanding everyday life: Toward the reconstruction of sociological knowledge* (pp. 221–238). London: Routledge and Kegan Paul.

Chapter 18

Stories and social networks

Warren Sack
University of California, Berkeley, CA

Introduction

What's so important about stories? The Internet has engendered a myriad of new social relations. These social relations, or "social networks" (see, for example, Wasserman & Galaskiewicz 1994) are forged by individuals through electronic mail and Internet-based chat. Some of the very active interchanges focus on movies, television programs, and news stories. In other words, a non-trivial portion of these social networks are based on discussions of widely circulated *stories*. Virtual, on-line communities are a result of these net-mediated, story-based relations.

To imagine that these new social relations (and the resultant virtual communities) are important, one must grasp the significance of storytelling in the first place. It matters which stories people know, which stories they tell, how they tell them, and how they are referred to. Narration, methods of citation and quotation, specific narratives, and general narrative forms constitute a kind of common sense upon which virtual and imaginary communities have been built (e.g., Anderson 1983). "... [C]ommon sense is our storehouse of narrative structures, and it remains the source of intelligibility and certainty in human affairs." (Schafer 1981). These presuppositions are the presuppositions of media studies (Hall 1982) and have also been integrated into some artificial intelligence (AI) research projects. The work of Roger Schank, Robert Abelson and their students is notable in this regard. Its close affinities with certain questions of media studies is unsurprising given the genealogy of the work. Robert Abelson did political analysis with media studies colleagues before his work in AI (e.g., de Sola Pool, Abelson & Popkin 1965).

A rather blurry line separates the Internet-based practices of relating and retelling widely-circulated stories authored by mass-media producers (e.g., Hollywood, CNN, etc.) from the practices of independently producing stories for Internet distribution. It is the former sort of practice that is the concern of this paper. Quotation, citation, and fragmentary repetition of stories are the life-blood of audience discussions and analysis of mass-produced stories Henry Jenkins discusses these audience practices as tactics of "poaching"; see Jenkins (1992). Audience members recirculate famous lines from movies (e.g., "Frankly my dear I don't give a damn," "I'll be back," "Make my day," etc.), comment on the plots and characters of known stories, summarize and retell pieces of stories for one another. The technology presented here is a first step towards a better understanding of story quotations, citations, and repetitions as the "threads" that weave people together into online, social networks.

A social network-based approach to story understanding differs from the standard approaches to "story understanding" that have been pursued by researchers in symbolic AI. Rather than examining stories as cognitive structures internal to individuals, the social network perspective is to see stories as shared ties that gather people into communities or social networks. An analogous difference in approaches to narrative theory was described by Mikhail Bakhtin in his critique of Russian Formalist approaches to literature and in his advocacy for a sociolinguistic method. See, for example, (Medvedev 1978). Bakhtin's "dialogical" approach to language and literature has been widely employed in contemporary literary theory, sociology, and media studies.

Moreover, unlike various media studies content analyses and structuralist analyses of narrative and film, it assumes the existence of an active, creative audience and uses audience activity (e.g., their discussion about a story) as the focus for gaining an understanding of stories. This distinction between research approaches in media studies (i.e., "content analysis" versus ethnographic approaches to the "active audience") has been recently summarized in books such as (Nightingale 1996). This alternative perspective shares some affinities with AI collaborative filtering techniques. Outside of AI, in the field of sociology, social network-based approaches to story understanding are not unusual, but the techniques of sociology can be improved through the use and development of an array of tools from natural language processing/computational linguistics. The research described here folds together insights from computational linguistics and the sociology of social networks to support the design of a new kind of story understanding technology; a technology predicated on the existence of verbally active story audiences.

A large amount of AI research is justified or motivated by pragmatic goals and there may in fact be pragmatic goals that would justify why we need a new technology of story understanding. In contrast, the poetics of AI have almost always been articulated around the need to get to know ourselves better. This poetics of the design and construction of intelligent, non-human entities has long been a theme of science fiction and science fantasy (not to mention its importance in philosophy since at least the time of Socrates when it was expressed as the ethical imperative "Know yourself.") Sherry Turkle nicely illustrates the how AI programs can function as a "second self" (Turkle 1984). It is within this tradition of poetics – what the philosopher Michel Foucault has described as "technologies of the self" (Foucault 1997) – that I would argue we need a new technology of story understanding. As new narrative forms are developed and new media proliferate, we need to invent new means for understanding how and where we are located in the emerging social networks.

1. Methodology

Methodology = Computational Sociolinguistics = Computational Linguistics + Quantitative Sociology.

Within the field of sociology, a number of computational approaches to understanding the social significance of literatures have been developed. Most prominently these methods have been applied to the literatures of science. For example, the methods of co-citation analysis (Garfield 1979) are routinely applied to determine the relative importance of a scientific article: its significance is thought to be a function of the number of other articles that cite it. AI elaborations of the techniques of co-citation analysis include (Lehnert et. al. 1990). The methods of social network theory (Wasserman & Galaskiewicz 1994) and actor-network theory (Callon et. al. 1986; Latour & Teil 1995) provide technologies akin to co-citation analysis, but have their own particular strengths and weaknesses. Co-word analysis, the computational technique associated with actor-network analysis, is basically the calculation of mutual probabilities between nouns in scientific abstracts and so this technique probably has more affinities with techniques in computational linguistics than with those developed by other sociologists.

