Lee, J.; et. al. “Modern Manufacturing”
Mechanical Engineering Handbook
Ed. Frank Kreith
Boca Raton: CRC Press LLC, 1999

Modern Manufacturing

Jay Lee
National Science Foundation

Robert E. Schafrik
National Research Council

Steven Y. Liang
Georgia Institute of Technology

Trevor D. Howes
University of Connecticut

John Webster
University of Connecticut

Ioan Marinescu
Kansas State University

K. P. Rajurkar
University of Nebraska-Lincoln

W. M. Wang
University of Nebraska-Lincoln

Talyan Altan
Ohio State University

Weiping Wang
General Electric R & D Center

Alan Ridilla
General Electric R & D Center

Matthew Buczek
General Electric R&D Center

S. H. Cho
Institute for Science and Technology, Republic of Korea

Ira Pence
Georgia Institute of Technology

Toskiaki Yamaguchi
NSK Ltd.

Yashitsuga Taketomi
NSK Ltd.

(continued on next page)

Carl J. Kempf
NSK Ltd.

John Fildes
Northwestern University

Yoram Koren
University of Michigan

M. Tomizuka
University of California-Berkeley

Kam Lau
Automated Precision, Inc.

Tai-Ran Hsu
San Jose State University

David C. Anderson
Purdue University

Tien-Chien Chang
Purdue University

Hank Grant
University of Oklahoma

Tien-I. Liu
California State University at Sacramento

J. M. A. Tanchoco
Purdue University

Andrew C. Lee
Purdue University

Su-Hsia Yang
Purdue University

Takeo Nakagawa
University of Tokyo

H. E. Cook
University of Illinois at Urbana-Champaign

James J. Solberg
Purdue University

Chris Wang
IBM

13.1 Introduction

Jay Lee and Robert Shafrik

Manufacturing is the means by which the technical and industrial capability of a nation is harnessed to transform innovative designs into well-made products that meet customer needs. This activity occurs through the action of an integrated network that links many different participants with the goals of developing, making, and selling useful things.

Manufacturing is the conversion of raw materials into desired end products. The word derives from two Latin roots meaning *hand* and *make*. Manufacturing, in the broad sense, begins during the design phase when judgments are made concerning part geometry, tolerances, material choices, and so on. Manufacturing operations start with manufacturing planning activities and with the acquisition of required resources, such as process equipment and raw materials. The manufacturing function extends throughout a number of activities of design and production to the distribution of the end product and, as necessary, life cycle support. Modern manufacturing operations can be viewed as having six principal components: materials being processed, process equipment (machines), manufacturing methods, equipment calibration and maintenance, skilled workers and technicians, and enabling resources.

There are three distinct categories of manufacturing:

- *Discrete item manufacturing*, which encompasses the many different processes that bestow physical shape and structure to materials as they are fashioned into products. These processes can be grouped into families, known as unit manufacturing processes, which are used throughout manufacturing.
- *Continuous materials processing*, which is characterized by a continuous production of materials for use in other manufacturing processes or products. Typical processes include base metals production, chemical processing, and web handing. Continuous materials processing will not be further discussed in this chapter.
- *Micro- and nano-fabrication*, which refers to the creation of small physical structures with a characteristic scale size of microns (millionths of a meter) or less. This category of manufacturing is essential to the semiconductor and mechatronics industry. It is emerging as very important for the next-generation manufacturing processes.

Manufacturing is a significant component of the U.S. economy. In 1995, 19% of the U.S. gross domestic product resulted from production of durable and nondurable goods; approximately 65% of total U.S. exports were manufactured goods; the manufacturing sector accounted for 95% of industrial research and development spending; and manufacturing industries employed a work force of over 19 million people in 360,000 companies. In the modern economy, success as a global manufacturer requires the development and application of manufacturing processes capable of economically producing high-quality products in an environmentally acceptable manner.

Modern Manufacturing

Manufacturing technologies address the capabilities to design and to create products, and to manage that overall process. Product quality and reliability, responsiveness to customer demands, increased labor productivity, and efficient use of capital were the primary areas that leading manufacturing companies throughout the world emphasized during the past decade to respond to the challenge of global competitiveness. As a consequence of these trends, leading manufacturing organizations are flexible in management and labor practices, develop and produce virtually defect-free products quickly (supported with global customer service) in response to opportunities, and employ a smaller work force possessing multidisciplinary skills. These companies have an optimal balance of automated and manual operations.

To meet these challenges, the manufacturing practices must be continually evaluated and strategically employed. In addition, manufacturing firms must cope with design processes (e.g., using customers' requirements and expectations to develop engineering specifications, and then designing components), production processes (e.g., moving materials, converting materials properties or shapes, assembling products or components, verifying processes results), and business practices (e.g., turning a customer order into a list of required parts, cost accounting, and documentation of procedures). Information technology will play an indispensable role in supporting and enabling the complex practices of manufacturing by providing the mechanisms to facilitate and manage the complexity of manufacturing processes and achieving the integration of manufacturing activities within and among manufacturing enterprises. A skilled, educated work force is also a critical component of a state-of-the-art manufacturing capability. Training and education are essential, not just for new graduates, but for the existing work force.

Manufacturing is evolving from an art or a trade into a science. The authors believe that we must understand manufacturing as a technical discipline. Such knowledge is needed to most effectively apply capabilities, quickly incorporate new developments, and identify the best available solutions to solve problems. The structure of the science of manufacturing is very similar across product lines since the same fundamental functions are performed and the same basic managerial controls are exercised.

13.2 Unit Manufacturing and Assembly Processes

Robert E. Schafrik

There are a bewildering number of manufacturing processes able to impart physical shape and structure to a workpiece. However, if these processes are broken down into their basic elements and then examined for commonality, only a few fundamental processes remain. These are the building blocks, or unit processes, from which even the most complicated manufacturing system is constructed. This section describes these unit processes in sufficient detail that a technically trained person, such as a design engineer serving as a member of an integrated product and process design team comprised of members from other specialties, could become generally knowledgeable regarding the essential aspects of manufacturing processes. Also, the information presented in this section will aid such an individual in pursuing further information from more specialized manufacturing handbooks, publications, and equipment/tool catalogs.

Considering the effect that a manufacturing process has on workpiece configuration and structure, the following five general types of unit manufacturing process can be identified (Altan et al., 1983; NRC, 1995):

Material removal processes — Geometry is generated by changing the mass of the incoming material in a controlled and well-defined manner, e.g., milling, turning, electrodischarge machining, and polishing.

Deformation processes — The shape of a solid workpiece is altered by plastic deformation without changing its mass or composition, e.g., rolling, forging, and stamping.

Primary shaping processes — A well-defined geometry is established by bulk forming material that initially had no shape, e.g., casting, injection molding, die casting, and consolidation of powders.

Structure-change processes — The microstructure, properties, or appearance of the workpiece are altered without changing the original shape of the workpiece, e.g., heat treatment and surface hardening.

Joining and assembly processes — Smaller objects are put together to achieve a desired geometry, structure, and/or property. There are two general types: (1) consolidation processes which use mechanical, chemical, or thermal energy to bond the objects (e.g., welding and diffusion bonding) and (2) strictly mechanical joining (e.g., riveting, shrink fitting, and conventional assembly).

Unit Process Selection

Each component being manufactured has a well-defined geometry and a set of requirements that it must meet. These typically include

- Shape and size
- Bill-of-material
- Accuracy and tolerances
- Appearance and surface finish
- Physical (including mechanical) properties
- Production quantity
- Cost of manufacture

In order to satisfy these criteria, more than one solution is usually possible and trade-off analyses should be conducted to compare the different approaches that could be used to produce a particular part.

Control and Automation of Unit Processes

Every unit process must be controlled or directed in some way. The need for improved accuracy, speed, and manufacturing productivity has spurred the incorporation of automation into unit processes regarding both the translation of part design details into machine instructions, and the operation of the unit process itself and as a subsystem of the overall production environment. The section of this chapter on computer-aided design/computer-aided manufacturing (CAD/CAM) discusses the technology involved in creating and storing CAD files and their use in CAM. The expectations of precision are continuing to change, as indicated in Figure 13.2.1. This drive for ever-tighter tolerances is helping spur interest in continual improvements in design and manufacturing processses.

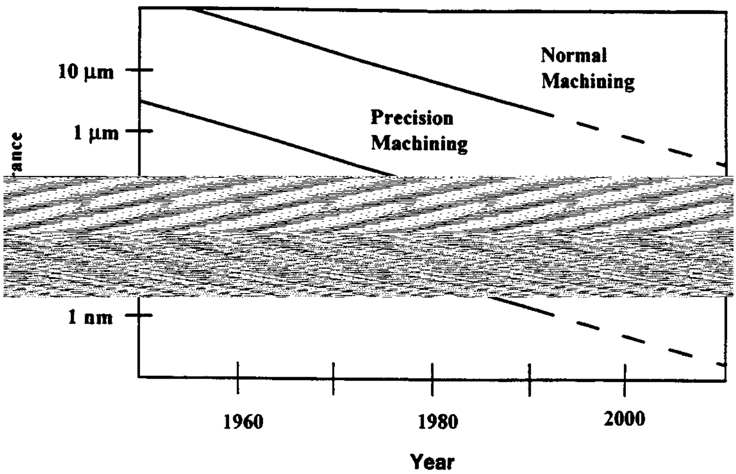

FIGURE 13.2.1 Precision machining domains. (From NRC, *Unit Manufacturing Processes,* National Academy Press, Washington, D.C., 1995, 169. With permission.)

Modern machine tool controls are emphasizing two areas: adaptive control and communication. For *adaptive control* the controller must adapt its control gains so that the overall system remains at or near the optimal condition in spite of varying process dynamics. Expanded *communication* links the data collected by a unit process controller to other segments of the manufacturing operation. Data regarding production time and quantity of parts produced can be stored in an accessible database for use by inventory control and quality monitoring. This same database can then be used by production schedulers to avoid problems and costs associated with redundant databases.

At the factory level, machining operations employing two or more numerically controlled (NC) machine tools may use a separate mainframe computer that controls several machine tools or an entire shop. The system is often referred to as *distributed numerical control* (DNC).

Today many factories are implementing *flexible manufacturing systems* (FMS), an evolution of DNC. An FMS consists of several NC unit processes (not necessarily only machine tools) which are interconnected by an automated materials handling system and which employ industrial robots for a variety of tasks requiring flexibility, such as loading/unloading the unit process queues. A single computer serves as master controller for the system, and each process may utilize a computer to direct the lower-order tasks. Advantages of FMS include:

- A wide range of parts can be produced with a high degree of automation
- Overall production lead times are shortened and inventory levels reduced
- Productivity of production employees is increased
- Production cost is reduced
- The system can easily adapt to changes in products and production levels

Unit Processes

In the following discussion, a number of unit processes are discussed, organized by the effect that they have on workpiece configuration and structure. Many of the examples deal with processing of metals since that is the most likely material which users of this handbook will encounter. However, other materials are readily processed with the unit processes described in this chapter, albeit with suitable modifications or variations.

Mechanical assembly and material handling are also discussed in this section. On average, mechanical assembly accounts for half of the manufacturing time, and processes have been developed to improve the automation and flexibility of this very difficult task. Material handling provides the integrating link between the different processes — material-handling systems ensure that the required material arrives at the proper place at the right time for the various unit processes and assembly operations.

The section ends with a case study that demonstrates how understanding of the different unit processes can be used to make engineering decisions.

- Material removal (machining) processes
  - Traditional machining
    Drill and reaming
    Turning and boring
    Planing and shaping
    Milling
    Broaching
    Grinding
    Mortality
  - Nontraditional machining
    Electrical discharge machining
    Electrical chemical machining
    Laser beam machining
    Jet machining (water and abrasive)
    Ultrasonic machining
- Phase-change processes
  - Green sand casting
  - Investment casting
- Structure-change processes
  - Normalizing steel
  - Laser surface hardening
- Deformation processes
  - Die forging
  - Press-brake forming
- Consolidation processes
  - Polymer composite consolidation
  - Shielded metal-arc welding
- Mechanical assembly
- Material handling
- Case study: Manufacturing and inspection of precision recirculating ballscrews

References

Altan, T., Oh, S.I., and Gegel, H. 1983. *Metal Forming — Fundamentals and Applications*, ASM International, Metals Park, OH.

ASM Handbook Series, 10th ed., 1996. ASM International, Metals Park, OH.

Bakerjian, R., Ed. 1992. *Design for Manufacturability,* Vol. VI, *Tool and Manufacturing Engineers Handbook,* 4th ed., Society of Manufacturing Engineers, Dearborn, MI.
DeVries, W.R. 1991. *Analysis of Material Removal Processes*, Springer-Verlag, New York.
Kalpakjian, S. 1992. *Manufacturing Engineering and Technology*, Addison-Wesley, Reading, MA.
National Research Council (NRC), 1995. *Unit Manufacturing Processes — Issues and Opportunites in Research,* National Academy Press, Washington, D.C.

Material Removal Processes

These processes, also known as machining, remove material by mechanical, electrical, laser, or chemical means to generate the desired shape and/or surface characteristic. Workpiece materials span the spectrum of metals, ceramics, polymers, and composites, but metals, and particularly iron and steel alloys, are by far the most common. Machining can also improve the tolerances and finish of workpieces previously shaped by other processes, such as forging. Machining is an essential element of many manufacturing systems (ASM, 1989b; Bakerjian, 1992).

Machining is important in manufacturing because

- It is precise. Machining is capable of creating geometric configurations, tolerances, and surface finishes that are often unobtainable by other methods. For example, generally achievable surface roughness for sand casting is 400 to 800 μin. (10 to 20 μm), for forging 200 to 400 μin. (5 to 10 μm), and for die casting 80 to 200 μin. (2 to 5 μm). Ultraprecision machining (i.e., super-finishing, lapping, diamond turning) can produce a surface finish of 0.4 μin (0.01 μm) or better. The achievable dimensional accuracy in casting is 1 to 3% (ratio of tolerance to dimension) depending on the thermal expansion coefficient and in metal forming it is 0.05 to 0.30% depending on the elastic stiffness, but in machining the achievable tolerance can be 0.001%.
- It is flexible. The shape of the final machined product is programmed and therefore many different parts can be made on the same machine tool and just about any arbitrary shape can be machined. In machining, the product contour is created by the path, rather than the shape, of the cutter. By contrast, casting, molding, and forming processes require dedicated tools for each product geometry, thus restricting their flexibility.
- It can be economical. Small lots and large quantities of parts can be relatively inexpensively produced if matched to the proper machining process.

The dominating physical mechanism at the tool/workpiece interface in conventional machining is either plastic deformation or controlled fracture of the workpiece. Mechanical forces are imposed on the workpiece by the application of a tool with sharp edges and higher hardness than the workpiece. However, many new materials are either harder than conventional cutting tools or cannot withstand the high cutting forces involved in traditional machining. Nontraditional manufacturing (NTM) processes can produce precision components of these hard and high-strength materials. NTM processes remove material through thermal, chemical, electrochemical, and mechanical (with high impact velocity) interactions.

Machinability is defined in terms of total tool life, power requirements, and resultant workpiece surface finish. To date, no fundamental relationship incorporates these three factors and thus machinability must be empirically determined by testing.

Process Selection

Machine tools can be grouped into two broad categories:

- Those that generate surfaces of rotation
- Those that generate flat or contoured surfaces by linear motion

Selection of equipment and machining procedures depends largely on these considerations:

- Size of workpiece
- Configuration of workpiece
- Equipment capacity (speed, feed, horsepower range)
- Dimensional accuracy
- Number of operations
- Required surface condition and product quality

For example, Figure 13.2.2 graphically indicates the various tolerance levels that can be typically achieved for common machining unit processes as a function of the size of the workpiece. Such data can help in identifying candidate unit processes that are capable of meeting product requirements.

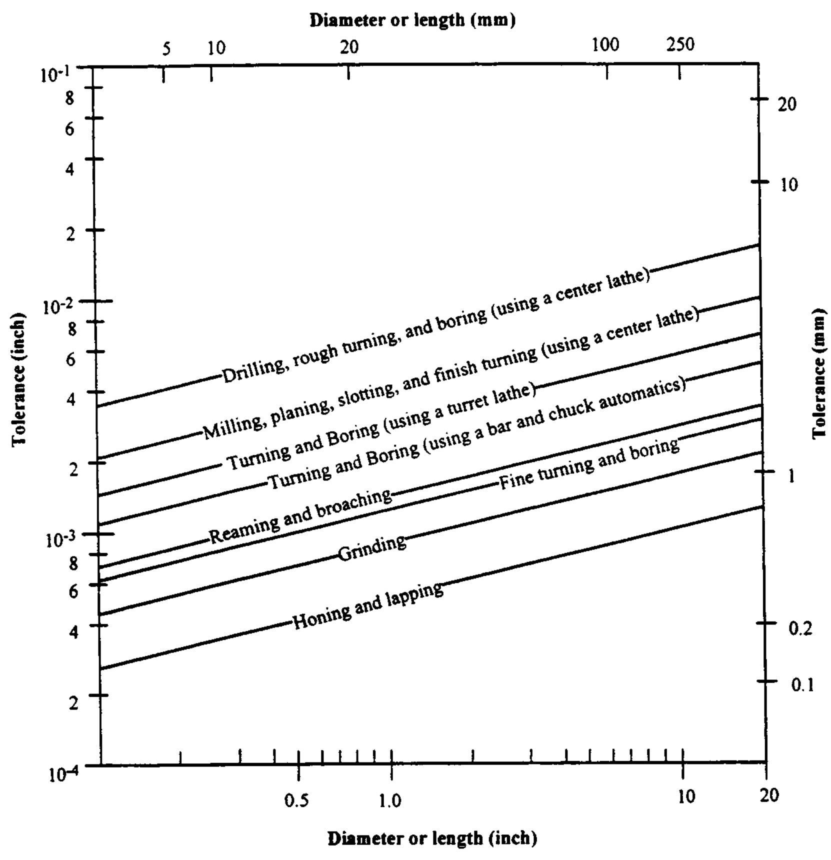

FIGURE 13.2.2 Tolerance vs. dimensional data for machining processes. (From NRC, *Unit Manufacturing Processes*, National Academy Press, Washington, D.C., 1995, 168. With permission.)

Traditional Machining

Steven Y. Liang

Traditional machining processes remove material from a workpiece through plastic deformation. The process requires direct mechanical contact between the tool and workpiece and uses relative motion between the tool and the workpiece to develop the shear forces necessary to form machining chips. The tool must be harder than the workpiece to avoid excessive tool wear. The unit processes described here are a representative sample of the types most likely to be encountered. The reference list at the end of the section should be consulted for more detailed information on the unit processes discussed below, plus those that are not included here.

Process Kinematics in Traditional Machining. In all traditional machining processes, the surface is created by providing suitable relative motion between the cutting tool and the workpiece. There are two basic components of relative motion: primary motion and feed motion. Primary motion is the main motion provided by a machine tool to cause relative motion between the tool and workpiece. The feed motion, or the secondary motion, is a motion that, when added to the primary motion, leads to a repeated or continuous chip removal. It usually absorbs a small proportion of the total power required to perform a machining operation. The two motion components often take place simultaneously in orthogonal directions.

The functional definitions of turning, milling, drilling, and grinding are not distinctively different, but machining process specialists have developed terminology peculiar to a given combination of functions or machine configurations. Commonly used metal-cutting machine tools, however, can be divided into three groups depending upon the basic type of cutter used: single-point tools, multipoint tools, or abrasive grits.

Dynamic Stability and Chatter. One of the important considerations in selecting a machine tool is its vibrational stability. In metal cutting, there is a possibility for the cutter to move in and out of the workpiece at frequency and amplitude that cause excessive variations of the cutting force, resulting in poor surface quality and reduced life of the cutting tool (Welboum, 1970).