These sorts of sociological "story understanding" technologies are very different from the story understanding technologies of an older, symbolic AI, but they have some affinities with techniques of newer AI work in agent-based architectures for information filtering and recommendation. Thus, for example, the "meaning" of a movie or television show for a collaborative filtering system (see Resnick & Varian 1997) is the set of ratings members of a user community have assigned to it. Users of such a system can be said to form a group to the extent that they have given similar ratings to the same items (cf., Lashkari 1995). For the most part these newer technologies (from sociology and from AI collaborative filtering research) for understanding stories as locations in and/or producers of social networks pay scant attention to the form and content of the stories: from this perspective stories are mostly "black boxes."

While the sociologists and AI, collaborative filtering researchers "black box" the form and content of stories, the corpus-based, computational linguistics and information retrieval researchers "black box" the social context of the stories they index (cf., Manning & Schutze 2000). Corpus-based computational linguistics is most often performed on large corpora described as, for instance, "10 million words from several volumes of the Wall Street Journal," or "1 million words from a wide variety of text genres." How the authors of the texts included in the corpora interact with one another or are related to one another is not factored into the analysis of the corpus. The one exception to this anonymity of authors is the use of corpus-based techniques for author identification purposes. But, even in these cases, the task is usually to determine who, among a small set of possible candidates, is the most likely author of a given text. The social network that incorporates (or the fact that no known social network incorporates) the set of candidate authors is not something that is often taken into account in the design of the corpus-based, computational linguistic methods of analysis.

The techniques of corpus-based, computational linguistics are oftentimes technically related to the techniques employed by sociologists since both sets of techniques can depend upon similar tools from statistics and information theory (e.g., measures of mutual information and entropy). But the techniques are inverses of one another due to the fact that what the sociologists black-box in their analyses is almost exactly what the corpus-based linguistics and information technology researchers do not black-box in their own research, and vice versa.

Any significantly new methodology for the development of a technology of story understanding should involve the combination of these two approaches. To understand a story as both (1) embedded in and (re)productive of both a network of related stories and other forms of discourse, and (2) as a facilitator or inhibitor of social networks, it is necessary to explore how social and semantic networks overlap. This intersection of social network and content analysis has been envisioned in sociology and linguistics (e.g., Milroy 1978). However, attempts to design and implement computer programs that combine sophisticated computational linguistic analysis with social network analysis are as yet unrealized.

Technology

System Design and Implementation. I have been analyzing Usenet newsgroup, audience discussions of popular television programs in an attempt to understand how the stories of television are pulled apart, reiterated, quoted, summarized, and – in general – appropriated into and used for the social networks of television viewers.

To analyze these and other newsgroups the *Conversation Map* system has been designed and implemented. The input to the system is an archive of thousands of messages from a newsgroup. The output of the system is four-fold and is pictured in the figure below.

1. Social Networks: The upper left-hand panel displays a social network showing who is in conversation with whom. The nodes of the network are labeled with the names of the participants in the newsgroup conversation. If two names are connected and close to one another, then the two participants have been responding to or quoting from each other more frequently than if they are connected but far apart from one another. Two names are connected if both participants have responded to or quoted from the other. In other words, the social network diagrams *reciprocity*. If someone in the conversation posts a lot

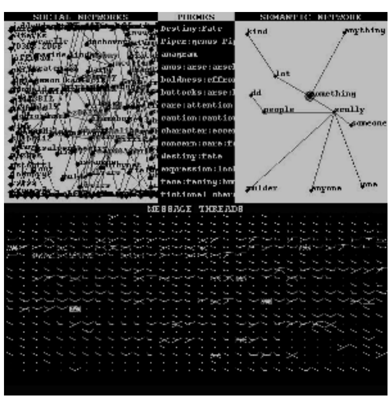

Figure 1. The Conversation Map interface.

of messages, but no one responds to those messages, then that someone will not show up in the social network.

2. Themes: The upper middle panel is a menu of discussion themes. Themes listed at the top of the menu are those themes that are most commonly used in the conversation. The list of discussion themes is extracted from the archives by examining the words and synonyms of words in quotations and replies to previous messages. In linguistics, this analysis is properly described as an analysis of *lexical cohesion* between messages (see Halliday & Hasan 1976). The links between participants in the social network are labeled with the discussion themes from the menu of themes.

3. Semantic Network: The upper right-hand panel displays a semantic network. If two terms in the semantic network are linked together, then those two terms have been found to be synonyms – or terms that may have similar meanings – in the conversation. The semantic network is produced through the application of corpus-based linguistics techniques referred to in the literature as techniques of "semantic extraction" and "automatic thesaurus construction" (cf., Hindle 1990; Hearst 1992).

4. Message Threads: The panel that occupies the lower half of the window is a graphical representation of all the messages that have been exchanged in the newsgroup conversation over a given period of time. The messages are organized into "threads," i.e., groups of messages that are responses, responses to responses, etc. of some given initial message. The threads are organized chronologically, from upper-left to lower-right. The oldest messages can be found in the upper left-hand corner.