Forced vibrations during cutting are associated with periodic forces resulting from the unbalance of rotating parts, from errors of accuracy in some driving components, or simply from the intermittent engagement of workpiece with multipoint cutters. *Self-excited vibrations* occur under conditions generally associated with an increase in machining rate. These vibrations are often referred to as chatter. All types of chatter are caused by a *feedback loop* within the machine tool structure between the cutting process and the machine frame and drive system. The transfer function of the machine tool, in terms of the stiffness and damping characteristics, plays a critical role in the stability of the overall feedback system. The static stiffness of most machine tools, as measured between the cutting tool and the workpiece, tends to be around 10^5 lb-ft/in. A stiffness of 10^6 lb-ft/in. is exceptionally good, while stiffness of 10^4 lb-ft/in. is poor but perhaps acceptable for low-cost production utilizing small machine tools.

Basic Machine Tool Components. Advances in machine-tool design and fabrication philosophy are quickly eliminating the differences between machine types. Fifty years ago, most machine tools performed a single function such as drilling or turning, and operated strictly stand-alone. The addition of automatic turrets, tool-changers, and computerized numerical control (CNC) systems allowed lathes to become *turning centers* and milling machines to become *machining centers*. These multiprocess centers can perform a range of standard machining functions: turning, milling, boring, drilling, and grinding (Green, 1992).

The machine tool *frame* supports all the active and passive components of the tool — spindles, table, and controls. Factors governing the choice of frame materials are: resistance to deformation (hardness), resistance to impact and fracture (toughness), limited expansion under heat (coefficient of thermal expansion), high absorption of vibrations (damping), resistance to shop-floor environment (corrosion resistance), and low cost.

Guide ways carry the workpiece table or spindles. Each type of way consists of a *slide* moving along a track in the frame. The slide carries the workpiece table or a spindle. The oldest and simplest way is the *box way.* As a result of its large contact area, it has high stiffness, good damping characteristics, and high resistance to cutting forces and shock loads. Box slides can experience stick-slip motion as a result of the difference between dynamic and static friction coefficients in the ways. This condition introduces positioning and feed motion errors. A *linear way* also consists of a rail and a slide, but it uses a rolling-element bearing, eliminating stick-slip. Linear ways are lighter in weight and operate with less friction, so they can be positioned faster with less energy. However, they are less robust because of the limited surface contact area.

Slides are moved by hydraulics, rack-and-pinion systems, or screws. *Hydraulic pistons* are the least costly, most powerful, most difficult to maintain, and the least accurate option. Heat buildup often significantly reduces accuracy in these systems. Motor-driven *rack-and-pinion* actuators are easy to maintain and are used for large motion ranges, but they are not very accurate and require a lot of power to operate. Motor-driven screws are the most common actuation method. The screws can either be lead screws or ballscrews, with the former being less expensive and the latter more accurate. The *recirculating ballscrew* has very tight backlash; thus, it is ideal for CNC machine tools since their tool trajectories are essentially continuous. A disadvantage of the ballscrew systems is the effective stiffness due to limited contact area between the balls and the thread. (*Note:* a case study at the end of this section discusses the manufacture of precision ballscrews.)

Electric motors are the prime movers for most machine tool functions. They are made in a variety of types to serve three general machine tool needs: spindle power, slide drives, and auxiliary power. Most of them use three-phase AC power supplied at 220 or 440 V. The design challenge with machine tools and motors has been achieving high torque throughout a range of speed settings. In recent years, the operational speed of the spindle has risen significantly. For example, conventional speeds 5 years ago were approximately 1600 rpm. Today, electric motors can turn at 12,000 rpm and higher. Higher speeds cause vibration, which makes use of a mechanical transmission difficult. By virtue of improvement in motor design and control technology, it is now possible to quickly adjust motor speed and torque. Mechanical systems involving more than a three-speed transmission are becoming unnecessary for most high-speed and low-torque machines. Spindle motors are rated by horsepower, which generally ranges from 5 to 150 hp (3.7 to 112 kW) with the average approximately 50 hp (37 kW). Positioning motors are usually designated by torque, which generally ranges from 0.5 to 85 lb-ft (0.2 to 115 Nm).

The *spindle* delivers torque to the cutting tool, so its precision is essential to machine tool operation. The key factors influencing precision are bearing type and placement, lubrication, and cooling.

Cutting Tool Materials. The selection of cutting tool materials is one of the key factors in determining the effectiveness of the machining process (ASM, 1989b). During cutting, the tool usually experiences high temperatures, high stresses, rubbing friction, sudden impact, and vibrations. Therefore, the two important issues in the selection of cutting tool materials are hardness and toughness. *Hardness* is defined as the endurance to plastic deformation and wear; hardness at elevated temperatures is especially important. *Toughness* is a measure of resistance to impact and vibrations, which occur frequently in interrupted cutting operations such as milling and boring. Hardness and toughness do not generally increase together, and thus the selection of cutting tool often involves a trade-off between these two characteristics.

Cutting tool materials are continuously being improved. Carbon steels of 0.9 to 1.3% carbon and tool steels with alloying elements such as molybdenum and chromium lose hardness at temperatures above 400°F (200°C) and have largely been replaced by *high-speed steels* (HSS). HSS typically contains 18% tungsten or 8% molybdenum and smaller amounts of cobalt and chromium. HSSs retain hardness up to 1100°F (600°C) and can operate at approximately double the cutting speed with equal life. Both tool steels and HSS are tough and resistive to fracture; therefore, they are ideal for processes involving interrupted engagements and machine tools with low stiffness that are subject to vibration and chatter.

Powder metallurgy (P/M) high-speed tool steels are a recent improvement over the conventionally cast HSSs. Powder metallurgy processing produces a very fine microstructure that has a uniform distribution of hard particles. These steels are tougher and have better cutting performance than HSS. Milling cutters are becoming a significant application for these cutting tool materials.

Cast cobalt alloys, popularly known as Stellite tools, were introduced in 1915. These alloys have 38 to 53% cobalt, 30 to 33% chromium, and 10 to 20% tungsten. Though comparable in room temperature hardness to HSS tools, cast cobalt alloy tools retain their hardness to a much higher temperature, and they can be used at 25% higher cutting speeds than HSS tools.

Cemented carbides offered a four- or fivefold increase in cutting speeds over conventional HSS. They are much harder, but more brittle and less tough. The first widely used cemented carbide was tungsten carbide (WC) cemented in a ductile cobalt binder. Most carbide tools in use now are a variation of the basic WC-Co material. For instance, WC may be present as single crystals or a solid solution mixture of WC-TiC or WC-TiC-TaC. These solid solution mixtures have a greater chemical stability in the cutting of steel. In general, cemented carbides are good for continuous roughing on rigid machines, but should avoid shallow cuts, interrupted cuts, and less rigid machines because of likely chipping.

A thin layer of TiC, TiN, or Al_2O_3 can be applied to HSS or carbide substrate to improve resistance to abrasion, temperature, friction, and chemical attacks. The *coated tools* were introduced in the early 1970s and have gained wide acceptance since. Coated tools have two or three times the wear resistance of the best uncoated tools and offer a 50 to 100% increase in speed for equivalent tool life.

Ceramic tools used for machining are based on alumina (Al_2O_3) or silicon nitride (Si_3N_4). They can be used for high-speed finishing operations and for machining of difficult-to-machine advanced materials, such as superalloys (Komanduri and Samanta, 1989). The alumina-based materials contain particles of titanium carbide, zirconia, or silicon carbide whiskers to improve hardness and/or toughness. These materials are a major improvement over the older ceramic tools. Silicon nitride–based materials have excellent high-temperature mechanical properties and resistance to oxidation. These materials also have high thermal shock resistance, and thus can be used with cutting fluids to produce better surface finishes than the alumina tools.

These tools can be operated at two to three times the cutting speeds of tungsten carbide, usually require no coolant, and have about the same tool life at higher speeds as tungsten carbide does at lower speeds. However, ceramics lack toughness; therefore, interrupted cuts and intermittent application of coolants can lead to premature tool failure due to poor mechanical and thermal shock resistance.

Cermets are titanium carbide (TiC) or titanium carbonitride particles embedded in a nickel or nickel/molybdenum binder. These materials, produced by the powder metallurgy process, can be considered as a type of cemented carbide. They are somewhat more wear resistant, and thus can be used for higher cutting speeds. They also can be used for machining of ferrous materials without requiring a protective coating.

Cubic boron nitride (CBN) is the hardest material at present available except for diamond. Its cost is somewhat higher than either carbide or ceramic tools but it can cut about five times as fast as carbide and can hold hardness up to 200°C. It is chemically very stable and can be used to machine ferrous materials.

Industrial *diamonds* are now available in the form of polycrystalline compacts for the machining of metals and plastics with greatly reduced cutting force, high hardness, good thermal conductivity, small cutting-edge radius, and low friction. Recently, diamond-coated tools are becoming available that promise longer-life cutting edges. Shortcomings with diamond tools are brittleness, cost, and the tendency to interact chemically with workpiece materials that form carbides, such as carbon steel, titanium, and nickel.

Wear of Cutting Tool Materials. Cutting tools are subjected to large forces under conditions of high temperature and stress. There are many mechanisms that cause wear.

- Adhesion. The tool and chip can weld together; wear occurs as the welded joint fractures and removes part of the tool material, such as along a tool cutting edge.
- Abrasion. Small particles on the wear surface can be deformed and broken away by mechanical action due to the high localized contact stresses; these particles then abrade the cutting tool. Typically, this is the most common wear mode.
- Brittle fracture. Catastrophic failure of the tool can occur if the tool is overloaded by an excessive depth of cut and/or feed rate.
- Diffusion. Solid-state diffusion can occur between the tool and the workpiece at high temperatures and contact pressures, typically at an area on the tool tip that corresponds to the location of

maximum temperature, e.g., cemented carbide tools used to machine steel. High-speed machining results in higher chip temperatures, making this an increasingly important wear mode.

- Edge chipping.
- Electrochemical. In the presence of a cutting fluid, an electrochemical reaction can occur between the tool and the workpiece, resulting in the loss of a small amount of tool material in every chip.
- Fatigue.
- Plastic deformation.

Single-Point Cutting Tool Geometry. Figure 13.2.3 depicts the location of various angles of interest on a single-point cutting tool. The most significant angle is the *cutting-edge* angle, which directly affects the shear angle in the chip formation process, and therefore greatly influences tool force, power requirements, and temperature of the tool/workpiece interface (ASM, 1989a). The larger the positive value of the cutting-edge angle, the lower the force, but the greater the load on the cutting tool. For machining higher-strength materials, negative rake angles are used. *Back rake* usually controls the direction of chip flow and is of less importance than the side rake. Zero back rake makes the tool spiral more tightly, whereas a positive back rake stretches the spiral into a longer helix. *Side rake* angle controls the thickness of the tool behind the cutting edge. A thick tool associated with a small rake angle provides maximum strength, but the small angle produces higher cutting forces than a larger angle; the large angle requires less motor horsepower.

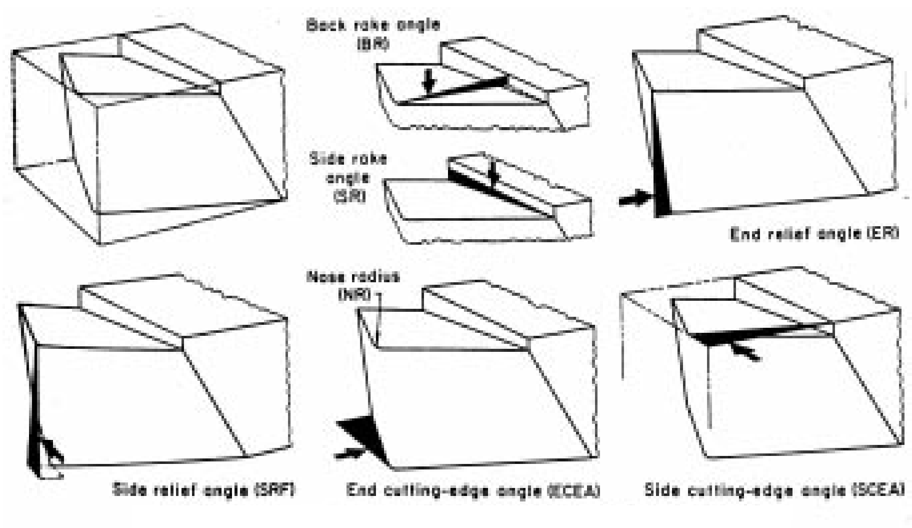

FIGURE 13.2.3 Standard nomenclature for single-point cutting tool angles. (From *ASM Handbook, Machining*, Vol. 16, 9th ed., ASM International, Metals Park, OH, 1989, 141. With permission.)

The *end relief angle* provides clearance between the tool and the finished surface of the work. Wear reduces the angle. If the angle is too small, the tool rubs on the surface of the workpiece and mars the finish. If the angle is too large, the tool may dig into the workpiece and chatter, or show weakness and fail through chipping. The *side relief angle* provides clearance between the cut surface of the work and the flank of the tool. Tool wear reduces the effective portion of the angle closest to the workpiece. If this angle is too small, the cutter rubs and heats. If the angle is too large, the cutting edge is weak and the tool may dig into the workpiece. The *end cutting-edge angle* provides clearance between the cutter and the finished surface of the work. An angle too close to zero may cause chatter with heavy feeds, but for a smooth finish the angle on light finishing cuts should be small.

Machinability. Optimum speed and feed for machining depend on workpiece material, tool material, characteristics of the cut, cutting tool configuration, rigidity of setup, tolerance, and cutting fluid. Consequently, it is not possible to recommend universally applicable speeds and feeds.

Drilling and Reaming

Description and Applications. *Drilling* is the most widely used process for making circular holes of moderate accuracy. It is often a preliminary step to other processes such as tapping, boring, or reaming. *Reaming* is used to improve the accuracy of a hole while increasing its diameter. Holes to be reamed are drilled undersize.

Key System Components. Drills are classified by the material from which they are made, method of manufacture, length, shape, number and type of helix or flute, shank, point characteristics, and size series (Table 13.2.1). Selection of drill depends on several factors (ASM, 1989b):

TABLE 13.2.1 Common Drill Types

Drill Type	Description	Application
Core	Has large clearance for chips	Roughing cuts; enlarging holes
General-purpose (jobber)	Conventional two-flute design; right-hand (standard) or left-hand helix. Available with flute modification to break up long chips	General-purpose use, wide range of sizes
Gun	Drill body has a tube for cutting fluid; drill has two cutting edges on one side and counter-balancing wear pads on the other side	Drill high-production quantities of holes without a subsequent finishing operation
High helix	Wide flutes and narrow lands to provide a large bearing surface	Soft materials, deep holes, high feed rates
Low helix	Deep flutes facilitate chip removal	Soft materials, shallow holes
Oil hole	Has holes through the drill body for pressurized fluid	Hard materials, deep holes, high feed rate
Screw-machine	Short length, short flutes, extremely rigid	Hard materials; nonflat surfaces
Step	Two or more drill diameters along the drill axis	Produce multiple-diameter holes, such as for drilling/countersinking
Straight flute	Flutes parallel to the drill axis minimize torquing of the workpiece	Soft materials; thin sheets

- Hardness and composition of the workpiece, with hardness being more important
- Rigidity of the tooling
- Hole dimensions
- Type of drilling machine
- Drill application — originating or enlarging holes
- Tolerances
- Cost

The most widely used drill is the general-purpose twist drill, which has many variations. The flutes on a twist drill are helical and are not designed for cutting but for removing chips from the hole. Typical twist drills are shown in Figure 13.2.4.

Machining forces during *reaming* operations are less than those of drilling, and hence reamers require less toughness than drills and often are more fragile. The reaming operation requires maximum rigidity in the machine, reamer, and workpiece.

Most *reamers* have two or more flutes, either parallel to the tool axis or in a helix, which provide teeth for cutting and grooves for chip removal. The selection of the number of flutes is critical: a reamer with too many flutes may become clogged with chips, while a reamer with too few flutes is likely to chatter (Table 13.2.2).

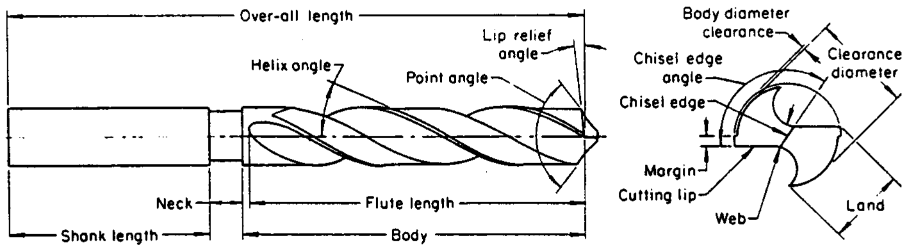

FIGURE 13.2.4 Design features of a typical straight-shank twist drill. (From *ASM Handbook, Machining,* Vol. 16, 9th ed., ASM International, Metals Park, OH, 1989, 218. With permission.)

TABLE 13.2.2 Common Reamer Types

Reamer Type	Description	Application
Adjustable	Tool holder allows adjustment of the reamer diameter to compensate for tool wear, etc.	High-rate production
End-cutting	Cutting edges are at right angles to the tool axis	Finish blind holes, correct deviations in through-holes
Floating blade	Replaceable and adjustable cutting edges to maintain tight tolerances	High-speed production (workpiece rotated, tool stationary)
Gun	Hollow shank with a cutting edge (e.g., carbide) fastened to the end and cutting fluid fed through the stem	High-speed production (workpiece rotated, tool stationary)
Shell	Two-piece assemblies, mounted on arbors, can be adjusted to compensate for wear	Used for finishing operations (workpiece rotated, tool stationary)
Spiral flute	Flutes in a helix pattern, otherwise same as straight-flute reamer	Difficult to ream materials, and holes with irregularities
Straight flute	Flutes parallel to the tool axis, typically pointed with a 45° chamfer	General-purpose, solid reamer

Machining Parameters. The optimal speed and feed for drilling depend on workpiece material, tool material, depth of hole, design of drill, rigidity of setup, tolerance, and cutting fluid. For *reaming* operations, hardness of the workpiece has the greatest effect on machinability. Other significant factors include hole diameter, hole configuration (e.g., hole having keyways or other irregularities), hole length, amount of stack removed, type of fixturing, accuracy and finish requirements, size of production run, and cost. Most reamers are more easily damaged than drills; therefore, the practice is to ream a hole at no more than two thirds of the speed at which it was drilled.

Capabilities and Process Limitations. Most drilled holes are $^1/_8$ to 1 in. (3.2 to 40 mm) in diameter. However, drills are available for holes as small as 0.001 in. (0.03 mm) (microdrilling), and special drills are available as large as 6 in. (150 mm) in diameter. The range of length-to-diameter (L/D) of holes that can be successfully drilled depends on the method of driving the drill and the straightness requirements. In the simplest form of drilling in which a rotating twist drill is fed into a fixed workpiece, best results are obtained when L/D is <3. But by using special tools, equipment, and techniques, straight holes can be drilled with L/D = 8 or somewhat greater. Nonconventional machining processes can also generate high-aspect-ratio holes in a wide variety of materials.

Reaming and boring are related operations. Hole diameter and length, amount of material to be removed, and required tolerance all influence which process would be most efficient for a given application (ASM, 1989b). Most holes reamed are within the size range of $^1/_8$ to 1 inch (3.2 to 40 mm), although larger and smaller holes have been successfully reamed. For most applications with standard reamers, the length of a hole that can be reamed to required accuracy ranges from slightly longer to much shorter than the cutting edges of the reamer, but there are many exceptions to this general rule-of-thumb. Tolerances of 0.001 to 0.003 in. (0.03 to 0.08 mm) with respect to the diameter are readily achievable in production reaming operations. Surface finish for annealed steels can be held within the range of 100 to 125 μin. (2.50 to 3.20 μm), but a surface as smooth as 40 μin. (1 μm) can be obtained under appropriate processing conditions (ASM, 1989b).