For a newsgroup devoted to a television program, the computed themes and terms in the semantic network often include names of characters and episodes from the television show. Thus, these are the pieces of the television story that one can empirically observe as being appropriated into and employed by the audience's discussions of the story. Obviously, with a more sophisticated set of computational linguistic analysis tools one might observe larger portions of the narrative structure being woven into the audience's discussion (e.g., like the sorts of appropriations observed by Jenkins (1992)). However, the set of computational linguistic procedures we employ and have developed expressly for our system are more sophisticated than any others compared to contemporary, computational work on the social and linguistic analysis of Usenet newsgroup discussions (cf., Smith 1997; Best 1998; Donath et. al. 1999).

The analysis engine of the Conversation Map system performs the following steps on an archive of Usenet newsgroup messages in order to compute the four outputs described above:

1. Messages are threaded.
2. Quotations are identified and their sources (in other messages) are found.
3. A table of posters (i.e., newsgroup participants) to messages is built.
4. For every poster, the set of all other posters who replied to the poster is recorded. Posters who reciprocally reply to one another's messages are linked together in the social network.
5. The "signatures" of posters are identified and distinguished from the rest of the contents of each message.
6. The words in the messages are divided into sentences. The tool described in (Reynar & Ratnaparkhi 1997) is used.
7. Discourse markers (e.g., connecting words like "if", "therefore", "consequently", etc.) are tagged in the messages. We use a list of discourse markers compiled by (Marcu 1997).
8. Every word of every message is tagged according to its part-of-speech (e.g., "noun", "verb" "adjective", etc.) A simple trigram-based tagger is used to accomplish the part-of-speech tagging.
9. Every word is morphologically analyzed and its root is recorded. The database containing morphological and syntactic information comes from the University of Pennsylvania (Karp et al. 1992).
10. The words of the messages are parsed into sentences using a partial parser. The Conversation Map incorporates a re-implementation and revision of the parser described in (Grefenstette 1994).
11. An analysis of lexical cohesion is performed on every pair of messages where a pair consists of one message of a thread followed by a message that follows the message in the thread by either referencing it or quoting a passage from it. The lexical cohesion analysis procedure we have developed is akin to, but different than, the one described in (Hirst & St-Onge 1998). This analysis produces an approximation of the themes of discussion. The themes of the discussion label the arcs of the calculated social network. This allows one to see, for any given pair of posters, the theme of the posters' discussion.
12. The lexical and syntactic context of every noun in the archive is compared to the lexical and syntactic context of every other noun in the archive. An

> algorithm similar to the one in (Grefenstette 1994) is used. Nouns that are used or discussed in the same manner are calculated to be similar and are placed close to one another in the semantic networks. One can understand this semantic network as a crude approximation to the sorts of metaphors of discourse identified by linguists like George Lakoff (Lakoff & Johnson 1980). Thus, for example, if the noun "economy" and the noun "plant" are often associated with the same verbs and adjectives (e.g., "plants grow", the economy grows", "plants have roots", "the economy has roots", "we have a healthy economy", "we have a healthy plant" etc.) the two words will be closely coupled in the word associations network and one can read that network as stating something like "the economy is like a plant."

Three parts of the fourfold output of the system (social networks, themes, and semantic networks) correspond to the three *metafunctions* of language defined by the linguist Michael Halliday (Halliday 1994): the *interpersonal* (language connects people together), the *textual* (language connects itself together by referencing other pieces of language through practices like quotation), and the *ideational* (language contains or carries ideas in it that are associated with other ideas). The vast amount of research that has been done in sociolinguistics within a Hallidayean framework illustrates ways in which the current system could be improved if – for the kinds of work sociolinguists have been doing by hand – analogous computational linguistic techniques can be developed. A Hallidayean framework is also being applied by other researchers working on similar corpora, but with simpler computational text analysis procedures; see, for example, (Yates 1996).

A user's manual for the Conversation Map system and interfaces for several archives can be found on the web at this address: http://www.sims.berkeley.edu/∼sack/CM/. With the Conversation Map interface, the interested reader can explore example messages, social and semantic networks, and themes like those discussed in the following section.

Message archives

Two message archives will be discussed. Both archives contain messages posted to the Usenet newsgroup alt.tv.x-files, a group devoted to discussion of the internationally broadcast television show entitled *The X-files*. The Usenet newsgroup discussion is archived and publicly available at a variety of websites in-

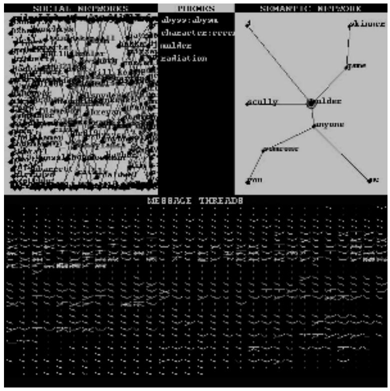

Figure 2. Conversation Map interface for Archive 2.

cluding www.deja.com. The staff at DejaNews was kind enough to provide us with the two archives discussed here.

The X-files is a weekly show produced by Twentieth Century Television in association with Fox Broadcasting Company. The show has two main characters, FBI Agents Dana Scully and Fox Mulder (played by actors Gillian Anderson and David Duchovny, respectively), who investigate cases reported to involve extraterrestrials, paranormal phenomena, and government conspiracy. It is an award winning television show now in its sixth season. More information about the show and short descriptions of the episodes can be found at the official *X-files* website: http://www.thex-files.com/.