Turning and Boring

Description and Applications. *Turning* produces external cylindrical surfaces by removing material from a rotating workpiece, usually with a single-point cutting tool in a lathe. *Boring* is this same process applied for enlarging or finishing internal surfaces of revolution.

Key System Components. The basic equipment for turning is an *engine lathe* that consists of a bed, a headstock, a carriage slide, a cross slide, a tool holder mounted on the cross slide, and a source of power for rotating the workpiece (Table 13.2.3). Engine lathes are often modified to perform additional types of machining operations through the use of attachments. Most turning machines can also perform boring operations, but boring machines may not be able to perform turning operations. Sizes of lathes range from fractional horsepower to greater than 200 hp.

TABLE 13.2.3 Typical Lathes Used for Turning

Lathe	Description	Applications
Bench lathe	An engine lathe that can be placed on a workbench	Small workpieces and prototype parts
Engine lathe	Has a leadscrew that moves the slide uniformly along the bed; available with chucking or centering headstock	Chucking type allows centering and clamping for rotation, e.g., holding castings or forgings Centering type secures workpiece between pointed centers, e.g., for turning long workpieces, such as shafts
Gap-frame lathe	Modified engine lathe for turning larger diameter parts	Workpieces requiring off-center mounting or irregular protuberances
Numerically controlled lathe	Uses a computer program to control the lathe to generate the desired shape	Produces consistent parts in a CAD/CAM environment
Tracer-controlled lathe	A duplicating lathe that uses a stylus moving over a template to control the cutting tool	Manufacture of prototype parts and low-rate production

TABLE 13.2.4 Typical Machines Used for Boring

Boring Machine	Description	Applications
Bar (screw) machine	Modified turret lathe to handle bars and tubes	Parts made from bars or tubes
Engine lathe	Versatile machine; essentially same machine as used for turning	Bores one hole at a time in a single part; limitations regarding workpiece size and configuration
Horizontal boring mill	Workpiece remains stationary and tool rotates	Wide variety of parts; cost-effective for a relatively high production volume
Precision boring machine	Vertical and horizontal models	Parts requiring extreme tolerances
Special-purpose machines	Boring machine modified for specialized application	Single-purpose applications with high production rates
Turret lathe	Has rotating turret on a lathe, tooled for multiple machining operations	More versatile than engine lathe; supports high production rates
Vertical boring mill	Same basic components as a lathe	Very large, heavy, or eccentric workpieces
Vertical turret lathe	Same features as vertical boring mill; may also have a second vertical head	Flexible machine, useful in CAD/CAM environment; simultaneous multiple machining operations possible

Machines used for boring are noted for their rigidity, adaptability, and ability to maintain a high degree of accuracy (Table 13.2.4). For extremely large workpieces, weighing thousands of pounds, the boring cutting tool is rotated and the workpiece is fixed.

Machine Tool and Machining Parameters. In turning and boring operations, a single-point tool is traversed longitudinally along the axisymmetric workpiece axis parallel to the spindle. A tangential force is generated when the cutting tool engages the rotating work. This force is generally independent of the cutting speed and directly proportional to the depth of cut for a particular material, tool shape (particularly side rake angle), and feed rate. That force, when multiplied by the surface speed of the workpiece, estimates the net horsepower required to remove material. The extent to which workpiece material affects required machining power is illustrated in Figure 13.2.5. Moving the tool longitudinally requires much less power (ASM, 1989b).

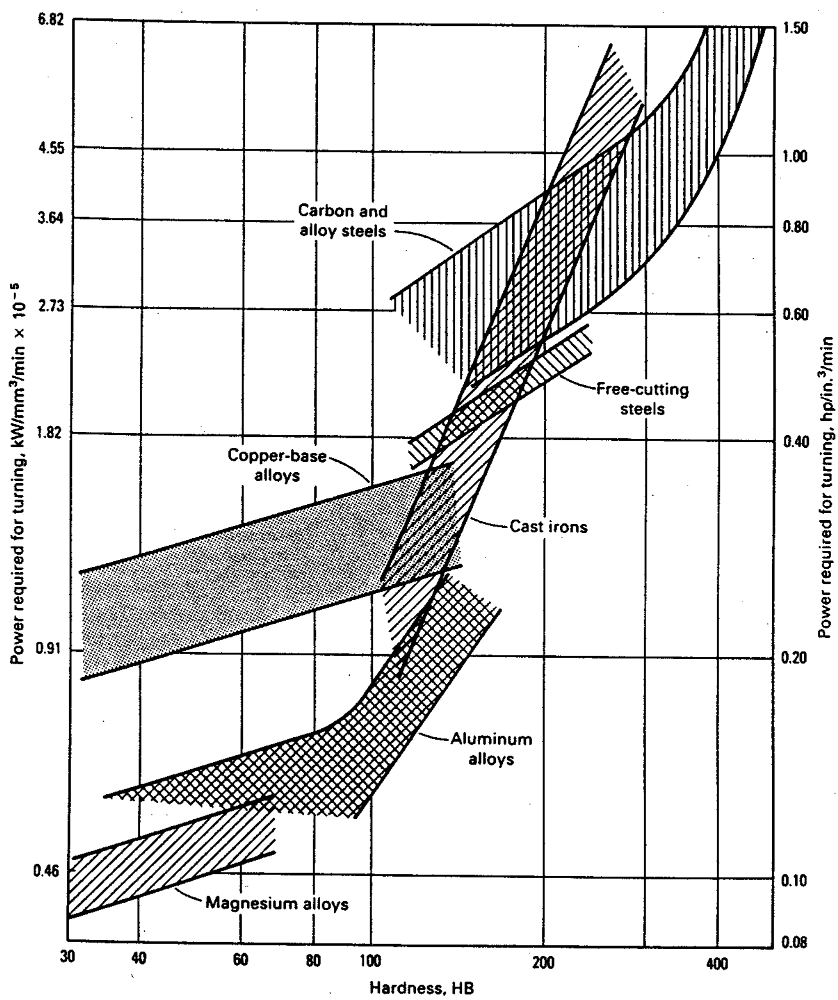

FIGURE 13.2.5 Effect of workpiece composition and hardness on power required for turning. (From *ASM Handbook, Machining,* Vol. 16, 9th ed., ASM International, Metals Park, OH, 1989, 136. With permission.)

To minimize the number of cuts required, the depth of cut should be as great as possible, which is limited by the strength of the part and the fixturing, and the power output of the machine tool. The feed rate is a function of the finish desired and the strength and rigidity of the part and machine tool.

Capabilities and Limitations. Components that range in size from those used in watches up to large steel propeller shafts more than 80 ft (24 m) long are regularly turned. Aluminum parts over 10 ft (3 m) in diameter have been turned. In practice, the weight of the workpiece per unit of volume determines the size of the workpiece that is practical to turn. Large, heavy parts can be turned in a vertical boring machine. Irregular-shaped parts, such as crankshafts, may require the use of counter-weighting to achieve dynamic balance for vibration-free turning.

For both turning and boring, the rotation speed, feed, and depth of cut determine the rate of material removal and resulting surface quality. Feed rate for most applications falls between 0.005 and 0.020 in./rev (0.13 and 0.51 mm/rev). Finishing cuts have a significantly lower feed rate (e.g., 0.001 in./rev, 0.03 mm/rev), and roughing cuts are made at a significantly higher feed rate (e.g., 0.25 in./rev, 6.35 mm/rev). Boring is not limited by the L/D ratio of a hole — this ratio can be as great as 50 if the tool bar and workpiece are adequately supported.

There are many potential sources of tolerance error in turning and boring operations. The more common errors are summarized in Tables 13.2.5 and 13.2.6.

TABLE 13.2.5 Factors Affecting Dimensional Accuracy in Turning

Quantity	Basic Cause of Inaccuracy	Corrective Actions
Diametrical roundness	Lathe spindle runout	Preload angular bearings to eliminate side and end movement
Diameter variation (taper)	Parallelism of spindle to longitudinal travel	Establish true parallelism between the axis of rotation of the workpiece and the longitudinal travel of the cutting tool
Face flatness	Lack of true normality of cross slide to axis of rotation	Precision align cross slide to axis of rotation
Length dimensions parallel to the axis of rotation	Improper positioning of longitudinal slide	Adjust positioning of longitudinal slide
Diameter accuracy	Inadequate gauging	Employ proper measurement technique

TABLE 13.2.6 Factors Affecting Dimensional Accuracy in Boring

Quantity	Basic Cause of Inaccuracy	Corrective Actions
Diameter accuracy	Inadequate gauging, or heat generated by the machining action	Employ proper measurement techniques; use cutting fluid to control temperature
Taper of the cylindrical bore	Deflection of the boring bar	Reduce the boring bar unsupported length; use a higher-stiffness material for the boring bar
Roundness, as determined by the variation in radius about a fixed axis	Finish cut is not concentric with the previous cut; or out-of-balance workpiece or holder	Begin with a semifinish cut, followed by a finish cut; carefully balance the initial setup
Concentricity of one surface with another	Too great a clamping force on the workpiece	Redesign clamping fixture, or use a precision boring machine
Squareness and parallelism of holes in relation to other features of the work	Dimensional changes in machine components	Maintain constant ambient temperature; maintain cutting fluid at constant temperature; or stabilize oil temperature

Planing and Shaping

Description and Applications. *Planing* is a widely used process for producing flat, straight surfaces on large workpieces. A variety of contour operations and slots can be generated by use of special attachments. It is often possible to machine a few parts quicker by planing than by any other method. *Shaping* is a process for machining flat and contour surfaces, including grooves and slots.

Principle of Operation. *Planers* develop cutting action from straight-line reciprocating motion between one or more single-point tools and the workpiece; the work is reciprocated longitudinally while the tools are fed sideways into the work. Planer tables are reciprocated by either mechanical or hydraulic drives, with mechanical drives predominating.

Shapers use a single-point tool that is supported by a ram which reciprocates the tool in a linear motion against the workpiece. The workpiece rests on a flat bed and the cutting tool is driven toward it in small increments by ram strokes. Shapers are available with mechanical and hydraulic drives, with mechanical drives predominating.

Machine Tool and Machining Parameters. Planing and shaping are rugged machining operations during which the workpiece is subjected to significant cutting forces. These operations require high clamping forces to secure the workpiece to the machine bed.

In general, it is advisable to plane steel with as heavy a feed and as high a speed as possible to promote good chip-formation conditions so that chip breakers are not needed. Carbide cutters allow cutting speed to be increased from 225 to 300 surface feet per minute (sfm) (70 to 90 m/min). For best results, uniform cutting speed and feed are maintained throughout the entire stoke (ASM, 1989b).

In general, speeds are related to workpiece material characteristics and associated machinability. Feeds are influenced by the workpiece machinability, but also by ram speed, depth of cut, and required dimensional accuracy and surface finish. Common practice in shaping is to make roughing cuts at as high a feed and slow a speed as practical, and make finishing cuts at a low feed rate and high speed (ASM, 1989b). For low carbon steel, a typical speed for a roughing cut is 50 sfm (15 m/min), while for a finishing cut it is 80 sfm (25 m/min) using a conventional cutting tool. Similarly for aluminum, a roughing cut of 150 sfm (45 m/min) is typically followed by a finishing cut of 200 sfm (60 m/min).

There is a practical lower bound on minimum feed rate. Feed rates that are too low will cause the tool to chatter; feed rates less than 0.005 in. (0.125 mm) are seldom used in shaping. Similarly, shallow cuts (less than 0.015 in., 0.38 mm) will cause chatter during shaping.

Key System Components. Planers are available in a wide range of sizes (Table 13.2.7). Tools are available in a variety of configurations for undercutting, slotting, and straight planing of either horizontal or vertical surfaces (ASM, 1989b).

TABLE 13.2.7 Types of Planers

Planer Type	Description	Application
Double housing	Two vertical uprights support the crossrail which in turn supports the tools	Rigid machine; restricts the width of workpiece
Open side	A single upright column supports a cantilevered crossrail; less rigid than the double-housing type	Accommodates wide workpieces which can overhang one side of the table without interfering with the planer operation

Shapers are available in a large variety of sizes (Table 13.2.8), ranging from small models with a maximum stroke length of less than 6 in. (150 mm) to large machines with a maximum stroke of 36 in. (914 mm). On each machine, the length of stroke can be varied from its maximum to slightly less than 1 in. (25 mm) for the largest machine, and to $^1/_8$ in. (3.2 mm) for the smallest machine.

TABLE 13.2.8 Types of Shapers

Shaper Type	Description	Application
Horizontal	The ram drives the tool in the horizontal direction; uses plane or universal table (rotates on three axes)	Gears, splined shafts, racks, and so on; not used for rate production
Vertical	The ram operates vertically, cutting on the downstroke	Slots, grooves, keyways; matching die sets, molds, fixtures; not used for rate production

Capabilities and Process Limitations. Planing is a precision process in which flatness can be held within 0.0005 in. (0.013 mm) total indicated runout (TIR) on workpieces up to 4 ft^2 (0.4 m^2). Although planing is most widely used for machining large areas, it is also used for machining smaller parts, although 12 in. is about the minimum distance for a planing stoke. Size of the workpiece that can be planed is limited by the capacity of the planing equipment.

Shaping is a versatile process in which setup time is short and relatively inexpensive tools can be used. Under good conditions, a shaper can machine a square surface of 18 in. on a side (0.2 m^2) to a flatness within 0.001 in. (0.025 mm); under optimum conditions this can be improved to 0.0005 in. (0.013 mm). The size of the workpiece that can be shaped is limited by the length of the stoke, which is usually about 36 in. (914 mm). Shaping should be considered for machining flat surfaces in these instances:

- Required flatness cannot be achieved by another method
- Production quantity is insufficient to justify the tooling costs of milling or broaching

Planing and shaping are interrupted cutting processes, and are comparatively inefficient means of metal removal; for example, shaping costs five times that of milling, exclusive of the tooling and setup costs.

Milling

Description and Applications. Milling is a versatile, efficient process for metal removal. It is used to generate planar and contour surfaces through the action of rotating multiple-tooth cutters. Surfaces having almost any orientation can be machined because both the workpiece and cutter can move in more than one direction at the same time.

Principle of Operation. Cutters with multiple cutting edges rotate in a spindle. The machining process is interrupted as the teeth of the milling cutter alternately engage and disengage from the workpiece.

Key System Components. Most milling is done in machines designed for milling (Table 13.2.9). Milling can also be done by any machine tool that can rigidly hold and rotate a cutter while feeding a workpiece into the cutter. Milling machines are usually classified in terms of their appearance: knee-and-column, bed-type, planar-type, and special purpose. The knee-and-column configuration is the simplest milling machine design. The workpiece is fixed to a bed on the knee and the tool spindle is mounted on a column, as depicted in Figure 13.2.6. For very large workpieces, gantry or bridge-type milling machines are used. Machines having two columns can provide greater stability to the cutting spindle(s). Special-purpose machines are modifications of the three basic models.

The usual power range for knee-and-column machines is 1 to 50 hp (0.75 to 37 kW). Bed-type machines are available in a wide range of sizes, up to 300 hp (225 kW). Planar-type machines are available from 30 to 100 hp (22 to 75 kW).

There is a wide variety of milling cutters, using the full range of cutting tool materials; there are three basic constructions (ASM, 1989b):

- Solid — Made from a single piece of HSS or carbide; cutters can be tipped with a harder material; teeth can designed for specific cutting conditions; low initial cost.
- Inserted blade — Usually made from HSS, carbide, or cast alloy; individual blades can be replaced as they wear out, saving replacement cost; ideal for close-tolerance finishing.
- Indexable insert — Cutter inserts are made from carbide, coated carbide, ceramic, or ultrahard material such as diamond; each insert has one or more cutting edges; as inserts wear, they are repositioned to expose new cutting surface or indexed to bring another cutting insert on line. These inserts, widely used in computer-controlled machines due to their performance and flexibility, can produce a rougher surface than the other tool constructions and require somewhat higher cutting forces to remove metal.

TABLE 13.2.9 Types of Milling Machines

Milling Machine	Description	Application
Knee-and-column	Six basic components: • Base — the primary support • Column — houses spindle and drive • Overarm — provides support for the arbor-mounted cutting tools • Knee — supports the table, saddle, and workpiece; provides vertical movement • Saddle — provides 1° of horizontal motion • Table — directly supports the workpiece and provides a second degree of horizontal motion	Widely used for low production milling; provides three-axis movement; primary drawback is lack of rigidity due to the number of joints
Bed-type	Table and saddle mounted on a bed in fixed vertical position; vertical motion obtained by movement of the spindle carrier; available with horizontal or vertical spindle	Very rigid machine; permits deep machining cuts and close dimensional control
Planar-type (adjustable rail)	Can accommodate almost any type of spindle for driving cutters and boring bars; utilizes several milling heads	Use for mass-production milling; can perform simultaneous milling and boring operations
Special purpose	Many possible configurations involving major modifications or combinations of the basic types of milling machines; adapted for automated control	Optimized for high-volume production; includes profilers and machining centers; these machines are capable of performing multiple simultaneous machining operations

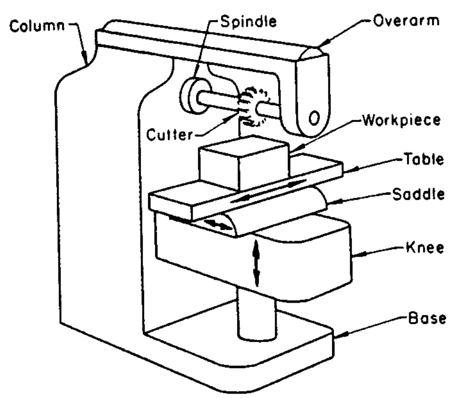

FIGURE 13.2.6 Principal components of a plain knee-and-column milling machine with a horizontal spindle. (From *ASM Handbook, Machining,* Vol. 16, 9th ed., ASM International, Metals Park, OH, 1989, 304. With permission.)

Milling cutters are also described by the location of the cutting edges, as described in Table 13.2.10. Several cutters are depicted in Figure 13.2.7.

Machine Tool and Machining Parameters. The angular relationships of the cutting edge greatly influence cutting efficiency, analogous to single-point cutting tools. A milling cutter should have enough teeth to ensure uninterrupted contact with the workpiece, yet not so many so as to provide too little space between the teeth to make chip removal difficult.

Milling speed varies greatly depending on workpiece material composition, speed, feed, tool material, tool design, and cutting fluid. Speeds as low as 20 sfm (6.1 m/min) are employed for milling low machinability alloys, while speeds as high as 20,000 sfm (6100 m/min) have been reported for milling aluminum (ASM, 1989b). If the setup is sufficiently rigid, carbide or carbide-tipped cutters can be operated three to ten times faster than HSS cutters; top speed is usually constrained by onset of tool chatter.

TABLE 13.2.10 Types of Milling Cutters

Milling Cutters	Description	Application
Peripheral mills	Cutting is primarily done by teeth on the periphery of the cutting tool; mounted on an arbor having its axis parallel to the machined surface	Removing metal from simple flat surfaces; milling contoured surfaces and surfaces having two or more angles or complex forms
Face mills	Machining action is accomplished by the bevel cutting edge located along the circumference of the mill; driven by a spindle on an axis perpendicular to the surfaced being milled	Can be more efficient at removing material than peripheral milling; very rigid tool setup possible; can achieve tight tolerances
End mills	Incorporate cutting edges on both the face and the periphery; can be used for face cuts and periphery cuts	Allow multiple operations without changing cutters; cutters can have difficulty in maintaining dimensional accuracy due to long unsupported length
Special mills	Can be almost any design	Optimized for a particular task

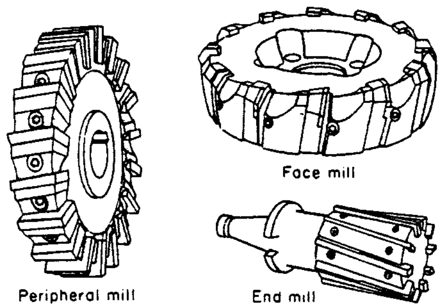

FIGURE 13.2.7 Three typical milling cutters. (From *ASM Handbook, Machining,* Vol. 16, 9th ed., ASM International, Metals Park, OH, 1989, 311. With permission.)