Message Archive 1: These messages were exchanged during the week following the airing of the episode entitled “Quagmire” (4 May 1996 – 10 May 1996). In the “Quagmire” episode a Loch Ness monster-like creature is suspected of killing several people. About 700 participants posted over 1900 messages to the Usenet newsgroup alt.tv.x-files during this week after this episode was shown. A sketch of the analyzed messages from this archive can be seen in Figure 1.

Message Archive 2: These messages were exchanged during the week following the airing of the episode entitled “Hell Money” (30 March 1996 – 5 April 1996). The “Hell Money” episode concerns a high-stakes gambling game in which the players risk their own organs (e.g., their eyes and kidneys). Approximately 900 participants posted 2400 messages to the Usenet newsgroup after this episode. Figure 2 shows the Conversation Map automatically generated from the analysis of messages posted that week.

Preliminary discussion

Before proceeding to a closer examination of the Conversation Maps, two points need to be made.

Firstly, in many structuralist, formalist, and/or older Marxist-inspired analyses of narrative and media audiences, the audience member is often assumed to be a "cultural dupe." That is to say, it is assumed that a story delivered through the media (e.g., radio, television, the Internet, etc.) is not really open to interpretation and/or appropriation and means, more or less, one – and only one – thing. Moreover, the one and only meaning of the story is exactly what the audience member receives and, in this reception, is seen to be "programmed" by the story to behave or think in a certain manner by the story. This description is an over simplification, but it underlies the heat generated in arguments over which stories should or should not be taught in schools (i.e., the debate over the so-called "canon") and also is a preferred viewpoint for many writers of non-fiction as well as that of past builders of AI technologies for "story understanding" who believed a machine could be built to understand "the point" of a story. On the other end of the realist-to-relativist spectrum are many post-structuralist and cultural studies-inspired media scholars who have tended to emphasize the extraordinary creativity of audience members. Stories, and media productions in general, are seen as raw materials for audience members to rewrite, reinterpret, and recreate in novel and undetermined ways.

By spending some time with the Conversation Maps of audiences' online conversations, it should become clear that neither of these idealisms is empirically supported. On the one hand, the range of responses to the television stories is very diverse both in content and in genre. The "genres" of response include these: some responses are close intertextual analyses of the plot and characters of the episode, others are simple questions (e.g., "What's your favorite X-files episode?"), others are wildly tangential (e.g., "I have two kittens, one named Mulder, the other Scully, and I'm looking for someone to adopt them…"). On the other hand, only someone who is very easily amused will be likely to see the messages contained in these archives as wildly creative.

Thus, as a first point, I maintain that a machine-assisted, empirical examination of audience conversation makes it quite easy to resolve an issue that is often a point of debate in narrative theory and media studies: audience members are not "cultural dupes," but, neither are they more likely than any of the rest of us to be wildly creative with the "raw material" of the stories seen, heard, or read.

The second point also concerns the computational form of the analyses presented here. It has often been the case that audience studies have been formulated and written in a specialist's language (e.g., the vocabulary of academic media studies) and presented in a medium unlike the medium of the story and unlike the media used by the audience members to communicate amongst themselves. For example, studies of television audiences are oftentimes written up as academic books. For Internet-based audiences, it is now possible to build technologies that are designed to be accessible to the audience members and specialists alike. The Conversation Map system has been designed to be available online.

My second preliminary point is this: audience-accessible, networked, media studies cannot – as previous work repeatedly has – treat audiences as commodities or scientific objects because the network provides a means for the audience members to dispute the interpretations offered by the specialists. Consequently, what is presented below can best be understood as one place to begin an examination of the audiences' understandings of the two X-files episodes. It is not a definite, final discovery of those understandings.

Two conversation maps

In what follows, the social networks, themes, and semantic networks displayed in the Conversation Maps of the two message archives will be more closely examined.

Social networks

Figures 3 and 4 are enlargements of the social networks visible in Figures 1 and 2 respectively. In Figures 3 and 4 the names of the newsgroup participants have been turned off to allow one to see the topology of the networks more clearly.

What should be clear in Figures 3 and 4 is that participants are grouped into many small networks. The small networks are not connected to one another although it can be seen that the social networks shown in Figure 3 are more highly connected than the networks shown in Figure 4. In Figure 3, for example, the circled participant is a "lynchpin" of sorts holding together several smaller networks.

The lack of connections in the social networks makes it immediately apparent that the newsgroup is a space in which many different, probably unrelated,

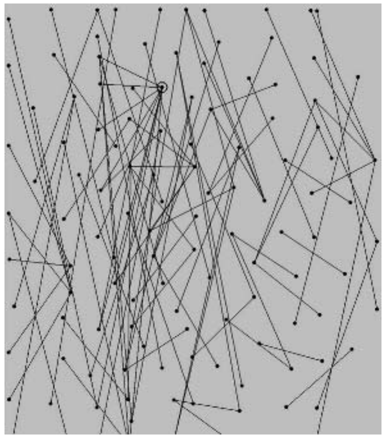

Figure 3. Social Network for Archive 1.