For highest efficiency in removing metal while minimizing chatter conditions, the feed per tooth should be as high as possible. The optimum feed rate is influenced by a number of factors (ASM, 1989b): type of cutter, number of teeth on the cutter, cutter material, workpiece machinability, depth of cut, width of cut, speed, rigidity of the setup, and machine power. The surface finish obtainable by milling can be quite good. A finish of 125 µin. (3.2 µm) can be readily achieved under normal circumstances with HSS mills, and finishes of 63 µin. (1.6 µm) are common if carbide tools are used. With careful selection of cutters and stringent control of process conditions, a finish of 10 min (0.25 mm) can be produced.

Process Limitations. The initial cost of a milling machine is considerably greater than that of a planar or a shaper that can machine workpieces of similar size to similar finishes. Milling tools usually cost up to 50 times as much as tools for planers and shapers, and the setup time is usually longer. However, milling is far more efficient in removing material, and milling machines are commonly highly automated. Therefore, milling is preferred for production operations.

Grinding is often preferred to milling when the amount of metal to be removed is small and the dimensional accuracy and surface finish are critical. Milling and grinding are frequently used in combination.

Broaching

Description and Applications. Broaching is a precision machining process. It is very efficient since both roughing cuts and finishing cuts are made during a single pass of the broach tool to produce a smooth surface, and further finishing is usually not necessary. Consequently, close tolerances can be readily achieved at a reasonable cost for high rates of production.

Broaches are expensive multitoothed cutting tools. Thus, the process is usually employed for low or high production when broaching is the only practical method to produce the required dimensional tolerance and surface quality. An example of the latter case is the dovetail slots in jet engine turbine disks.

Principle of Operation. Broaching is a machining process similar to planing. A broach is essentially a tapered bar into which teeth are cut, with the finishing teeth engaging last on the end with the larger diameter. A single broach has teeth for rough cutting, semifinishing, and finishing. Broaching involves pushing or pulling a broach in a single pass through a hole or across a surface. As the broach moves along the workpiece, cutting is gradual as each successive tooth engages the workpiece, removing a small amount of material. Overall machining forces are much greater than that of other machining methods, and consequently broaching is considered to be the most severe of all machining operations.

Key System Components. Broaching machines are categorized as horizontal or vertical, depending on the direction of broach travel. Industry usage is almost evenly divided between these two categories. The selection of machine type depends heavily on the configuration of the workpiece and available space in the factory, considering both floor space and vertical clearance requirements.

Broaches can be categorized by the method through which they are actuated (push or pull), by type of cut (internal or external), and by the construction of the broach body. Figure 13.2.8 depicts typical internal and external broaching operations. Table 13.2.11 describes broaches according to their construction.

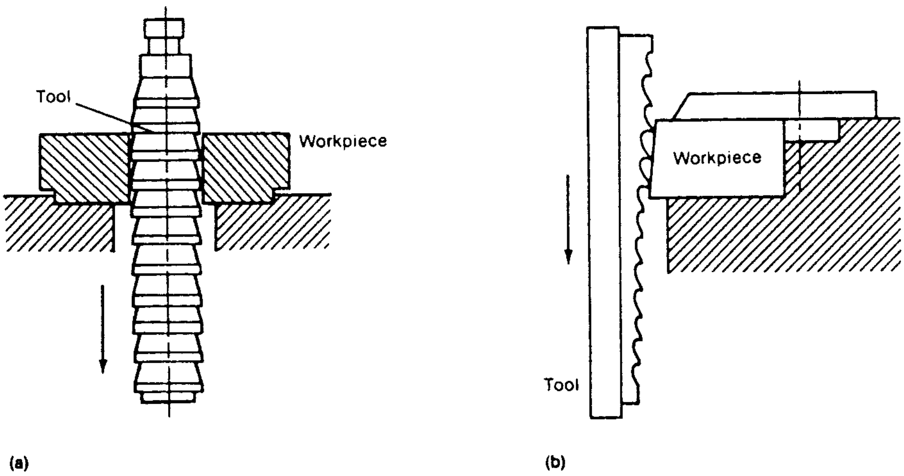

FIGURE 13.2.8 Internal (a) and external (b) broaching. (From *ASM Handbook, Machining,* vol. 16, 9th ed., ASM International, Metals Park, OH, 1989, 195. With permission.)

TABLE 13.2.11 Types of Broaches

Broaches	Description	Application
Solid	One-piece broach produced from tapered bar stock; repair of broken teeth is difficult	Parts which require high dimensional accuracy and concentricity
Shell	Multipiece broach consisting of a main body, an arbor section over which a removable shell fits, and a removable shell containing the cutting edges; worn or damaged sections can be replaced	Internal and selected external broaching; sacrifices some accuracy and concentricity as the tool is not as stiff as a solid broach
Insert-type	Effectively a tool holder with inserts to perform the actual cutting; inserts are typically made from HSS or carbides; worn or damaged inserts can be readily replaced	Broaching large, flat surfaces

Machine Tool and Machining Parameters. Length and depth of cut have the most influence on determining the required broaching tool length. For internal cutting operations, as the cut length increases, more chip storage capacity must be provided between the cutting edges for the same amount of tooth advance. Cutting fluids are useful in preventing the work metal from adhering to the broach, and thus result in higher-quality surface finishes and increased broach life.

The primary consideration in the selection of optimum broaching speed is the trade-off between speed and wear rate. In general, steels are broached at 10 to 30 sfm (3 to 9 m/min); the harder the steel, the slower the broach speed (ASM, 1989b).

Capabilities and Process Limitations. Broaching can maintain tight tolerances during long production runs since metal-cutting operations are distributed among the different roughing and finishing teeth. Also, broach teeth can be repeatedly sharpened, allowing cutting efficiency and accuracy to be maintained.

Broaching is an extremely fast, precise machining operation. It is applicable to many workpiece materials over a wide range of machinability, can be accomplished in seconds, is readily automated, and can easily be done manually. For example, for low-carbon steels, tolerances of 0.002 in. (0.05 mm) can be readily attained with a surface finish of 60 μin. (1.55 μm); if desired, tighter tolerances and surface finishes of 30 μin. (0.8 μm) are possible without much additional effort. For difficult-to-machine superalloys, tolerances of 0.001 in. (0.025 mm) and surface finishes of 30 μin. (0.8 μm) are commonly achieved in production (ASM, 1989b).

Broaching is rarely used for removing large amounts of material since the power required would be excessive. It is almost always more effective to use another machining method to remove the bulk of material and use broaching for finishing.

Since a broach moves forward in a straight line, all surface elements along the broach line must be parallel to the direction of travel. Consequently, the entire surface of a tapered hole cannot be broached. Also, cutting is done sequentially with the finishing teeth engaging last. Therefore, a blind hole can be broached only if a sufficiently long recess is provided to permit full travel of the broach.

The direction of travel of the broach cannot realistically be changed during a broaching stroke, except for rotating the tool. Thus, surfaces having compound curves cannot be broached in a single operation. On external surfaces, it is impossible to broach to a shoulder that is perpendicular to the direction of broach movement.

Grinding

Trevor D. Howes, John Webster, and Ioan Marinescu

Description and Applications. Grinding, or abrasive machining, refers to processes for removing material in the form of small chips by the mechanical action of irregularly shaped abrasive grains that are held in place by a bonding material on a moving wheel or abrasive belt (Green, 1992). In surface-finishing operations (e.g., lapping and honing) these grains are suspended in a slurry and then are embedded in a roll-on or reference surface to form the cutting tool. Although the methods of abrasion may vary, grinding and surface-finishing processes are used in manufacturing when the accuracy of workpiece dimensions and surface requirements are stringent and the material is too hard for conventional machining.

Grinding is also used in cutoff work and cleaning of rough surfaces, and some methods offer high material-removal rates suitable for shaping, an area in which milling traditionally has been used.

Grinding is applied mainly in metalworking because abrasive grains are harder than any metal and can shape the toughest of alloys. In addition, grinding wheels are available for machining plastics, glass, ceramics, and stone. Conventional precision metal and ceramic components and ultraprecision electronic and optical components are produced using grinding.

Mechanics of Grinding. Three types of energy are involved in grinding (Andrew et al., 1985). *Rubbing energy* is expended when the grains (cutting edges) of the grinding wheel wear down. As they wear, they cut less and produce increasing friction, which consumes power but removes less material. *Plowing energy* is used when the abrasive does not remove all of the material but rather plows some of it aside plastically, leaving a groove behind. *Chip-formation energy* is consumed in removing material from the workpiece as the sharp abrasive grain cuts away the material (or chip) and pushes it ahead until the chip leaves the wheel.

The grinding wheel experiences *attritious wear* as the abrasive grains develop wear flats from rubbing on the workpiece, or when grains break free from the bond material. Attritious wear gives rise to rubbing energy resulting from friction, and thus can lead to thermal damage as power consumption increases without an increase in material removal rate. The wheel can wear through *fracture*, predominating at relatively high in-feed rates. In this case, pieces of the abrasive grain break free and expose a new, sharp surface.

Materials can be classified as either easy to grind or difficult to grind. For *easy-to-grind materials,* most of the power consumption becomes invested in chip formation; thus, rubbing and plowing energy are minimal. *Difficult-to-grind materials* involve considerable rubbing and plowing energy since the force required to remove chips is comparatively high.

Types of Grinding. In *surface grinding,* the grinding wheel traverses back and forth across the workpiece. Grinding can take place by using either the periphery or side face of the wheel. The table holding the part may also reciprocate. Surface grinding is done most commonly on flat surfaces and surfaces with shapes formed of parallel lines, e.g., slots.

Creep-feed grinding is a form of surface grinding in which the wheel feeds into the workpiece at a low rate (0.4 to 40 in./min, 10 to 1000 mm/min) while grinding at a large depth of cut (0.04 to 0.4 in., 1 to 10 mm, or deeper). A large amount of material can be removed with one pass of the wheel, compared with conventional surface grinding in which the wheel makes many quick passes over the workpiece at slight depths of cut. This process is limited by the large amount of heat generated at the grinding arc which can result in thermal damage (grinding “burn”). Application of coolant is critical in creep-feed operations. CBN wheels, with their good heat transfer property, can also reduce the severity of burn (King and Hahn, 1986).

Cylindrical grinding produces round workpieces, such as bearing rings, although some machines can also grind tapered parts. The workpiece is mounted to a spindle and rotates as the wheel grinds it. The workpiece spindle has it own drive motor so that the speed of rotation can be selected. Both inner surfaces (internal cylindrical grinding) and outer (external cylindrical grinding) can be worked, although usually the same machine cannot do both.

There are three variants for *external grinding*:

- *Plain grinding* — the wheel carriage is brought to the workpiece on its spindle and in-feeds until the desired dimensions are reached
- *Traverse grinding* — the rotating workpiece is mounted on a table that reciprocates under the wheel; the grinding wheel is stationary except for its downward feed into the workpiece
- *Plunge grinding* — the table with the rotating workpiece is locked while the wheel moves into the workpiece until the desired dimensions are attained

Centerless grinding is a form of cylindrical grinding. In this method workpieces are not held in a centering chuck but instead rotate freely between a support, regulating wheel, and the grinding wheel. The force of the rotating grinding wheel holds the workpiece against the support. The supports are usually stationary and so a flow of lubricant is required to reduce friction between workpiece and support. An example of centerless grinding is shown in Figure 13.2.9.

Abrasive belt machines use a flexible fabric coated with an abrasive stretched between two rollers, one of which is driven by a motor. Usually, the abrasive coating is aluminum oxide for steels and bronzes and silicon carbide for hard or brittle materials. In the metal industries, common use of such machines is for dry grinding of metal burrs and flash and polishing of surfaces. However, some fabrics permit use

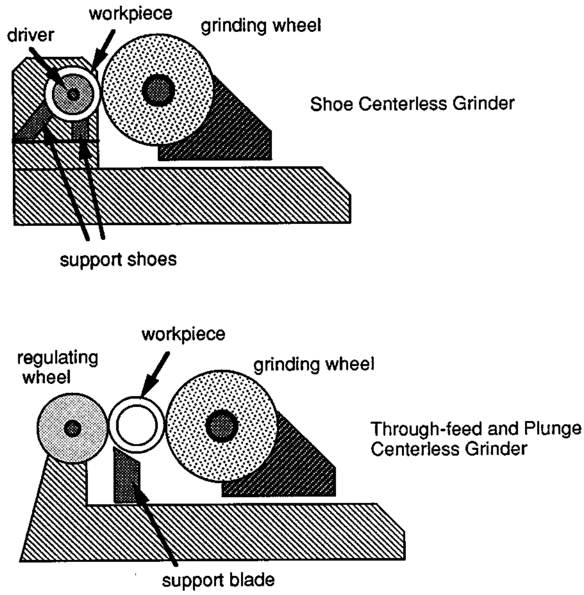

FIGURE 13.2.9 Through-feed and plunge centerless grinding. (Courtesy of T. D. Howes, John Webster, and I. Marinescu.)

of grinding fluids to enhance chip removal and provide cooling and lubrication, which results in better cutting action and longer belt life.

Honing, lapping, and polishing use abrasives to improve the accuracy of the form of a workpiece or the surface finish beyond the capabilities of grinding.

- *Honing* is a low-surface-speed operation, usually performed on an internal, cylindrical surface but possible also on external ones. Stock is removed by the shearing action of abrasive grains: a simultaneous rotary and reciprocating motion of fixed abrasive in the form of a stone or stick. Finishes range from under 1 to 50 μin. (0.025 to 1.3 μm). The development of CBN has revolutionized the honing process because this material easily outperforms conventional abrasives such as aluminum oxide, lasting up to 100 times longer.
- *Superfinishing,* like honing, uses fixed abrasives in the form of a stone. Unlike honing, which has a helical motion inside a bore, superfinishing uses high-speed, axial reciprocation combined with slow rotation of the outside diameter of the cylindrical component being processed. The geometry produced by a previous operation generally is not improved.
- *Lapping* is a fine-finishing, abrasive machining process used to obtain superior finish and dimensional accuracy. Lapping is unlike other finishing processes because some of the abrasive is loose rather than bonded. In general, lapping occurs when abrasive grains in a liquid vehicle (called a slurry) are guided across a workpiece by means of a rotating plate.
- *Polishing* uses free abrasive, as in lapping, but requires a soft support unlike the relatively hard support used in lapping. The total depth of cut during polishing can be as little as nanometers where chemical interactions will play a stronger role than mechanical or physical interactions. When the depth of cut is greater than 1 μin. (0.025 μm), the interactions are usually of a mechanical nature. Many industrial components, especially electronic and optical, required highly polished surfaces.

Key System Components. A *grinding wheel* consists of thousands of small, hard grains of abrasive material held on the surface in a matrix of bond material (Table 13.2.12). The bond material is matched to the characteristics of the grain to retain the grain sufficiently to maximize its use before shedding it.

TABLE 13.2.12 Common Grinding Wheel Abrasives

Abrasive	Characteristic	Grinding Application
Aluminum oxide	Friable	Steel: soft or hardened, plain or alloyed
Seeded gel aluminum oxide	More friable and expensive than aluminum oxide	Steel: soft or hardened, plain or alloyed; use at higher stock removal rates
CBN	Tough; increased life at higher speeds; high accuracy and finish	Hardened steels; tough superalloys
Synthetic diamond	Hardest of all abrasives; can be friable; seldom need dressing/truing	Grinding hardened tool steels, cemented carbides, ceramics, and glass; cutting and slicing of silicon and germanium wafers
Silicon carbide	Friable	Cast iron; nonferrous metals; nonmetallics

The structure of the wheel formed by specific types of grains and bonds determines its characteristics. The grains are spaced apart depending on the cutting required. Widely spaced grains (open structure) cut aggressively, which is useful for hard materials or high rates of material removal, but which tends to produce coarse finishes. Closely packed grains (dense structure) make fine and precise cuts for finish grinding.

Grain spacing is also important for temporary storage of chips of material removed from the workpiece. An open structure is best for storing chips between the grains, which are then released after wheel rotation moves the grains away from the workpiece. An open structure also permits more coolant to enter the spaces to dissipate heat.

The bonding material is important to grinding performance. This material is weaker than the cutting grains so that ideally, during grinding, the grain is shed from the wheel surface when it becomes dull, exposing new sharp grains. For instance, wheels with friable abrasives that fracture to expose new, sharp grains must retain the grains longer to maximize the use of the abrasive; these wheels use stronger bonding materials. The four types of bond material are vitrified, resinoid, rubber, and metal. Table 13.2.13 shows their properties and uses.

TABLE 13.2.13 Characteristics of Grinding Wheel Bonds

Bond Type	Characteristic	Application
Rubber	Relatively flexible	Wet cut-off wheels; high-finish work; regulating wheels for centerless grinding
Resinoid (thermoplastic)	Relatively flexible	Rough grinding; portable grinders
Vitrified (glasslike)	Endure high temperatures; resist chemical effects of coolants; sensitive to impacts	Most widely used of all bonding materials
Metal	Electrodeposited nickel or sintered metal powder often used to bond CBN and diamond abrasives; has long wheel life; electrically conductive	Aggressive cutting operations such as creep-feed and deep grinding; electrically conductive grinding (e.g., electrochemical methods)

Grinding wheels must be resharpened on occasion. *Dressing*, not always required, sharpens the grains before grinding. *Truing* operations ensure the wheel conforms to the required cutting shape and will rotate concentrically to its spindle.

Because grinding wheels have relatively high mass and high operating speed, they must be *precisely balanced*. Imbalance causes vibrations that reduce the quality of the workpiece, hasten the wear of the spindle and bearings of the machine, affect other devices mounted on the grinder, and possibly transmit vibration from the grinder through the shop floor to other machines. Mounting of the wheel on the spindle and subsequent wheel wear can degrade the balance of the rotating system. Wheels are balanced by moving counterweights on balancing flanges. Some machines have an automatic balancer that shifts internal counterbalance masses.

Coolants are usually sprayed on the grinding zone to cool the wheel and workpiece, lubricate the surface to reduce grinding power, and flush away the chips. Excessive heat can damage both the wheel and workpiece by inducing undesirable physical changes in materials, such as metallurgical phase changes or residual stresses, or softening of the bond material in the grinding wheel. Coolant application is especially important in creep-feed grinding where the wheel-to-workpiece contact arc is long, heat generation is high, and the chips produced and abrasive lost from the wheel must be flushed away.

Selecting the type of *coolant system* depends on many factors, including the grinding wheel speed, material removal rate, depth of cut, and wheel/workpiece materials. The type of fluids used in this system requires consideration of both physical and environmental issues. Use of oil fluids can favor the formation of preferred residual stress patterns and better surface finish, and these oils can be recycled for long periods. However, oils present health risks, potential for groundwater contamination, and fire risks (especially with high-sparking superalloys). Water-based fluids offer far fewer environmental problems. Disadvantages of water-based fluids lie in their limited life expectancy of 3 to 12 months. Also, the relatively low viscosity of a water-based fluid at high velocity promotes a dispersed jet which reduces cooling capacity.

Capabilities and Process Limitations. Surface grinding can be a cheaper, faster, and more precise method than milling and planing operations. For profiled shapes, the grinding wheel can be dressed with less cost and inconvenience than changing milling setups for different parts. Grinding can be used as a high-stock-removal process; for example, creep-feed grinding has a depth of cut more typical of milling operations (0.1 in., 2.54 mm, and deeper). Creep-feed grinding is used for machining materials that are too difficult to work by other machining methods.

High-speed grinding can be extremely efficient. CBN abrasive allows high rates of material removal because CBN transfers heat away from the grinding zone due to its relatively high thermal conductivity, and CBN does not react with steel.

Considerable effort has been expended on modeling and testing the thermal limitations of grinding (Malkin, 1989). Nearly all models depend on a fundamental model which depends on sliding contact theory. All models confirm the following guidelines for grinding with conventional abrasives when burn is a limitation: decrease wheel speed, increase workpiece speed, use softer-grade wheels.

Grinding operations can be limited by two types of vibration:

- *Forced vibration.* Typical causes are out-of-balance wheels, nonuniform wheels, couplings, belts, noise from hydraulic systems, bearing noise, and forces transmitted from the floor to the grinding machine.
- *Self-excited vibration.* Typical causes are wheel wear, workpiece surface error regeneration, wheel loading, and wear flats. Typical solutions include softer grinding wheels, flexible (for dampening) wheel structures, and stiffer machine structures.

References

Altan, T., Oh, S.I., and Gegel, H. 1983. *Metal Forming — Fundamentals and Applications,* ASM International, Metals Park, OH.

Andrew, C., Howes, T.D., and Pearce, T.R.A. 1985. *Creep Feed Grinding,* Holt, Rinehart, and Winston, London.

ASM. 1989a. Turning, in *Machining, ASM Handbook,* Vol. 16, 9th ed., ASM International, Metals Park, OH, 142–149.

ASM. 1989b. *Machining, ASM Handbook,* Vol. 16, 9th ed., ASM International, Metals Park, OH.

Bakerjian, R., Ed. 1992. *Design for Manufacturability,* Vol. VI, *Tool and Manufacturing Engineers Handbook,* 4th ed., Society of Manufacturing Engineers, Dearborn, MI.

DeVries, W.R. 1991. *Analysis of Material Removal Processes,* Springer-Verlag, New York.

Green, R.E., Ed. 1992. *Machinery's Handbook,* 24th ed., Industrial Press, New York.

Kalpakjian, S. 1991. *Manufacturing Processes for Engineering Materials*, Addison-Wesley, Reading, MA.
Kalpakjian, S. 1992. *Manufacturing Engineering and Technology*, Addison-Wesley, Reading, MA.
King, R.I. and Hahn, R.S., Eds. 1986. *Handbook of Modern Grinding Technology*, Chapman and Hall, New York.
Komanduri, R. and Samanta, S.K. 1989. Ceramics, in *Machining, ASM Handbook*, Vol. 16, 9th ed., ASM International, Metals Park, OH, 98–104.
Malkin, S. 1989. *Grinding Technology — Theory and Applications of Machining with Abrasives*, John Wiley, New York.
Shaw, M.C. 1984. *Metal Cutting Principles*, Oxford Science Publications. New York.
Welboum, D. 1970. *Machine-Tool Dynamics*, Cambridge University Press, New York.

Nontraditional Machining

K. P. Rajurkar and W. M. Wang

The processes described below are representative of the types most likely to be encountered. The references listed at the end of the section contain detailed information on these processes, plus those that are not described here.

Electrical Discharge Machining (EDM)

Description and Applications. The EDM process machines hard materials into complicated shapes with accurate dimensions. EDM requires an electrically conductive workpiece. Process performance is unaffected by the hardness, toughness, and strength of the material. However, process performance is a function of the melting temperature and thermal conductivity. EDM is currently widely used in aerospace, machinery, and die and mold industries.

There are two types of EDM processes:

- Die-sinking EDM uses a preshaped tool electrode to generate an inverted image of the tool on the workpiece; commonly used to generate complex-shaped cavities and to drill holes in different geometric shapes and sizes on hard and high-strength materials.
- Wire EDM (WEDM) uses a metal wire as the tool electrode; it can generate two- or three-dimensional shapes on the workpiece for making punch dies and other mechanical parts.

Principle of Operation. EDM removes workpiece materials by harnessing thermal energy produced by pulsed spark discharges across a gap between tool and workpiece. A spark discharge generates a very small plasma channel having a high energy density and a very high temperature (up to 10,000°C) that melts and evaporates a small amount of workpiece material. The spark discharges always occur at the highest electrical potential point that moves randomly over the machining gap during machining. With continuous discrete spark discharges, the workpiece material is uniformly removed around the tool electrode. The gap size in EDM is in the range of 400 μin. to 0.02 in. (0.01 to 0.5 mm), and is determined by the pulse peak voltage, the peak discharge current, and the type of dielectric fluid.

EDM Power System. The discharge energy during EDM is provided by a direct current pulse power generator. The EDM power system can be classified into RC, LC, RLC, and transistorized types. The transistorized EDM power systems provide square waveform pulses with the pulse on-time usually ranging from 1 to 2000 msec, peak voltage ranging from 40 to 400V, and peak discharge current ranging from 0.5 to 500 A. With the RC, LC, or RLC type power system, the discharge energy comes from a capacitor that is connected in parallel with the machining gap. As a result of the low impedance of plasma channel, the discharge duration is very short (less than 5 msec), and the discharge current is very high, up to 1000 A. The peak voltage is in the same range of transistorized power systems.

The transistorized power systems are usually used in die-sinking EDM operations because of their lower tool wear. Capacitive power systems are used for small hole drilling, machining of advanced materials, and micro-EDM because of higher material removal rate and better process stability. WEDM power generator usually is a transistor-controlled capacitive power system that reduces the wire rupture risk. In this power system, the discharge frequency can be controlled by adjusting the on-time and off-time of the transistors that control the charging pulse for the capacitor connected in parallel with the machining gap.

Key System Components. The machining gap between tool and workpiece during EDM must be submerged in an electrically nonconductive *dielectric fluid*. In die-sinking EDM, kerosene is often used as a dielectric fluid because it provides lower tool wear, higher accuracy, and better surface quality. Deionized water is always used as a dielectric fluid in WEDM to provide a larger gap size and lower wire temperature in order to reduce the wire rupture risk. This fluid also serves to flush debris from the gap and thus helps maintain surface quality.

Copper and graphite are commonly used as die-sinking *EDM tool materials* because of the high electrical conductivity and high melting temperature and the ease of being fabricated into complicated shapes. The wire electrode for WEDM is usually made of copper, brass, or molybdenum in a diameter ranging from 0.01 to 0.5 mm. Stratified copper wire coated with zinc brass with diameter of 0.25 mm is often used.

In the traditional die-*sinking EDM process*, the tool is fabricated into a required shape and mounted on a ram that moves vertically. The spark discharges can only occur under a particular gap size that determines the strength of electric field to break down the dielectric. A servo control mechanism is equipped to monitor the gap voltage and to drive the machine ram moving up or down to obtain a dischargeable gap size and maintain continuous sparking. Because the average gap voltage is approximately proportional to the gap size, the servo system controls the ram position to keep the average gap voltage as close as possible to a preset voltage, known as the servo reference voltage.

In a WED machine, the wire electrode is held vertically by two wire guides located separately above and beneath the workpiece, with the wire traveling longitudinally during machining. The workpiece is usually mounted on an *x-y* table. The trajectory of the relative movement between wire and workpiece in the *x-y* coordinate space is controlled by a CNC servo system according to a preprogrammed cutting passage. The CNC servo system also adjusts the machining gap size in real time, similar to the die-sinking EDM operation. The dielectric fluid is sprayed from above and beneath the workpiece into the machining gap with two nozzles.

The power generators in WED machines usually are transistor-controlled RC or RLC systems that provide higher machining rate and larger gap size to reduce wire rupture risks. In some WED machines, the machining gap is submerged into the dielectric fluid to avoid wire vibration to obtain a better accuracy. The upper wire guide is also controlled by the CNC system in many WED machines. During machining, the upper wire guide and the *x-y* table simultaneously move along their own preprogrammed trajectories to produce a taper and/or twist surface on the workpiece.

Machining Parameters. The polarity of tool and workpiece in EDM is determined in accordance with the machining parameters. When the discharge duration is less than 20 μsec, more material is removed on the anode than that on the cathode. However, if the discharge duration is longer than 30 μsec, the material-removal rate on the cathode is higher than that on the anode. Therefore, with a transistorized power system, if the pulse on-time is longer than 30 μsec, the tool is connected as anode and the workpiece is connected as cathode. When the on-time is less than about 20 μsec, the polarity must be reversed. With an RC, LC, or RLC power system, since the discharge duration is always shorter than 20 μsec, the reversed polarity is used.

With transistorized EDM power systems, the machining rate and surface finish are primarily influenced by the peak current. The machining rate increases with the peak current. The relationship between the machining rate and pulse on-time is nonlinear, and an optimal pulse on-time exists. Reducing peak current improves the surface finish, but decreases the machining rate.

Capabilities and Process Limitations. Die-sinking ED machines with transistorized power systems under good gap-flushing conditions can attain a material-removal rate as high as 12 mm^3/min/amp (for a steel workpiece). The wire cut EDM process can cut ferrous materials at a rate over 100 mm^2/min. A surface roughness value of 0.01 in. (0.2 mm) can be obtained with a very low discharge current. The tool wear ratio can be controlled within 1% during rough machining and semifinishing with the transistorized power generator. Dimensional tolerance of ±118 µin. (3 µm) and taper accuracy of 20 to 40 µin. (0.5 to 1 µm/mm) with both die-sinking and WEDM can be obtained.

EDM can machine materials having electrical conductivity of 10^{-2} W^{-1} cm^{-1} or higher. An average current density more than 4 A/cm^2 tends to cause substantial tool wear and unstable machining, and may lead to dielectric fire. This factor largely limits the productivity of EDM. During machining a deep cavity using die-sinking EDM under difficult flush condition, arc discharges occur, and the resultant thermal damage on workpiece substantially limits the productivity and the machined surface quality. Dielectric properties also impose additional constraints.

Electrical Chemical Machining (ECM)

Description and Applications. The ECM process uses the electrochemical anodic dissolution effect to remove workpiece material. Like die-sinking EDM, the tool electrode of ECM is preshaped according to the requirements of the workpiece. During machining, the inverted shape of the tool is gradually generated on the workpiece. ECM machines complex contours, irregular shapes, slots, and small, deep, and/or noncircular holes. Typical applications of ECM include machining nickel-based superalloy turbine blade dovetails, slots in superalloy turbine disks, engine castings, gun barrel rifles, and forging dies. ECM is also used for deburring, surface etching, and marking.

Principle of Operation. Electrolyte fluid is forced through the gap between tool and workpiece during ECM. A low-voltage and high-current DC power system supplies the machining energy to the gap. The tool electrode of ECM must be connected as cathode and the workpiece must be connected as anode. The electrochemical anodic dissolution phenomenon dissolves workpiece surface material into metal ions. The electrolyte fluid flushes the metal ions and removes heat energy generated by the deplating actions. The gap size in ECM is in the range of 0.004 to 0.04 in. (0.1 to 1 mm). ECM process performance is independent of the strength, hardness, and thermal behavior of workpiece and tool materials.

Key System Components. Copper, brass, stainless steel, and titanium are commonly used as *ECM tool electrode materials* due to their good electrical conductivity, resistance to chemical erosion, and ease of being machined into desired shapes. The geometric dimensions of the machined surface generated by ECM depend on the shape of tool and the gap size distribution.

The structure of an EC machine varies with the specific applications. An ECM must have a tool feed system to maintain the machining gap, a power system to supply the power energy, and a fluid-circulating system to supply the electrolyte and to flush the machining gap. The power system used in ECM is a DC power source with the voltage ranging from 8 to 30 V and a high current in the range of 50 to 50,000 A, depending on the specific design of the power system.

The *electrolyte* is the medium enabling the reaction of electrochemical dissolution occurring on the anode. The electrolyte can be classified into categories of aqueous and nonaqueous, organic and nonorganic, alkaline and neutral, mixed and nonmixed, and passivating and nonpassivating, and acidic. The electrolyte is selected according to the type of workpiece material, the desired accuracy, surface finish requirements, and the machining rate. Neutral salts are used in most cases. Acid electrolytes are used only for small hole drilling when the reaction products must be dissolved in the electrolyte.

Machining Parameters. ECM performance is mainly influenced by electrical parameters, electrolyte, and geometry of tool and workpiece. The electrical parameters include machining current, current density, and voltage. ECM systems with 50 to 50,000 A and 5 to 30 V DC are available, and the current density can be in the range of 10 to 500 A/cm^2. Key electrolyte parameters consist of flow rate, pressure, temperature, and concentration. Important parameters of tool and workpiece geometry are contour gradient, radii, flow path, and flow cross section. When the tool feed rate equals the machining rate, an equilibrium gap size is obtained; this is critical to maintaining shaping accuracy.

The electrolyte selection also plays an important role in ECM dimension control. In this regard, the sodium nitrate solution is preferable because the metal-removal rate at smaller gap size locations is higher than at other places. Therefore, the characteristics of current efficiency in ECM influence the uniformity of the gap size distribution. ECM accuracy has recently been shown to improve with pulsed voltage (instead of continuous voltage) and an appropriate set of pulse parameters (on-time, off-time, etc.).

Capabilities and Process Limitations. ECM is capable of machining any electrically conductive metallic material, and the process is generally unaffected by the hardness, strength, and thermal behaviors of materials. This process can be used to machine parts with low rigidity such as parts with thin walls. Machining rates of 2 to 2.5 cm^3/min/1000 A current, surface roughness of 4 to 50 μin. (0.1 to 1.2 μm), and accuracy of 400 μin. to 0.01 in. (10 to 300 μm) can be achieved with ECM. The available ECM equipment can machine a 0.04 to 80 in. (1 to 2000 mm) long workpiece. The typical energy consumption is 300 to 600 J/mm^3.

The machining rate and surface finish with ECM are much higher than that with EDM due to higher allowable current density and the molecular level of material removal. The machining accuracy, however, is substantially lower than EDM and is much more difficult to control.

The gap size distribution is influenced by many factors including the type of electrolyte, electrolyte flow rate and flow pattern, electrolyte temperature, current density of machining, etc. Therefore, the gap size distribution is not uniform in most cases and is difficult to determine analytically. The shape of machined surface by ECM will not be a perfect mirror image of the tool electrode. In order to achieve an acceptable accuracy, the tool shape must be modified by using trial-and-error method in test machining before machining of actual workpieces.

ECM cannot machine materials with electrical conductivity less than 10^3 W^{-1} cm^{-1}. ECM cannot produce very sharp corners (less than 800 μin., 0.02 mm, radius). The machining rate is limited by the electrolytic pressure (less than 5 MPa) and boiling point as well as the applied current (less than 50,000 A). The gap size which determines the final shape and accuracy is limited to 800 μin. to 0.004 in. (0.02 to 0.1 mm).

ECM generates a large amount of sludge and spent electrolyte. This waste requires significant processing before it can be safely disposed of.

Water and Abrasive Jet Machining

Description and Application. *Water jet machining* (WJM) and *abrasive water jet machining* (AWJM) are used in many applications. In the WJM process, relatively soft workpiece materials are cut by a high-velocity water jet; e.g., food, wood, paper, plastic, cloth, rubber, etc. The AWJM process uses the fine abrasive particles mixed in the water jet to machine harder workpiece materials. The AWJM is used for drilling, contour cutting, milling, and deburring operations on metal workpieces, as well as for producing cavities with controlled depths using multipass and non-through-cutting methods. The cutting path of the WJM and AWJM can be controlled by a CNC system according to a preprogrammed program.

Principle of Operation and Machine Structures. During WJM, the workpiece material is removed by the mechanical energy generated by the impact of a high-velocity water jet. In a water jet machine, a high-pressure pumping system increases the pressure of water in the pipe system, and the pressurized water is sprayed from a nozzle with a small diameter to generate a high-velocity water jet.

In the AWJM process, pressurized water is sprayed from an orifice in the nozzle body into a mixing chamber to generate a negative pressure that absorbs the abrasive particles (supplied by an abrasive feed hose) into the water jet. The water jet/abrasive grain mixture is then sprayed through a tungsten carbide nozzle. The abrasive grains in the high-velocity water jet provide small cutting edges that remove material. The relative distance between water jet nozzle and workpiece is controlled by a two- or three-dimensional CNC system. This process can be used to generate a complicated shape.

Process Parameters and Limitations. The key parameters are the water and/or abrasive flow velocity, abrasive grain size, and mixing tube (nozzle) length and diameter. The typical water flow velocity is in the range of 2000 to 3000 ft/sec (600 to 900 m/sec) as determined by the water pressure. The water pressure in WJM and AWJM is very high, up to 2.7×10^6 psi (400 MPa), and the nozzle diameter is in the range of 0.003 to 0.08 in. (0.8 to 2 mm). The abrasive flow rate is governed by the water flow rate and the mixing density, and can be controlled up to 10 g/sec. Abrasive particles are usually in the range of 60 to 150 mesh size. Increasing water jet flow velocity increases cutting depth. The taper error of cutting is determined by the traverse rate that describes the ratio between the material removal and the cutting depth and width. This parameter is influenced by the water and abrasive velocity and cutting speed. Proper selection of AWJ parameters is essential for the elimination of burrs, delaminations, and cracks.

Limitations of the process include stray cutting and surface waviness, high equipment costs, hazard from the rebounding abrasive, high noise levels, and short nozzle lifetimes due to wear and abrasion.

Ultrasonic Machining

Description and Applications. *Ultrasonic machining* (USM) is a process that uses the high velocity and alternating impact of abrasive particles on the workpiece to remove material. The abrasive particles are mixed in a slurry that fills the machining gap between the tool and workpiece. The alternating movement of abrasive particles is driven by the vibration of the frontal surface of tool at an ultrasonic frequency. The ultrasonic machining process can machine hard and brittle materials.

USM is often used for machining of cavities and drilling of holes on hard and brittle materials including hardened steels, glasses, ceramics, etc. Rotary ultrasonic machining (RUM) is a new application that uses a diamond grinding wheel as the tool for drilling, milling, and threading operations. During RUM, the tool is rotating at a high speed up to 5000 rpm and vibrating axially at ultrasonic frequency. This process is able to drill holes with diameter from 0.02 to 1.6 in. (0.5 to 40 mm) at depths up to 12 in. (300 mm). The material removal rate of 6 mm^3/sec can be obtained with the RUM process. The tolerance of ±300 μin. (±0.007 mm) can be easily achieved with both conventional and rotary ultrasonic processes.

Principle of Operation. In the USM process, the machining gap between tool and workpiece is filled with an abrasive slurry composed of an oil mixed with abrasive particles, with the frontal surface of the tool vibrating at ultrasonic frequency to provide the machining energy. The inverted shape of the tool is gradually generated on the workpiece. Material removal by the USM process is very complex. When the machining gap is small, the material may be removed as the frontal surface of the tool moves toward the workpiece, hitting an abrasive particle that impacts the workpiece surface. Material can also be removed by the impact of the abrasive particles when the machining gap is relatively large. In this case, the abrasive particles are accelerated by the pressure of slurry due to the ultrasonic vibration of the frontal surface of the tool. Also, ultrasonic-induced alternating pressure and cavitation in the slurry assist material removal.

Key System Components. The ultrasonic vibration in USM is generated by an *ultrasonic generator.* The ultrasonic generator consists of a signal generator, a transducer, and a concentrator. The signal generator produces an electrical signal whose voltage and/or current is changing at an ultrasonic frequency to drive the transducer. The frequency of the electrical signal can be adjusted in the range of 10 to 40 kHz.