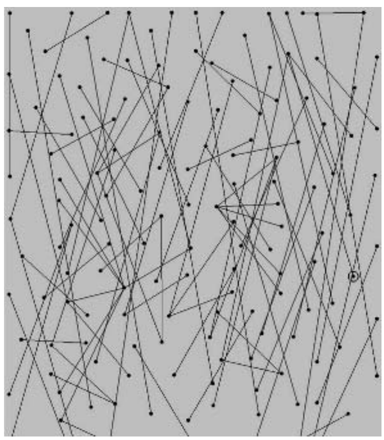

Figure 4. Social Network for Archive 2.

conversations are happening. Obviously the "effects" of a television story do not include the straightforward production of a cohesive social order.

It is interesting to compare the interconnections of these social networks with the social networks of other types of online discussions. Some of these can be seen here: www.sims.berkeley.edu/~sack/CM/.

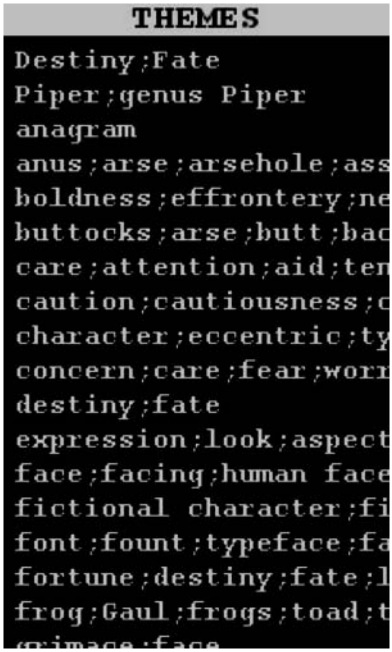

Figure 5. Themes Menu for Archive 1.

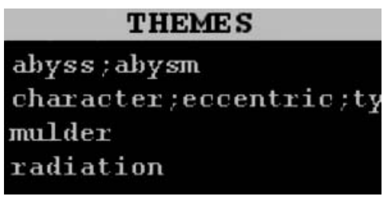

Figure 6. Themes Menu for Archive 2.

Themes

Another measure of the diversity of conversation in a newsgroup is provided by the menu of computed "discussion themes" (i.e., what in linguistics would more properly be described as the *lexical ties* between messages). Figures 5 and 6 list the tops of the theme menus for message archives 1 and 2 respectively. Themes in the menus of themes are ordered according to the number of arcs in the social network that they label.

Recall that an arc in the social network connects two newsgroup participants if and only if those two participants have replied to each other or cited from one another's' messages. Thus, for example, A and B are connected in the social network and the arc between A and B is labeled with a theme – e.g., "sports" – if and only if A and B have had at least one interchange like the following: A posts a message about baseball, B replies with a post about football, B posts a message about swimming, and A cites or responds to B's message with one about skiing. Since baseball, football, swimming, and skiing are all sports, the link between A and B might be labeled with the more abstract term "sports" (computed by the Conversation Map system using the WordNet thesaurus, version 1.6). So, the themes listed in the menus are only there if there has been one or more reciprocated responses in which the theme (or a semantically similar) term was mentioned in each of the exchanged messages.

Figure 6, showing the reciprocated discussion themes in the messages of archive 2, is a surprisingly short list. Usually the menu of themes lists many items. Clicking on the items to highlight the parts of the social network that they label shows even more clearly how fragmented the discussion of archive 2 is. All of the themes listed connect only one pair of posters. In short, only a small handful of the interchanges concerning the "Hell Money" episode are focused around a specific theme of discussion.

Figure 5, showing the reciprocated discussion themes in the messages of archive 2, again shows that the social interchange visible in the message archives is more cohesive in the first archive than it is in the second archive. This can be interpreted from the longer list of reciprocated themes for archive 1.

Semantic networks

The semantic networks shown in Figures 7 and 8 show that the conversations after both episodes are concerned with the main characters (Scully and Mulder). Moreover, it is interesting to see the computed similarities between the main characters and the more generic terms of "you," "me," "someone," "anyone" etc. These calculations provide a way of seeing how the audience members talk about themselves in ways comparable to the way they talk about the main characters. This calculation might be compared to analyses of character "identification" discussed in the literatures of film theory and other media studies.

Conclusions

A computational sociolinguistic analysis of stories has been proposed and implemented in the *Conversation Map* system. The significance of a story is seen as

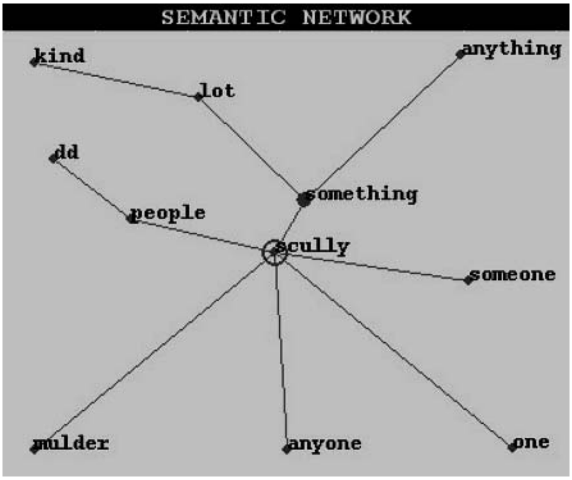

Figure 7. Semantic Network from Archive 1.

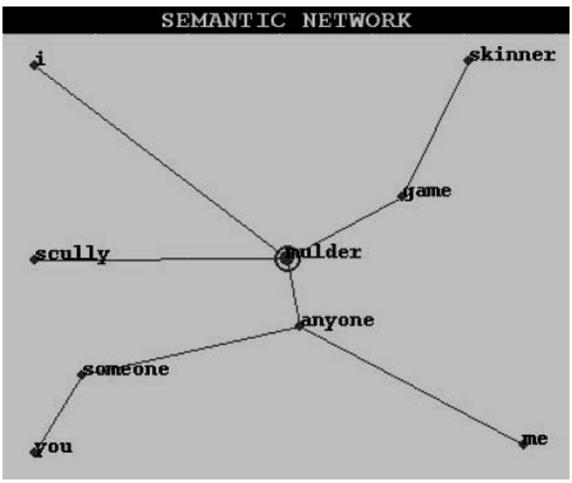

Figure 8. Semantic Network from Archive 2.

a function of the social network that it engenders and/or inflects. The proposed analysis method was compared to related work in AI collaborative filtering, sociology, and computational corpus-based linguistics. It was also briefly compared to the relatively unrelated work in story understanding done within the symbolic AI tradition. The Conversation Map system has been designed and implemented to perform a sociolinguistic analysis of Usenet newsgroup analysis postings and it has been employed in the analysis of television audiences' newsgroup discussions of stories from a popular television show. The output of the implemented system illustrates sociolinguistic analyses of the television stories as they are visible in the social networks and language of the television audiences' newsgroup postings. It is hoped that a tool like the Conversation Map system can be used as a future "technology of the self" with which audiences can critically reflect on the emergent social and semantic structures of their online discussions.

Acknowledgments

As an undergraduate and later a research assistant at the Yale University Artificial Intelligence Lab, I gained an early appreciation for the symbolic AI approach to story understanding with the help of many of my friends and acquaintances who were students and colleagues of Roger Schank and Robert Abelson. Thanks to Larry Birnbaum who taught the first natural language processing course I ever took and to many other friends and acquaintances from the old days of the Yale AI Lab. I first came to the MIT Media Laboratory at the invitation of Marc Davis and Mike Travers to present my work at the meeting of the Narrative Intelligence (NI) Reading Group in 1991. One of the main reasons I chose the Media Lab as the place to do my graduate work was the wonderful interchange I experienced with the NI Group. With them I learned more about narrative theory and media studies. I hope the present work reflects this and can be seen as an interesting outgrowth that embodies some of the concerns of the old Yale school, but diverges enough from that work to open up one of many, possible, new, NI approaches to examining story understanding with computational means. Although the MIT Media Lab NI Group no longer meets together face-to-face, we are still alive online as a mailing list. Thanks to Phoebe Sengers and Michael Mateas for organizing the AAAI Workshop that gave us a face-to-face meeting including some of the MIT Media Lab NI Group and many new faces too. Thanks also to Phoebe and Michael for discussions and online exchanges about this paper and other NI-related subjects.

References

Anderson, Benedict (1983). *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.

Berthold, Michael, Fay Sudweeks, Sid Newton, & Richard Coyne (1998). It makes sense: Using an autoassociative neural network to explore typicality in computer mediated discussions. In Fay Sudweeks, Margaret McLaughlin, and Sheizaf Rafaeli (Eds.), *Network and netplay: Virtual groups on the internet* (pp. 191-219). Cambridge, MA: AAAI/MIT Press.

Best, Michael L. (1998). Corporal ecologies and population fitness on the net. *Journal of Artificial Life, 3* (4).

Callon, Michel , John Law, & Arie Rip (Eds.) (1986). *Mapping the dynamics of science: sociology in the real world*. London: Macmillan.

Donath, Judith, Karrie Karahalios, & Fernanda Viegas (1999). Visualizing conversations. *Proceedings of HICSS-32*, Maui, HI, January 5–8.

Foucault, Michel (1997). Technologies of the self. In Paul Rabinow (Ed.), Robert Hurley & others (Trans.), *Ethics: Subjectivity and truth (Essential works of Foucault 1954–1984)* (pp. 223–251), Volume One. New York: The New Press.

Garfield, Eugene (1979). *Citation indexing: Its theory and applications in science, technology and humanities*. New York: John Wiley.

Grefenstette, Gregory (1994). *Explorations in automatic thesaurus discovery*. Boston: Kluwer Academic Publishers.

Hall, Stuart (1982). The rediscovery of 'ideology': Return of the repressed in media studies. In M. Gurevitch, T. Bennett, J. Curran, J. Woollacott (Eds.), *Culture, society and the media*. New York: Routledge.

Halliday, Michael A.K. (1994). *An introduction to functional grammar, second edition*. London: Edward Arnold.

Halliday, Michael A.K. & Ruqaiya Hasan (1976). *Cohesion in English.* New York: Longman.

Hearst, Marti A. (1992). Automatic extraction of hyponyms from large text corpora. In *Proceedings of the fourteenth international conference on computational linguistics* (pp. 183–191), Nantes, France: COLING-92.

Herring, Susan, Deborah A. Johnson, & Tamra DiBenedetto (1995). 'This discussion is going too far!': Male resistance to female participation on the Internet. In Kira Hall and Mary Bucholtz (Eds.), *Gender articulated: Language and the socially constructed self* (pp. 67–96). New York: Routledge.