The transducer converts the electrical voltage or current into the mechanical vibration. Two types of transducers are commonly used in USM: magnetostrictive and piezoelectric.

- The *magnetostrictive transducer* was extensively used prior to 1970. This transducer is constructed by surrounding a number of sheets of magnetostrictive material with a coil. When the strength of the electric current in the coil changes at an ultrasonic frequency, a mechanical ultrasonic vibration is generated in the magnetostrictive material. This transducer has a low energy conversion efficiency, usually less than 30%.

- The *piezoelectric ultrasonic transducer* is commonly used today. The geometrical dimensions of this transducer vary with the change in the applied electric field. A mechanical ultrasonic vibration is generated when the strength of the electric voltage applied across the transducer material changes at an ultrasonic frequency. This transducer has an extremely high energy conversion efficiency, up to 95%. The amplitude of the ultrasonic vibration generated directly by the transducer is very small, about 400 μin. (0.01 mm). A concentrator is used for amplifying the amplitude into a level that is acceptable for USM. The transducer is mounted on the larger end of the concentrator; the tool is mounted on the smaller end.

In the ultrasonic machine, the ultrasonic generator is held vertically on the ram that moves vertically, and the workpiece is mounted on an *x-y* table that determines the relative position between tool and workpiece. During machining, a force providing pressure between the tool and workpiece is added through the ram mechanism.

Process Parameters and Limitations. The material-removal rate during USM increases with an increase in the amplitude of ultrasonic vibration, grain size of the abrasive particles, and pressure between the tool and workpiece. The surface finish is essentially determined by the grain size for a given workpiece material; i.e., the smaller the grain size, the better surface finish. The abrasive grains used in USM are usually in the range of 100 to 900 mesh number.

The USM process is limited by the softness of the material. Workpiece materials softer than Rockwell C40 result in prohibitively long cycles. The best machining rate can be obtained on materials harder than Rockwell C60.

References

Benedict, G.F. 1987. *Nontraditional Manufacturing Processes,* Marcel Dekker, New York.

McGeough, J.A. 1988. *Advanced Methods of Machining,* Chapman and Hall, London.

Rajurakar, K.P. 1994. Nontraditional manufacturing processes, in *Handbook of Design, Manufacturing and Automation*, Dorf, R.C. and Kusiak, A., Eds., John Wiley & Sons, New York, 211–241.

Steen, W.M., 1991. *Laser Material Processing*, Springer-Verlag, New York.

Phase-Change Processes

Phase-change processes produce parts from materials originally in the liquid or vapor phase. These include processes such as metal casting and injection molding of polymers. The two most commonly used metal-casting processes are described below. The references listed at the end of the section contain detailed information on all phase-change unit processes.

Advantages and Applications of Metal Casting

Metal casting is one of the primary methods of producing bulk shapes. Very complex shapes can be cast from nearly every metal, making casting an extremely versatile process. Castings are made in sizes that range from fractions of an ounce to hundreds of tons.

The selection of the best molding and casting process for an application can be complex, and is governed by many factors which include casting size, variation in thickness of the casting sections, required mold strength, required surface finish and dimensional accuracy, production rates, environmental factors (e.g., reclamation of the sand and type of sand binder), and cost. Casting processes can be considered to fall into five categories (Kanicki, 1988):

- Conventional molding processes — green sand, shell, flaskless
- Precision molding and casting processes — investment, permanent mold, die casting
- Special molding and casting processes — vacuum molding, evaporative pattern casting, centrifugal casting

- Chemically bonded self-setting sand molding — no-bake, sodium silicate
- Innovative molding and casting processes — unbonded sand molding (Patz and Piwonka, 1988), rheocasting, squeeze casting, electroslag casting

This section discusses the most widely used conventional molding process, green sand, and the most widely used precision molding and casting process, investment casting. The reference list at the end of the section should be consulted for detailed information on all casting processes.

Casting Defects and Design Considerations

Properly designed and manufactured castings provide many advantages, and are competitive with other unit process methods. Success with casting processes, as with every process, requires design and process engineers knowledgeable regarding the advantages and limitations of casting processes so that appropriate design and process choices can be made that avoid or minimize the occurrence of defects.

Table 13.2.14 summarizes the most common defects that occur during casting and suggests design and process changes that can avoid or reduce the effect of the defects. However, the suggested mitigation strategies may introduce different casting defects or the evolution of other problems, so each change should be carefully evaluated with regard to the system as a whole.

Green Sand Casting Processes

Description and Applications. Sand-mold casting is adaptable to a very wide range of alloys, shapes, sizes, and production quantities. Hollow shapes can be produced in these castings through the use of cores. Sand-mold casting is by far the most common casting process used in industry; some estimates are that as many as 90% of industrial castings use the sand-mold casting process (O'Meara et al., 1988). Green sand molding is currently the most widely used of all sand casting methods, although dry sand methods are preferred for very large castings. "Green" sand refers to the fact that water is added to activate the clay binder. In dry sand molding, the moisture is removed prior to casting.

Green sand-mold casting involves mixing sand with a suitable clay binder (usually a bentonite clay) and other additives, and packing the sand mixture tightly around a pattern that is constructed from the part design, but the pattern is not an exact replica of the part since various dimensional allowances must be made to accommodate certain physical effects. After extracting the pattern from the sand mold, a cavity is left behind that corresponds to the shape of the pattern. Next, molten metal is poured into this cavity and solidifies into a cast replica of the desired part. After the casting cools, the sand is readily removed from the casting and may be reclaimed for further use.

Key System Components. A *mold pattern* is constructed from the casting design with suitable modifications. Depending on the complexity of the part, CAD/CAM programs are able to design patterns that require very little adjustment to achieve desired solidification control and dimensional accuracy in the resulting casting (Berry and Pehlke, 1988). The computation is complex and cannot be easily done for complex shapes. The principal adjustments that must be made to translate a part design to a mold design result from many considerations (ASM, 1988). One needs to

- Compensate for shrinkage of the sand during the drying/curing operations
- Compensate for expansion of the sand caused by the rapid introduction of the molten metal into the mold
- Compensate for contraction of the liquid metal during freezing
- Allow easy extraction of the pattern from the packed sand through a taper on the vertical sides of the pattern
- Add a gating network to allow molten metal to smoothly flow into the cavity
- Add risers (including size), as required in key locations to continue feeding molten metal into the solidifying casting
- Add provisions for core prints, as required, to anchor cores that produce internal cavities which could not be directly molded from the pattern

TABLE 13.2.14 Typical Casting Defects and Mitigation Strategy

Casting Defect	Description and Cause	Mitigation Strategy
Cold shuts	Appear as folds in the metal — occurs when two streams of cold molten metal meet and do not completely weld Possible causes: • Interruption in the pouring operation • Too slow a pouring rate • Improperly design gating	Pour as quickly as possible Design gating system to fill entire mold quickly without an interruption Preheat the mold Modify part design Avoid excessively long thin sections
Hot tears and cracks	Hot tears are cracklike defects that occur during solidification due to overstressing of the solidifying metal as thermal gradients develop Cracks occur during the cooldown of the casting after solidification is complete due to uneven contraction	Fill mold as quickly as possible Change gating system; e.g., use several smaller gates in place of one large gate Apply thermal management techniques within the mold (e.g., chills or exothermic material) to control solidification direction and rate Insulate the mold to reduce its cooling rate Modify casting design: • Avoid sharp transitions between thin and thick sections • Taper thin sections to facilitate establishment of appropriate solidification gradients • Strengthen the weak section with additional material, ribs, etc.
Inclusions	Presence of foreign material in the microstructure of the casting Typical sources: • Furnace slag • Mold and core material	Modify gating system to include a strainer core to filter out slag Avoid metal flow turbulence in the gating system that could cause erosion of the mold Improve hardness of the mold and core
Misruns	Incomplete filling of the mold cavity Causes: • Too low a pouring temperature • Too slow a pouring rate • Too low a mold temperature • High backpressure from gases combined with low mold permeability • Inadequate gating	Control mold and metal temperature Increase the pouring rate Increase the pouring pressure Modify gating system to direct metal to thinner and difficult-to-feed sections quicker
Porosity	Holes in the cast material Causes: • Dissolved or entrained gases in the liquid metal • Gas generation resulting from a reaction between molten metal and the mold material	Pour metal at lowest possible temperature Design gating system for rapid but uniform filling of the mold, providing an escape path for any gas that is generated Select a mold material with higher gas permeability
Microshrinkage	Liquid metal does not fill all the dendritic interstices, causing the appearance of solidification micro-shrinkage	Control direction of solidification • Design gating system to fill mold cavity so that solidification begins at the extremities and progresses toward the feed gate • Lower the mold temperature and increase the pouring temperature • Add risers, use exothermic toppings to maintain temperature longer • Control cooling rate using chills, insulators, etc. in selected portions of the mold

There are basically three types of molding methods which are categorized by the resulting hardness or density of the sand mold (O'Meara et al., 1988; Brown, 1988):

Low-density molding

- *Hand-ramming* is the oldest, slowest, and most variable method for packing sand around the pattern. It is rarely used for production, but can be employed for prototypes or very limited production runs.
- *Jolt* machines operate with the pattern mounted on a table which is attached to the top of an air piston. Typically, a flask is placed on the table with the pattern centered in the flask cavity. The flask is filled with sand. Compressed air lifts the piston, and then the air is released, allowing the entire assembly to fall with a sharp jolt. The sequence can be repeated. The sand is compacted by its own weight and is densest at the pattern plate.
- *Jolt-squeeze* molding machines employ the same pattern equipment as the jolt machines but after the jolting operation, the sand is squeezed by hydraulic pressure to improve the packing uniformity in the mold.

Medium-density molding

- *Rap-jolt machines* are improved versions of the low-pressure machines, capable of exerting higher pressure to compact the sand.
- *Sand slingers* direct sand into a mold from a rotating impeller. Sand-packing density is a function of the centrifugal velocity of the impeller. This method is particularly useful for large molds.

High-pressure molding

- *Pressure wave* methods allow the sand to gravity fill the mold. Then, the top of the mold is sealed and a high-pressure wave is created by a controlled explosion of a combustible gas or by the rapid release of air pressure.
- *Horizontal flaskless molding* utilizes a pattern carrier to support the top half (cope) and bottom half (drag) of the mold pattern. The cope and drag are spaced apart in the carrier, and the space is evacuated. The molding setup is depicted in Figure 13.2.10. Vents in the pattern cause sand to be drawn into the mold. When the mold is filled, it is tightly squeezed.

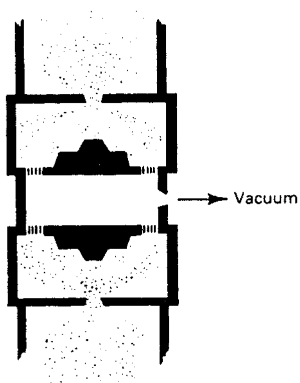

FIGURE 13.2.10 High-pressure vacuum-fill squeeze machine. (From *ASM Handbook, Machining,* Vol. 15, 9th ed., ASM International, Metals Park, OH, 1989, 343. With permission.)

The details of the *metal pouring* operation vary quite a bit depending on the metal, the material specification, the furnace type, and the foundry layout. A ladling method using gravity pressure is commonly used to transfer the molten metal, which may be alloyed, degassed, etc. immediately before it is discharged into the poring basin that feeds the gating system of the castings.

A *gating system* must allow complete fill of a mold cavity without causing flow turbulence that can entrap loose sand or slag and feed shrinkage as the liquid metal within the mold cools. The gating system should be designed to promote progressive solidification from the point most distant from the gate toward the gate.

A *riser* is a reservoir of molten metal that is attached to the casting. It feeds the voids that develop within the casting as the liquid metal cools and begins to solidify. A thin skin of frozen metal first forms in a shell around the outer part of the mold cavity immediately after the metal is poured. This rigid shell serves as a mold for the remainder of the casting. The volume lost by the shrinkage of the metal as it solidifies within this shell must be replaced from a liquid metal source, such as a riser or a feeding gate, to prevent internal porosity. Risers are subsequently removed from the casting.

Management of thermal gradients within the solidifying casting is essential to minimize shrinkage. The solidification of risers can be slowed by the use of an *exothermic material* placed on top of a riser. The heat generated by this material can allow the riser to continue to feed the casting until it is solidified. *Chills* can be used to reduce the local temperature at the mold-casting interface, and thus accelerate freezing in selected locations, thereby establishing the solidification direction. Chills are usually metal inserts strategically placed in the mold. The mold can also be *insulated* to reduce the overall cooling rate of the casting if necessary to reduce residual stresses, but the metallurgical effect must be carefully considered.

In most cases, *electric furnaces* are used for melting, and *pouring* is nonpressurized.

Influence of Process on Part Design. Properly designed gating systems and risers should produce completely solid castings. A number of design rules for the gating and risering system have been developed for casting various metals to achieve continuous feeding of the solidifying casting. These rules are embedded in computer programs that aid design of casting molds and layouts.

It may be less expensive and easier to change the design than to develop a complex thermal management system within the mold. Typical design changes to evaluate include

- Thickening thin sections that feed heavier, more remote sections
- Reducing the mass of the remote thick section
- Adding a riser to feed separately the remote thick section

Process Limitations. Each production step contributes to *dimensional variations* in a sand casting. For instance, dimensions are affected by the sand-packing density, by the process of withdrawing the pattern from the sand, by the moisture content of the sand, and by both the temperature of the molten metal and the speed with which it enters the mold cavity. The result is that sand casting is not inherently a precision process.

The surface finishes of sand castings are controllable only within rather wide limits (ASM, 1988). Normally, the maximum allowable surface roughness is specified, and any smoother surface is acceptable. For instance, casting steel in green sand can have a surface roughness varying from 500 to 2000 μin. (12 to 50 μm), and aluminum from 125 to 750 μin. (3 to 20 μm). These values can be improved in certain instances through the application of coatings on the sand mold.

Porosity cannot be prevented in all cases. Changes to the part design and postcasting processing, such as hot isostatic pressing (HIP), should be considered.

Investment Casting

Description and Applications. Investment castings are noted for their ability to reproduce extremely fine details and conform to very tight tolerances. As a result, these castings are used in critical, demanding structural applications, such as superalloy turbine airfoils and works of art.

The investment casting process employs a mold produced by enclosing an expendable pattern with a refractory slurry that is then hardened. The pattern, usually made from wax or plastic, is subsequently removed (e.g., by melting, dissolving, or burning), creating the desired mold cavity. The expendable patterns are themselves cast in a permanent pattern die. Ceramic cores can create internal passages within the casting.

Key System Components. Shell investment and solid investment processes are used in the production of investment castings (Horton, 1988). The two processes differ in the method of mold preparation, not in pattern preparation or pattern assembly. In the *shell investment process*, the pattern assembly is precoated, dipped in a coating slurry, and covered with granulated refractory until the shell is built up to desired thickness, usually less than 0.5 in. (20 μmm); the thickness depends on the casting size and weight, cluster size, and type of ceramic and binder. As thin a shell as possible is specified to maximize mold permeability. Ceramic shell molds are used for the investment casting of carbon and alloy steels, stainless steels, heat-resistant alloys, and other alloys with melting points above 1100°C (2000°F).

The *ceramic material* used in shell investments is often silica, zircon (zirconium orthosilicate), an aluminum silicate, or alumina (Horton, 1988). *Silica glass* (fused silica) is desirable because it is readily available, but it has a high coefficient of thermal expansion and an abrupt phase transition, and it cannot be used in vacuum casting because the silica decomposes at low vapor pressures, leading to severe metal–mold reaction. *Zircon* is readily available, is resistant to wetting by molten metal, and has a high refractoriness. Its use is limited to prime coats though, because it is not available in large grain sizes. *Aluminum silicates*, such as mullite, can be manufactured to a range of pellet sizes and over a range of compositions. *Alumina* is more refractory than silica or mullite and is not very reactive with many metals.

The *binders* most often employed in shell investments are colloidal silica, ethyl silicate, and sodium silicate. *Colloidal silica* is an excellent general-purpose binder, and is the most widely used binder. Its primary disadvantage is that it is slow to dry. *Ethyl silicate* produces a bond between the refractory material that is very similar to that of colloidal silica. It dries much faster, but poses a fire hazard and is more expensive. Liquid *sodium silicate* forms a strong, glassy bond. The material is inexpensive but its refractoriness is poor, and the bond deteriorates in the presence of steam used to remove the wax pattern.

In the *solid investment process*, the pattern assembly is placed in a flask, which is filled with a refractory mold slurry. This slurry hardens in air, forming a solid mass in which the pattern assembly is encased. The types of bonding materials and refractories differ depending on the pouring temperature of the metal. For nonferrous alloys, pouring temperature is usually below 2000°F (1100°C). In these cases, *alpha gypsum* is commonly used as both the refractory and the binder, with other refractories such as silica added to improve mold permeability. The process is primarily used for making dental and jewelry castings.

For extremely limited production and for the development of production process parameters, investment mold patterns can be directly machined from an expendable material, such a plastic. For production-level investment casting, however, the patterns are produced by injecting wax or plastic into *permanent pattern molding dies*. The dimensional tolerance of these permanent dies must be closely controlled.

A mixture of paraffin and microcrystalline *wax* is widely used for making investment casting patterns (Horton, 1988). Waxes are strong, stiff, and provide adequate dimensional control during pattern making. They are easy to remove with pressurized saturated steam or elevated temperatures. *Plastic* patterns have several advantages compared with wax: higher strength, less subject to handling damage, can withstand automatic ejection from the pattern mold, and can reproduce thinner sections, finer definition, sharper corners, and better surface finish. But a major disadvantage is that certain plastics, such as polystyrene, expand during burnout and can crack ceramic shell molds.

Some investment castings require complex internal cavities (e.g., holes and air passages). For complicated shapes, the pattern-die *cores* are used to form portions of the pattern that cannot be withdrawn after the pattern is made. The cores must subsequently be dissolved or etched out from the casting.

As much *gating* as possible is included in the wax patterns. This allows use of standard methods of joining patterns together so that a number of investment castings can be produced during one pouring operation.

Investment casting is done in air and in vacuum. *Gravity pouring* fills the pouring basin from a ladle or directly from a furnace. This method requires low equipment investment, but highly skilled operators. In *pressure pouring* the molds are filled from the furnace with an assist from a pressurized gas to fill rapidly thin sections. With *vacuum-assisted pouring,* a vacuum pump evacuates air from a mold ahead of the stream of molten metal to minimize flow resistance. *Centrifugal casting* uses a spinning mold assembly to develop added pressure to fill the mold.

Capabilities and Process Limitations. Investment castings:

- Produce complex shapes that are difficult to make by other means
- Reproduce fine detail, high dimensional accuracy, and smooth surfaces requiring only minimal finishing
- Adapt to most metal alloys
- Allow control of metallurgical properties such as grain size and grain orientation

A tolerance of ±0.002 in. (0.05 mm) can be held on investment castings for each inch in its maximum dimension; however, a tolerance of ±0.005 in. (0.13 mm) is more typical (Horton, 1988). A surface finish of 125 min. (3 mm) can be readily achieved, and surface finishes as smooth as 30 to 40 min. (0.8 to 1.0 mm) can be produced with suitable process control.

The size and weight of castings that can be investment cast are usually limited by physical and economic considerations. Generally, the process can be applied cost-effectively to casting weighing up to 10 lb (4.5 kg); investment castings weighing 50 lb (22.5 kg) are not unusual, and castings as large as 1000 lb (450 kg) are feasible. The initial tooling costs of investment casting can be high.

References

ASM. 1988. in *Casting,* Vol. 15, *ASM Handbook,* 9th ed., ASM International, Metals Park, OH.

Berry, J.T. and Pehlke, R.D. 1988. Modeling of solidification heat transfer, in *Casting,* Vol. 15, *ASM Handbook,* 9th ed., ASM International, Metals Park, OH, 860ff.

Brown, R.B. 1988. Sand processing, in *Casting,* Vol. 15, *ASM Handbook,* 9th ed., ASM International, Metals Park, OH, 341–351.

Horton, R.A. 1988. Investment casting, in *Casting,* Vol. 15, *ASM Handbook,* 9th ed., ASM International, Metals Park, OH, 253–269.

Isayev, A.I. Ed., 1987. *Injection and Compressive Molding Fundamentals,* Marcel Dekker, New York.

Kanicki, D.P. 1988. Casting advantages, applications, and market size, in *Casting,* Vol. 15, *ASM Handbook,* 9th ed., ASM International, Metals Park, OH, 37–45.

O'Meara, P., Wile, L.E., Archibald, J.J., Smith, R.L., and Piwanka, T.S. 1988. Bonded sand molds, in *Casting,* Vol. 15, *ASM Handbook,* 9th ed., ASM International, Metals Park, OH, 222–230.

Patz, M. and Piwonka, T.S. 1988. Unbonded sand molds, in *Casting,* Vol. 15, *ASM Handbook,* 9th ed., ASM International, Metals Park, OH, 230–237.

Structure-Change Processes

Structure-change processes alter the microstructure of a workpiece. These changes can be achieved through thermal treatment involving heating and cooling (quenching) under controlled conditions, sometimes in combination with mechanical forces, in order to effect desired solid-state phase transformations. These processes include those that diffuse selected species into a surface layer to modify its composition or to create a thin layer of material that does not increase the dimensions of the workpiece.

The two structure-change processes described here are representative examples. *Normalization of steel* is a process that changes bulk properties of a workpiece. *Laser surface hardening* of steel only changes its surface properties; it does not affect the bulk properties. Even though both examples relate to ferrous materials, structure-change processes are used to impart desired properties to many material systems; e.g., age hardening of aluminum alloys. The references listed at the end of this section contain detailed information on structure-change processes.

Normalizing Steel

Description and Applications. Normalizing is a heat-treating process that results in a relatively uniform steel microstructure. Essentially all the standard carbon steels can be normalized. The resulting phases and their size/distribution depend heavily on the carbon content of the steel. Normalization treatments are performed for a variety of reasons (Ruglic, 1991):

- Refine the dendritic grain structure remaining from casting
- Eliminate severe texture (and hence anisotropic properties) that results from forging and rolling operations
- Reduce residual stresses
- Improve the response of a steel to further processing, such as machining or surface hardening
- Improve mechanical properties by precipitating desirable phases

Principle of Operation. The steel workpiece must be heated sufficiently high to transform the entire structure to austenite, a face-centered cubic phase which essentially solutionizes all the carbon (at room temperature iron exists as ferrite, a body-centered cubic phase which has a very low carbon solubility). Diffusion-controlled solid-state phase transformations require time at temperature to occur. Therefore, the workpiece must be held at temperature long enough for austenite to dissolve the carbon.

The workpiece is then cooled slowly enough to avoid trapping the carbon in a supersaturated solution as the iron transforms back from austenite to ferrite. A time–temperature-transformation curve depicts which phase transformations will occur for different cooling gradients. A typical T-T-T curve is shown in Figure 13.2.11, which depicts the difference between a normalizing cool rate and that for annealing. For the normalization treatment to be successful, the regions in the microstructure with a carbon content of 0.8% carbon precipitate fine lamellae of ferrite and iron carbide (Fe_3C) on cooling, known as pearlite. Those areas low in carbon content should precipitate ferrite grains during the initial phase of the cooling cycle, followed by pearlite precipitation. The regions high in carbon should precipitate iron carbide in the austenite grain boundaries, followed by pearlite precipitation (Ruglic, 1990).

The end result of normalization is a microstructure, and hence the mechanical properties characteristic of the composition of the steel (primarily governed by carbon content) as opposed to a microstructure that was principally shaped by its previous thermomechanical processing.

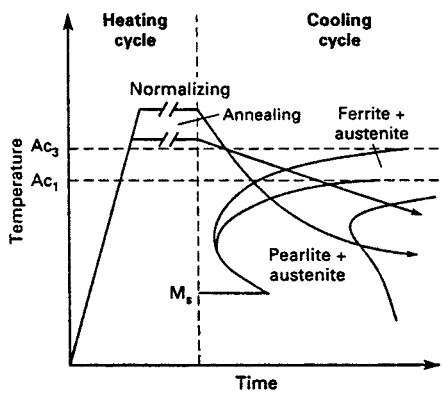

FIGURE 13.2.11 Time–temperature-transformation curve for normalizing compared to annealing. (From *ASM Handbook, Heat Treating,* Vol. 4, 10th ed., ASM International, Metals Park, OH, 1991, 35. With permission.)

Key System Components. Conventional heat-treating furnaces are used for normalization, such as batch or continuous furnaces (Ruglic, 1991). The rate of heating is not critical. The furnaces must be able to heat the component to about 100°F (55°C) into the austenitizing region — this temperature depends on the composition of the steel.

Control of the cooling rate is critical. In the usual case, a workpiece can be removed from the furnace and allowed to air cool uniformly until the diffusion-controlled phase transformations are completed.

Process Limitations. The ability to normalize the microstructure of steel is governed by thermodynamics. Phase stability and kinetics of the phase transformations are crucial. But heat transfer considerations also play a major role. For instance, it may not be possible to achieve the desired cooling gradient in the center of a thick section, and hence the properties of such sections will not be uniform.

For complex shapes or for those workpieces having a high degree of residual stresses, some distortion will occur during the normalization process. The extent of distortion can be reduced by appropriate fixturing, but it cannot be eliminated. Hence, some tolerance allowance must be made.

Laser Surface Hardening

Description and Applications. Laser surface hardening is used as an alternative to flame hardening and induction hardening ferrous materials. The rapid heating rate achievable by the laser minimizes part distortion and can impart surface hardness to low-carbon steels. The ability to locate the laser some distance from the workpiece can also be advantageous. The entire operation can be performed in air. This process is used to harden selected areas of machine components, such as gears, cylinders, bearings, and shafts.

Laser surface hardening imparts wear resistance and strength to the surface of a component without affecting its overall dimensions or changing its bulk properties. It is applied to selected areas which can be accessed by a laser beam. The process relies on rapid laser heating, followed by rapid quenching, to effect the necessary degree of hardening through phase transformation. The result is a very fine grain structure that is extremely hard. Typical case depth is a function of the composition of the ferrous material, but it will usually not exceed 0.1 in. (0.25 cm). For low- and medium-carbon steels, the case depth will range from 0.01 to 0.05 in. (0.03 to 0.13 cm), with the case depth increasing as the carbon content increases (Sandven, 1991).

Principle of Operation. An industrial laser rapidly heats a thin surface layer into the austenite phase region (austenite is a face-centered cubic allotropic phase of iron that has a high solubility for carbon). The interior of the workpiece is unaffected. When the laser beam is moved, the heated surface quickly cools. Consequently, the carbon does not have time to diffuse as the iron attempts to transform back to its ferrite (body-centered cubic) structure. The resulting microstructure is extremely hard since the trapped carbon atoms distort the iron crystal structure into a highly strained body-centered tetragonal form, known as martensite.

Key System Components. The majority of *industrial metalworking lasers* are either solid-state Nd:YAG or carbon dioxide type. Either pulsed or continuous mode can be used for surface treatment. The power output range for YAG lasers is 50 to 500 W. Carbon dioxide lasers are available in much higher power levels, up to 25 kW.

The surface to be hardened is usually *coated* to improve its ability to absorb laser radiation. A typical coating is manganese phosphate. *Paints* containing graphite, silicon, and carbon are also used. These coatings/paints can increase the absorption of laser energy to 80 to 90% (Sandven, 1991).

The output beam of the laser must be shaped and directed by an *optical system* to generate a laser spot of desired shape and size at the correct location on the workpiece surface. Reflective optical components are used since they are sturdy and easily adapted to an industrial environment.

Process Parameters. Many factors affect the end result of laser surface hardening. Important is the hardenability of the workpiece material which is affected by its composition and prior thermomechanical history. For the laser process, the key parameters are beam power density, uniformity of the beam, and processing speed. Following are some general processing guidelines (Sandven, 1991).

- The range of usable power densities for laser surface hardening is 3200 W/in.^{2} (500 W/cm^2) to 32,000 W/in.^{2} (5000 W/cm^2) with beam dwell times ranging from 0.1 to 10 sec; higher power levels would melt the surface.
- Alloys with high hardenability can be processed at low speed with low power density to produce relatively thick cases.
- Alloys with low hardenability should be processed at high speed with high power density; the result is a shallow case.
- Beam configuration can be rectangular, square, or round; uniform energy density within the beam is very important.
- Maximum achievable surface temperature is proportional to the square root of the processing speed; thus doubling the beam power density requires the processing speed to be increased by a factor of four to maintain the equivalent maximum surface temperature.
- Smaller workpieces are not as effective a heat sink as larger workpieces, and hence self-quenching may have to be assisted by quenching media.

Process Limitations. The depth of hardness that can be practically achieved is limited by the surface melting point. Because of the high beam energy density, heat flow on complex-shaped surfaces, particularly those involving sharp corners or edges, can cause unexpected surface melting. Therefore, power density and process conditions must be carefully controlled.

It may be necessary to overlap passes of the laser beam, such as at the end of a complete pass around a cylinder. As a result, some tempering of the area already hardened occurs. The slower the processing speed, the greater the degree of tempering.

References

ASM. 1995. *Heat Treater's Guide: Practices and Procedures for Irons and Steels,* 2nd ed., ASM International, Metals Park, OH.

Boyer, H.E. 1982. *Practical Heat Treating,* ASM International, Metals Park, OH.

Brooks, C.R. 1982. *Heat Treatment, Structure and Properties of Nonferrous Alloys,* ASM International, Metals Park, OH.

Ruglic, T. 1991. Normalizing of steel, in *Heat Treating, ASM Handbook,* Vol. 4, 10th ed., ASM International, Metals Park, OH, 35–42.

Sandven, O.A. 1991. Laser surface hardening, in *Heat Treating, ASM Handbook,* Vol. 4, 10th ed., ASM International, Metals Park, OH, 286–296.

Sudarshan, T.S., Ed. 1989. *Surface Modification Technologies,* Marcel Dekker, New York.

Deformation Processes

Deformation processes change the shape of an object by forcing material to flow plastically from one shape into another shape without changing mass or composition (Table 13.2.15). The initial shape is usually simple. This shape is plastically deformed between tools or dies to obtain the final desired geometry, properties, and tolerances. A sequence of such processes is generally used to progressively form material. Deformation processes, along with casting and machining, have been the backbone of modern mass production.

In addition to shape change, forming processes alter the microstructure of the workpiece and can improve material properties. Deformation processes are normally considered when (Semiatin, 1988):

- Part geometry is moderately complex
- Component properties and structural integrity are important

TABLE 13.2.15 Significant Factors in Modeling a Deformation Process

Process Component	Characteristics
Input material	Flow stress; workability; surface condition
Output material	Geometry; mechanical properties; dimensional accuracy and tolerances; surface finish
Deformation zone	Deformation mechanics; stress state; temperature
Tooling	Material and geometry; surface conditions; temperature
Tool/material interface	Friction and lubrication; heat transfer
Process equipment	Speed and production rate; power range; precision

- Sufficient production volume can amortize tooling costs

Deformation processes can be classified (Semiatin, 1988) as bulk forming processes and sheet forming processes.

Bulk forming processes (e.g., rolling, extrusion, and forging) are characterized by

- Input material form is a billet, rod, or slab
- Workpiece undergoes a significant change in cross section during forming

Sheet metal forming processes (e.g., stretching, flanging, and drawing) are characterized by

- Input material is a sheet blank
- Workpiece is deformed into a complex three-dimensional form without appreciably changing the cross section

The key to attaining desired shape and properties is controlling metal deformation (Altan et al., 1983). The direction of the metal flow, the magnitude of the deformation, the rate of deformation, and the processing temperatures greatly affect the properties of the formed part. Design of the end product and the required deformation process consists of these steps:

- Predict metal flow by analyzing kinematic relationships (e.g., shape, velocities, strain rates, and strains) between the deformed and undeformed part configurations
- Establish producibility limits
- Select the process equipment and tooling capable of operating within the producibility limits

A bulk forming process (forging) and a sheet forming process (bending) are described below as representative examples of deformation processes. The references listed at the end of the section should be consulted for further information on these unit processes.

Die Forging

Talyan Altan

Description and Applications. Forging involves the controlled plastic deformation of metals into useful shapes (ASM, 1988c). Deformation may be accomplished by means of pressure, impact blows, or a combination. In order to reduce the flow stress, forging is usually accomplished at an elevated temperature. Forging refines the microstructure of a metal and can improve its mechanical properties, especially in preferred directions. Forging can also be used for other purposes, such as to consolidate powder preforms by welding grains, eliminate porosity in castings, break up long inclusions in forgings, and demolish the dendritic structure resulting from casting (Altan, 1988a).

Forgings are generally considered when strength, reliability, fracture toughness, and fatigue resistance are important. Forgings are used in critical, high-load applications, such as connecting rods, crankshafts, transmission shafts and gears, wheel spindles, and axles. Military and commercial aircraft are major users of forgings for numerous critical items, such as bulkheads, beams, shafts, landing gear cylinders and struts, wheels, wing spars, and engine mounts. Similarly, jet engines depend on forgings for disks, blades, manifolds, and rings.

Types of Forging Processes. There are two broad categories of forging processes: closed-die forging and open-die forging. *Closed-die forging,* also known as *impression die forging,* employs precision-machined, matching die blocks to forge material to close dimensional tolerances. Large production runs are generally required to justify these expensive dies. During forging, the die cavity must be completely filled. In order to ensure this, a slight excess of material is forged. Consequently, as the dies close, the excess metal squirts out of the cavity in a thin ribbon of metal, called flashing, which must be trimmed.

Isothermal forging in heated superalloy dies minimizes the die quenching effect, preventing the rapid cooling of the workpiece in cold dies. This allows complete die fill and the achievement of close dimensional tolerances for difficult to process materials, such as superalloys.

Open-die forgings are the least refined in shape, being made with little or no tooling (Klare, 1988). These forgings are large, relatively simple shapes that are formed between simple dies in a large hydraulic press or power hammer. Examples are ship propeller shafts, rings, gun tubes, and pressure vessels. Since the workpiece is always larger than the tool, deformation is confined to a small portion of the workpiece at any point in time. The chief deformation mode is compression, accompanied by considerable spreading in the lateral directions.

Key System Components. There are two major classes of forging equipment as determined by their principle of operation: forging hammer, or drop hammer, which delivers rapid impact blows to the surface of the metal, and forging press, which subjects the metal to controlled compressive force. Each of these classes of forging equipment needs to be examined with respect to load and energy characteristics, its time-dependent characteristics, and its capability for producing parts to dimension with high accuracy.

Forging hammers generate force through a falling weight or ram (Altan, 1988b). These machines are energy restricted since the deformation results from dissipating the kinetic energy of the ram. The forging hammer is an inexpensive way to generate high forging loads. It also provides the shortest contact time under pressure, ranging from 1 to 10 msec. Hammers generally do not provide the forging accuracy obtainable in presses.

Forging presses are either mechanical or hydraulic (Altan, 1988c). *Mechanical forging presses* are stroke-restricted machines since the length of the press stroke and the available load at various positions of the stroke represent their capacity. Most mechanical presses utilize an eccentric crank to translate rotary motion into reciprocating linear motion of the press slide. The blow of the press is more like a squeeze than an impact of a hammer. Because of this, dies can be less massive and die life is longer than with a hammer. *Hydraulic presses* are load-restricted machines in which hydraulic pressure actuates a piston that squeezes the die blocks together. Full press load is available at any point during the stroke of the ram. A hydraulic press is relatively slow, resulting in longer process time; this may cause undesirable heat loss and die deterioration.

Design Considerations. Preform design is the most difficult and critical step in forging design. Proper preform design assures defect-free flow, complete die fill, and minimum flash loss. Although metal flow consists only of two basic types, extrusion (flow parallel to die motion) and upsetting (flow perpendicular to the direction of die motion), in most forgings both types of flow occur simultaneously, leading to a very complex flow field. An important step in understanding metal flow is to identify the neutral surfaces. Metal flows away from the neutral surface in a direction perpendicular to the die motion. Ideally, flow in the finishing step should be lateral toward the die cavity without additional shear at the die–workpiece interface. This type of flow minimizes forging load and die wear. A milestone in metalworking is the use of CAD in establishing the proper design for preforming and finishing dies in closed-die forging (Gegel and Malas, 1988). Figure 13.2.12 illustrates the relationships between forging process variables and those of a forging press that must be understood in order to estimate process performance for a hot-forging operation.

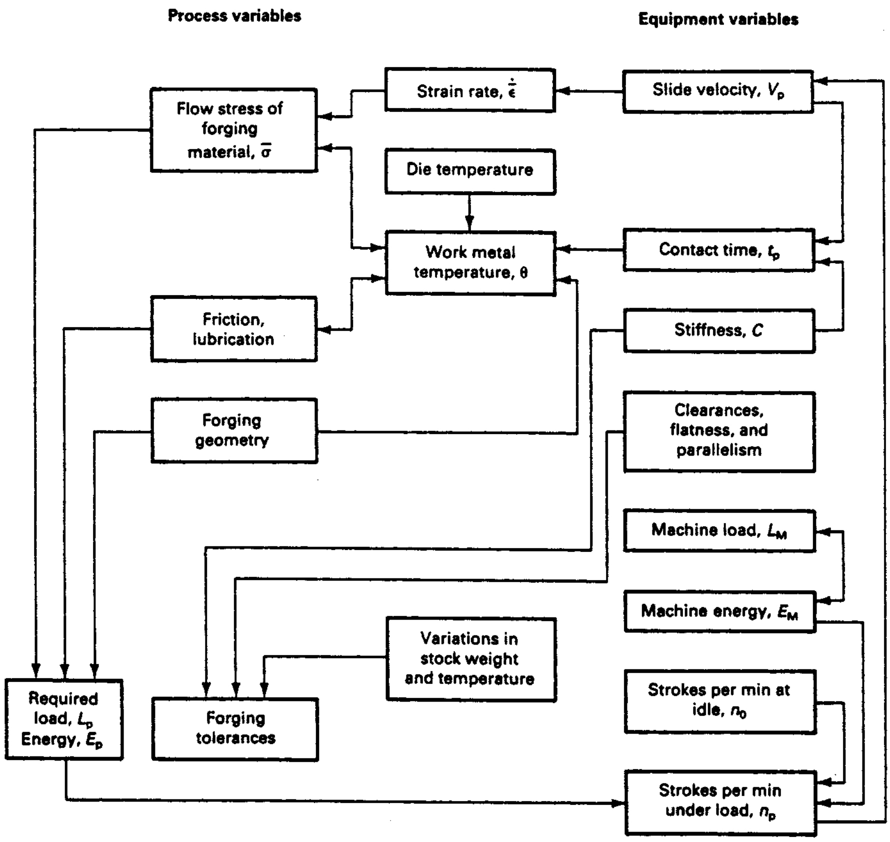

FIGURE 13.2.12 Relationships between process and machine variables in hot-forging processes conducted in presses. (From *ASM Handbook, Forming and Forging,* Vol. 14, 9th ed., ASM International, Metals Park, OH, 1988, 36. With permission.)