Hindle, Donald (1990). Noun classification from predicate-argument structures. In *Proceedings of the 27th Annual meeting of the association for computational linguistics* (pp. 268–275). Pittsburgh, Pennsylvania: ACL.

Hirst, Graeme & David St-Onge (1998). Lexical chains as representations of context for the detection and correction of malapropisms. In Christiane Fellbaum (Ed.), *WordNet: An Electronic Lexical Database* (pp. 305–332). Cambridge, MA: MIT Press.

Jenkins, Henry (1992). *Textual poachers: Television fans and participatory culture*. New York: Routledge.

Karp, Daniel, Yves Schabes, Martin Zaidel, & Dania Egedi (1992). A freely available wide coverage morphological analyzer for English. In *Proceedings of the 14th international conference on computational linguistics*, Nantes, France: COLING-92.

Lakoff, George & Mark Johnson (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lashkari, Yezdezard (1995). *Feature guided automated collaborative filtering*. MIT Media Laboratory, Master's Thesis.

Latour, Bruno & Geneviève Teil (1995). The Hume Machine: Can association networks do more than formal rules. *Stanford humanities review, 4* (2) (Special issue on Artificial Intelligence), 47–65.

Lehnert, Wendy, Claire Cardie, & Ellen Riloff (1990). Analyzing research papers using citation sentences. In *Proceedings of the 12th annual conference on cognitive science* (pp. 511–518). Cambridge, MA: Lawrence Erlbaum Associates.

Manning, Christopher D. and Heinrich Schutze (2000). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.

Marcu, Daniel (1997). *The rhetorical parsing, summarization, and generation of natural language texts*. Ph.D. Thesis. Toronto: Department of Computer Science, University of Toronto.

Medvedev, Pavel Nikolaevich [Mikhail Mikhailovich Bakhtin] (1978). *The formal method in literary scholarship: A critical introduction to sociological poetics*. Baltimore: Johns Hopkins University Press.

Milroy, Lesley (1978) *Language and social networks*. Baltimore: University Park Press.

Nightingale, Virginia (1996). *Studying audiences: The shock of the real*. New York: Routledge.

Reynar, Jeffrey C. & Adwait Ratnaparkhi (1997). A maximum entropy approach to identifying sentence boundaries. In *Proceedings of the fifth conference on applied natural language processing* (pp. 16–19). Washington, D.C.

Resnick, Paul and Hal Varian (1997). Recommender Systems, *Communications of the ACM, 40* (3).

Schafer, Roy (1981). Narration in the psychoanalytic dialogue. In W.J.Thomas Mitchell (Ed.), *On narrative* (pp. 29–53). Chicago: University of Chicago Press, pages

Smith, Marc (1997). Netscan: Measuring and mapping the social structure of usenet. Presented at the *17th Annual international sunbelt social network conference*, Bahia Resort Hotel, Mission Bay, San Diego, California, February 13–16 (see www.sscnet.ucla.edu/soc/csoc/papers/sunbelt97/).

de Sola Pool, Ithiel , Robert P. Abelson & Samuel L. Popkin (1965). *Candidates, issues and strategies: A computer simulation of the 1960 and 1964 presidential elections*. Cambridge, MA: MIT Press.

Turkle, Sherry (1984). *The second self: Computers and the human spirit*. New York: Simon and Schuster.

Wasserman, Stanley & Joseph Galaskiewicz (Eds.) (1994). *Advances in social network analysis: Research in the social and behavioral sciences*. Thousand Oaks, CA: Sage Pub.

Yates, Simeon J. (1996). Oral and written linguistic aspects of computer conferencing. In Susan C. Herring (Ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*. Philadelphia: John Benjamins Pub. Co.

Author index

Subject index

In the series ADVANCES IN CONSCIOUSNESS RESEARCH (AiCR) the following titles have been published thus far or are scheduled for publication:

1. GLOBUS, Gordon G.: *The Postmodern Brain.* 1995.
2. ELLIS, Ralph D.: *Questioning Consciousness. The interplay of imagery, cognition, and emotion in the human brain.* 1995.
3. JIBU, Mari and Kunio YASUE: *Quantum Brain Dynamics and Consciousness. An introduction.* 1995.
4. HARDCASTLE, Valerie Gray: *Locating Consciousness.* 1995.
5. STUBENBERG, Leopold: *Consciousness and Qualia.* 1998.
6. GENNARO, Rocco J.: *Consciousness and Self-Consciousness. A defense of the higher-order thought theory of consciousness.* 1996.
7. MAC CORMAC, Earl and Maxim I. STAMENOV (eds): *Fractals of Brain, Fractals of Mind. In search of a symmetry bond.* 1996.
8. GROSSENBACHER, Peter G. (ed.): *Finding Consciousness in the Brain. A neurocognitive approach.* 2001.
9. Ó NUALLÁIN, Seán, Paul MC KEVITT and Eoghan MAC AOGÁIN (eds): *Two Sciences of Mind. Readings in cognitive science and consciousness.* 1997.
10. NEWTON, Natika: *Foundations of Understanding.* 1996.
11. PYLKKÖ, Pauli: *The Aconceptual Mind. Heideggerian themes in holistic naturalism.* 1998.
12. STAMENOV, Maxim I. (ed.): *Language Structure, Discourse and the Access to Consciousness.* 1997.
13. VELMANS, Max (ed.): *Investigating Phenomenal Consciousness. Methodologies and Maps.* 2000.
14. SHEETS-JOHNSTONE, Maxine: *The Primacy of Movement.* 1999.
15. CHALLIS, Bradford H. and Boris M. VELICHKOVSKY (eds.): *Stratification in Cognition and Consciousness.* 1999.
16. ELLIS, Ralph D. and Natika NEWTON (eds.): *The Caldron of Consciousness. Motivation, affect and self-organization – An anthology.* 2000.
17. HUTTO, Daniel D.: *The Presence of Mind.* 1999.
18. PALMER, Gary B. and Debra J. OCCHI (eds.): *Languages of Sentiment. Cultural constructions of emotional substrates.* 1999.
19. DAUTENHAHN, Kerstin (ed.): *Human Cognition and Social Agent Technology.* 2000.
20. KUNZENDORF, Robert G. and Benjamin WALLACE (eds.): *Individual Differences in Conscious Experience.* 2000.
21. HUTTO, Daniel D.: *Beyond Physicalism.* 2000.
22. ROSSETTI, Yves and Antti REVONSUO (eds.): *Beyond Dissociation. Interaction between dissociated implicit and explicit processing.* 2000.
23. ZAHAVI, Dan (ed.): *Exploring the Self. Philosophical and psychopathological perspectives on self-experience.* 2000.
24. ROVEE-COLLIER, Carolyn, Harlene HAYNE and Michael COLOMBO: *The Development of Implicit and Explicit Memory.* 2000.
25. BACHMANN, Talis: *Microgenetic Approach to the Conscious Mind.* 2000.
26. Ó NUALLÁIN, Seán (ed.): *Spatial Cognition. Selected papers from Mind III, Annual Conference of the Cognitive Science Society of Ireland, 1998.* 2000.
27. McMILLAN, John and Grant R. GILLETT: *Consciousness and Intentionality.* 2001.

28. ZACHAR, Peter: *Psychological Concepts and Biological Psychiatry. A philosophical analysis.* 2000.
29. VAN LOOCKE, Philip (ed.): *The Physical Nature of Consciousness.* 2001.
30. BROOK, Andrew and Richard C. DeVIDI (eds.): *Self-reference and Self-awareness.* 2001.
31. RAKOVER, Sam S. and Baruch CAHLON: *Face Recognition. Cognitive and computational processes.* 2001.
32. VITIELLO, Giuseppe: *My Double Unveiled. The dissipative quantum model of the brain.* 2001.
33. YASUE, Kunio, Mari JIBU and Tarcisio DELLA SENTA (eds.): *No Matter, Never Mind. Proceedings of Toward a Science of Consciousness: Fundamental Approaches, Tokyo, 1999.* 2002.
34. FETZER, James H.(ed.): *Consciousness Evolving.* 2002.
35. Mc KEVITT, Paul, Seán Ó NUALLÁIN and Conn MULVIHILL (eds.): *Language, Vision, and Music. Selected papers from the 8th International Workshop on the Cognitive Science of Natural Language Processing, Galway, 1999.* 2002.
36. PERRY, Elaine, Heather ASHTON and Allan YOUNG (eds.): *Neurochemistry of Consciousness. Neurotransmitters in mind.* 2002.
37. PYLKKÄNEN, Paavo and Tere VADÉN (eds.): *Dimensions of Conscious Experience.* 2001.
38. SALZARULO, Piero and Gianluca FICCA (eds.): *Awakening and Sleep-Wake Cycle Across Development.* 2002.
39. BARTSCH, Renate: *Consciousness Emerging. The dynamics of perception, imagination, action, memory, thought, and language.* 2002.
40. MANDLER, George: *Consciousness Recovered. Psychological functions and origins of conscious thought.* 2002.
41. ALBERTAZZI, Liliana (ed.): *Unfolding Perceptual Continua.* 2002.
42. STAMENOV, Maxim I. and Vittorio GALLESE (eds.): *Mirror Neurons and the Evolution of Brain and Language.* 2002.
43. DEPRAZ, Natalie, Francisco VARELA and Pierre VERMERSCH.: *On Becoming Aware.* n.y.p.
44. MOORE, Simon and Mike OAKSFORD (eds.): *Emotional Cognition. From brain to behaviour.* 2002.
45. DOKIC, Jerome and Joelle PROUST: *Simulation and Knowledge of Action.* 2002.
46. MATHEAS, Michael and Phoebe SENGERS (ed.): *Narrative Intelligence.* 2003.
47. COOK, Norman D.: *Tone of Voice and Mind. The connections between intonation, emotion, cognition and consciousness.* 2002.
48. JIMÉNEZ, Luis: *Attention and Implicit Learning.* 2003.
49. OSAKA, Naoyuki (ed.): *Neural Basis of Consciousness.* 2003.
50. GLOBUS, Gordon G.: *Quantum Closures and Disclosures. Thinking-together post-phenomenology and quantum brain dynamics.* n.y.p.
51. DROEGE, Paula: *Caging the Beast. A theory of sensory consciousness.* n.y.p.