Designing a mechanical component that is to be made by forging together with the optimum geometry for the forging dies requires analysis of many factors (Altan, 1988a), including

- Design rules
- Workpiece material specification, and its critical temperatures
- Flow stress of the material at the process conditions (e.g., temperature, temperature gradient, strain rate, total strain)
- Workpiece volume and weight
- Frictional conditions in the die
- Flash dimensions
- Number of preforming steps and their configuration (flow field for the material)
- Load and energy requirements for each forging operation
- Equipment capability

In closed-die forging, it is particularly difficult to produce parts with sharp fillets, wide thin webs, and high ribs. Moreover, dies must be tapered to facilitate removal of the finished piece; draft allowance is approximately 5° for steel forgings (see Table 13.2.16).

TABLE 13.2.16 Typical Forging Defects and Mitigation Strategies

Defect	Description and Cause	Mitigation Strategy
Surface cracking	Fine cracks in the surface of the forging Possible causes: • Excessive working of the surface at too low a temperature • Brittle or low melting phases in the grain boundaries • Cracking at the die parting line	Increase the amount of preheating of the forging billet and forging die Change material specification Change furnace atmosphere to avoid diffusion of unwanted elements Increase the flash thickness Relocate the die parting line to a less critical location Stress relieve the forging prior to flash removal
Cold shut	Appears as a fold; occurs when two surfaces of metal fold against each other without welding Possible causes: • Poor metal flow in the die • Excessive chilling during forging • Poor die lubrication	Redesign forging die and/or forging preform to improve plastic flow of the metal during the forging operation Redesign the forging to avoid areas which are difficult to fill Increase the amount of preheating of the forging billet and forging die Improve die lubrication
Underfill	Incomplete forging in which all details are not produced Possible causes: • Debris residue in die • Scale on forging billet • Billet too small to completely fill die	Clean die thoroughly Completely descale the billet Redesign preform
Internal cracks	Cracks not visible from the surface, but detected during inspection and/or exposed during metal removal Possible causes: • Scale embedded in the internal structure of the forging • High residual tensile stresses	Completely descale the billet For open-die forgings, use concave dies Redesign for closed-die forging Use a hydraulic press and heated dies to avoid formation of excessive tensile stresses during forging

Press-Brake Forming

Description and Applications. Press-brake forming is a process used for bending sheet metal; the workpiece is placed over an open die and then pressed into the die by a punch that is actuated by a ram known as a press brake (ASM, 1988b). The main advantages of press brakes are versatility, ease and speed with which new setups can be made, and low tooling costs.

Press-brake forming is widely used for producing shapes from ferrous and nonferrous metal sheet and plate. Although sheet or plate 0.250 in. (10 mm) thick or less is commonly formed, metals up to 1 in. (25 mm) thick are regularly formed in a press brake. The length of a sheet is limited only by the size of the press brake. Forming can be done at room or elevated temperature. Low-carbon steels, high-strength low-alloy steels, stainless steels, aluminum alloys, and copper alloys are commonly formed in a press brake. Press-brake forming is applicable to any metal that can be formed by other methods, such as press forming and roll forming.

Press-brake forming is considered for bending sheet metal parts when the production quantities are small, dimensional control is not critical, or the parts are relatively long. In contrast, press forming would be considered when production quantities are large, tolerances are tight, or parts are relatively small. Contour roll forming would be another option for high-rate production applications (ASM, 1988c).

Principle of Operation. Bending is a method of forming sheet metal by stressing a material beyond its yield strength while remaining below its ultimate strength so that cracking is avoided. In press-brake forming the tooling and setup are relatively simple. A workpiece is placed over a die, typically having a V-shape. The bend angle is determined by the distance the workpiece is pressed into the die by the punch. The width of the die opening (top of the V) affects the force needed to bend the workpiece. The minimum width is determined by the thickness of the workpiece and the radius of the punch nose.

Key System Components. A *press brake* is basically a slow-speed punch press that has a long, relatively narrow bed and a ram mounted between end housings. Rams are mechanically or hydraulically actuated. Figure 13.2.13 depicts a typical setup for press-brake forming.

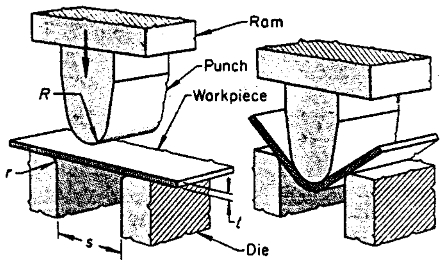

FIGURE 13.2.13 Typical setup for press-brake forming. (From *ASM Handbook, Forming and Forging,* Vol. 14, 9th ed., ASM International, Metals Park, OH, 1988, 533. With permission.)

V-bending dies and their corresponding punches are the tools most commonly used in press-brake forming. The width of the die opening is usually a minimum of eight times the sheet thickness.

Process Parameters. Capacities of commercial press brakes range from 8 to 2500 tons. Required capacity is governed by the size and bending characteristics of the work metal and by the type of bend to be made (ASM, 1988b).

The nose radius of the punch should not be less than the work-metal thickness for bending low-carbon steel, and must be increased as the formability of the workpiece material decreases. The radius of the V-bending die must be greater than the nose radius of the punch by an amount at least equal to the workpiece thickness to allow for the bottoming of the punch in the V.

It is preferable to orient a bend so that it is made across the rolling direction rather than parallel to it. Sharper bends can be made across the rolling direction without increasing the probability of cracking the material. If bends must be made in two or more directions, the workpiece should be oriented on the sheet layout such that none of the bends will be parallel to the rolling direction.

Springback after press-brake bending is considered only when close dimensional control is needed. It can be readily compensated for by overbending. Factors that affect springback include the mechanical properties of the work material, the ratio of the bend radius to stock thickness, the angle of bend, the method of bending, and the amount of compression in the bend zone. A greater amount of overbending is needed to correct for springback on small bend angles than on large bend angles.

Capabilities. The generally accepted tolerance for dimensions resulting from bending of metal sheet in the press brake is ±0.016 in. (±0.4 mm) up to and including 0.125 in. (3 mm) thickness (ASM, 1988b). For heavier gauges, the tolerance must be increased accordingly. Achievable tolerances are influenced by the part design, stock tolerances, sheet metal blank preparation, the condition of the machine and its tooling, and operator skill.

Design Factors. In press-brake forming, as in other forming processes, the metal on the inside portion of the bend is compressed or shrunk, and the metal on the outside portion is stretched. This results in a strain gradient across the thickness of the workpiece in the area of the bend with tensile strain on the outside and compressive strain inside. These residual strains (and resulting stresses) can lead to distortion of the part under loading conditions, heating, or cooling.

The formability of metals decreases as the yield strength approaches the ultimate strength. In press-brake forming, as the yield strength of the work metal increases, power requirements and springback problems also increase, and the degree of bending that is practical decreases.

There are several factors which will make it difficult to establish or maintain accurate placement of a bend line in a press brake. Corrective action may require a design change or change in processing sequence.

- Bends or holes that are located in close proximity to the required bend line can cause the position of the bend line to wander.
- Notches and cutouts located directly on the bend line make it difficult to maintain an accurate bend location.
- Offset bends will shift location unless the distance between bends in the offset is at least six times the thickness of the workpiece material.

If multiple bends must be made on a workpiece, it may not be possible to avoid a bend that is parallel to the rolling direction. Depending on the degree of texture in the sheet and the anisotropy of the material, a change to a higher-strength material may be necessary to achieve the desired geometry.

References

Altan, T. 1988a. Selection of forging equipment, in *Forming and Forging, ASM Handbook,* Vol. 14, 9th ed., ASM International, Metals Park, OH, 36–42.

Altan, T. 1988b. Hammers and presses for forging, in *Forming and Forging, ASM Handbook,* Vol. 14, 9th ed., ASM International, Metals Park, OH, 25–35.

Altan, T., Oh, S.I., and Gegel, H.L. 1983. *Metal Forming: Fundamentals and Applications,* ASM International, Metals Park, OH.

ASM. 1988a. Closed-die forging in hammers and presses, in *Forming and Forging, ASM Handbook,* Vol. 14, 9th ed., ASM International, Metals Park, OH, 75–80.

ASM. 1988b. Press-brake forming, in *Forming and Forging, ASM Handbook,* Vol. 14, 9th ed., ASM International, Metals Park, OH, 533–545.

ASM. 1988c. *Forming and Forging, ASM Handbook,* Vol. 14, 9th ed., ASM International, Metals Park, OH.

Gegel, H.L. and Malas, J.C. 1988. Introduction to computer-aided process design for bulk forming, in *Forming and Forging, ASM Handbook,* Vol. 14, 9th ed., ASM International, Metals Park, OH, 407ff.

Klare, A.K. 1988. Open-die forging, in *Forming and Forging, ASM Handbook,* Vol. 14, 9th ed., ASM International, Metals Park, OH, 61–74.

Kobayashi, S., Oh, S., and Altan, T. 1989. *Metalforming and Finite Element Methods,* Oxford Press, New York.

Lascoe, O.D. 1988. *Handbook of Fabrication Processes,* ASM International, Metals Park, OH.

Muccio, E.A. 1991. *Plastic Part Technology,* ASM International, Metals Park, OH.

Semiatin, S.L. 1988. Introduction to forming and forging processes, in *Forming and Forging, ASM Handbook,* Vol. 14, 9th ed., ASM International, Metals Park, OH, 17–21.

Consolidation Processes

Consolidation processes fuse smaller objects such as particles, filaments, or solid sections into a single solid part or component to achieve desired geometry, structure, and/or property. These processes use either mechanical, chemical, or thermal energy to bond the objects. Interaction between the material and the energy that produces the consolidation is a key feature of the process.

Consolidation processes are employed throughout manufacturing, from the initial production of the raw materials to final assembly. One group of processes involves the production of parts from powders of metals, ceramics, or composite mixtures. The resultant consolidated products are typically semifinished and require further processing. For instance, the consolidation of powders produces bar, rod, wire, plate, or sheet for upstream processes.

The consolidation of net shape composite structures (i.e., require minimal finishing work) is an increasingly important area. The design of the structural geometry, selection of material, and choice of consolidation processes all act together to provide the required level of performance. There are two types of matrix materials used: themosetting and thermoplastic. The consolidation process of each of these types of resins is different. A unit process described in this section addresses the consolidation of composites using polymeric thermosetting resins.

An important family of consolidation processes includes welding and joining processes used to permanently assemble subcomponents. Historically, welding and joining processes are developed empirically and quickly evaluated for benefit in manufacturing applications, driven by the promise of significant potential benefits. The need for welding and joining is substantial since only monolithic parts can be made without joining. The ideal joint would be indistinguishable from the base material and inexpensive to produce (Eagar, 1993). However, experience indicates that no universal joining process exists that can entirely satisfy the wide range of application needs, and thus design engineers must select the most appropriate joining methods that meet requirements. Shielded metal-arc welding, the most widely-used welding process, is described in this section.

The unit processes described here are a representative sample of the types most likely to be encountered. The references listed at the end of the section should be consulted for detailed information on these unit processes.

Polymer Composite Consolidation

Weiping Wang, Alan Ridilla, and Matthew Buczek

A composite material consists of two or more discrete materials whose combination results in enhanced properties. In its simplest form, it consists of a reinforcement phase, usually of high modulus and strength, surrounded by a matrix phase. The properties of the reinforcement, its arrangement, and volume fraction typically define the principal mechanical properties of a composite material. The matrix keeps the fibers in the correct orientation and transfers loads to the fibers.

Continuous-fiber-reinforced materials offer the highest specific strengths and moduli among engineering materials. For example, a carbon fiber/epoxy structural part in tensile loading has only about 20% of the weight of a steel structure of equal stiffness. Composite parts can integrate component piece parts, such as molded-in rib stiffeners, without the need for subsequent assembly operations and fasteners.

Description and Applications. There are many types of polymer composites in use. Composities are usually identified by their fiber material and matrix material. Fiber materials are necessarily strong, stable materials that can be processed into fiber formats. Typcial fibers are glass, graphite/carbon, aramid, and boron. *Glass* fibers represent the largest volume usage since they have excellent properties and are low cost. *Graphite/carbon* fibers are widely used for advanced composite applications in which stiffness and high performance are critical. These fibers are also expensive. *Aramid* fibers, a type of polyamide, have proved useful in applications where its performance in axial tension can be exploited without incurring too severe a penalty by the material's poor performance under compressive loading. *Boron* filaments have high strength and high modulus, but most applications requiring high performance, such as military aircraft structures, are now using carbon fibers.

Low-cost *polyester* resins are the most widely used matrix material for the general composite industry; the majority of these application use glass as the fiber. Many applications are found in the chemical process, construction, and marine industry. The most widely used polymers are *epoxy* resins, which are used with carbon, aramid, and boron fibers for many advanced applications, such as in aircraft structure and rocket motor fuel tanks. *Polyimides* are polymer resins that provide more temperature performance than is possible from the epoxy-based material, and these are used in advanced aircraft structures and jet engine components where heating of the structure will occur.

Principle of Operation. Each of the constituent materials in advanced composites acts synergistically to provide aggregate properties that are superior to the materials individually. The functional effectiveness of composites is principally due to the anisotropy of the materials and the laminate concept, where materials are bonded together in multiple layers. This allows the properties to be tailored to the applied load so that the structure can be theoretically more efficient than if isotropic materials were used. The reinforcements come in a variety of formats. Unidirectional tapes with all fibers along a common axis, woven fabrics constructed with fibers along both axes in the *x-y* plane, and multidimensional architectures with reinforcements in more than one axial direction are just a few of the available formats.

Consolidation in composites can be considered to occur at two levels: the fibers are infiltrated with the matrix to form a lamina or ply, and the individual laminae are consolidated together to form the final structure. In the prepreg process, these two levels are distinctly separated, since the fiber/matrix consolidation process forms the prepreg, which is then laid up to form the laminate or final component. In other processes, such as resin transfer molding, fiber/matrix infiltration and the consolidation of the final part are done in a single stage. Single-stage consolidation processes are attractive because they eliminate the additional cost associated with prepreg production; however, two-stage consolidation processes have major advantages that often outweigh the benefits of single-stage consolidation. These include flexibility in part geometry, high fiber content, excellent fiber wet-out, and better control of fiber volume fraction distribution. Because of these advantages, prepreg processing is firmly entrenched in high-value products, such as aerospace applications, in spite of its high cost.

Fabrication Methods. Typical steps in manufacturing continuous-fiber composites involve *preform fabrication* and consolidation/curing (Advani, 1994). Preform fabrication creates the structure by positioning material close to the final part shape (Table 13.2.17). The material comes in either the *dry fiber* (without resin) form or with resin included, called prepreg. Dry fibers are used in filament winding, weaving, braiding, and pultrusion. The resin can be introduced in the operation or downstream molding. *Prepreg,* at a higher material cost, eliminates the step of resin addition and provides the adhesion to hold the material together. Consolidation/curing involves compacting the preform to remove entrapped air, volatiles, and excess resins while developing the structural properties by increasing the polymer chain length and cross-linking.

Process Parameters. Thermosetting polymeric materials will not soften and flow upon reheating after polymerization because of the formation of a cross-linked polymer network. Hence, thermosetting polymer matrices must be cured *in situ* with the fibers to form the composite structure. The goals of a successful cure are good consolidation with low porosity and high conversion of initial monomeric constituents to polymer.

The challenge of the cure process is to manage the interactions of temperature distribution, degree of cure, laminate thickness, and void content by manipulating the applied temperature, pressure (or displacement), and vacuum. Temperature must be controlled so that resin temperature stays within limits. Both the duration and the magnitude of pressure application are important as excessive resin flow results in a resin-starved laminate. Similarly, pressure application too soon in the process can entrap volatiles in the material. Materials can also exhibit lot-to-lot variability. The problem is further complicated when processing a complex-shaped part or multiple parts of different geometry.

Composite cure processes are typically performed in an autoclave or in a heated press. An *autoclave* is essentially a heated pressure vessel. Nitrogen gas is normally used for pressurization. The temperature, pressure, and vacuum are controlled vs. time to effect the cure. Recent advancements in *intelligent processing* use sensors to determine the state of cure in real time, and make appropriate control adjustments to optimize the cure cycle (NRC, 1995). Figure 13.2.14, which plots data from an actual implementation, depicts the potential that intelligent processing has in reducing autoclave processing time and cost.

Press molding uses a high-pressure press and matched metal tools to form a part. The main components of a press-molding system include the tools, the ram, and the heated platens. Again, release materials must be applied to the tools to avoid laminate adhesion. Advantages of press molding include improved surface finish (since both sides are tooled) and the elimination of the vacuum bagging systems. However, the pressure inside the mold is not necessarily uniform and volatiles are not easily removed. Hence,

TABLE 13.2.17 Methods of Composite Preform Fabrication

Method	Description	Application
Weaving	Process of interlacing yarns to form a stable fabric construction that is flexible Less frequent interlacing results in better composite strength	Closely conforms to surfaces with compound curvature
Braiding	Intertwines parallel strands of fiber A tubular braid consists of two sets of yarn which are intertwined in "maypole dance" fashion; produced with varying diameter or circumferential size	Sporting equipment Good torsional stability for composite shafts and couplings Geometric versatility and manufacturing simplicity
Pultrusion	Reinforcing fibers pulled from a series of creels through a resin impregnating tank; preformed to the shape of the profile to be produced; enters a heated die and is cured to final shape	Produce constant cross-section pieces at high production rates
Filament winding	Pulling roving (unlisted bundles of fibers) over a mandrel by rotating the mandrel about a spindle axis; can cure on the mandrel	Cylindrical parts Can be low cost
Tube rolling	Material cut from prepreg tapes and laid on a flat surface; plies of different orientations joined together; a cylindrical mandrel is rolled on the material; curing is typically done on the mandrel	Low-cost method for making tubular structures or tapered tubes, e.g., golf shafts
Manual layup	Plies of different fiber orientations are cut from flat sheets of the prepreg material and laid up on a tool; bagging materials are applied and sealed to the tool for subsequent consolidation Flexible in producing complicated features at low start-up/tooling cost; building a curved, variable-thickness part can be complicated	Most common method used in the aerospace industry
Automatic tape layup	Computer-controlled machine tools or robots with a material delivery system • Automated tape layup machine uses prepreg tape, typically 3 to 12 in. wide, suitable for large surface of gentle contours • Fiber placement employs multiple tows for more complex surfaces	Provides lower cost and variability over manual layup, but requires higher capital investment

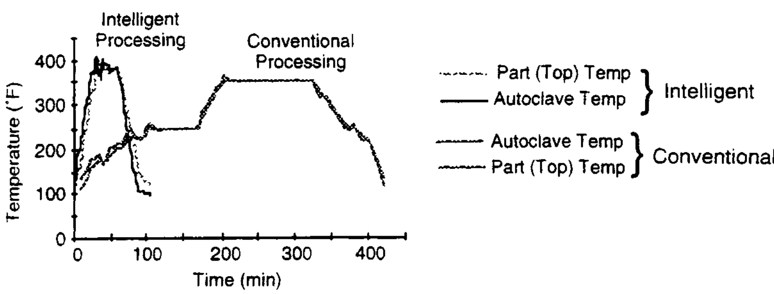

FIGURE 13.2.14 Intelligent processing of composites. (From NRC, *Expanding the Vision of Sensor Materials,* NMAB-470, National Academy Press, Washington, D.C., 1995, 39. With permission.)

press molding is generally suitable for composite systems which do not generate a significant amount of volatiles.

Shielded-Metal Arc Welding

Description and Applications. Bonding is achieved in fusion welding by interposing a liquid of substantially similar composition as the base metal between the surfaces to be joined (Eagar, 1993). The need for welding and joining is substantial since only monolithic parts can be made without joining. Traditional welding processes have unique advantages, which make them the processes of choice for a large number of applications. For example, in the fabrication of heavy structures, arc welding will dominate other assembly processes because of the inherent flexibility and economy of welding.

Principle of Operation. In the majority of arc-welding methods, the workpiece is made part of the electric welding circuit, which has as its power source a welding generator or transformer. To start a weld, an arc is struck by touching the workpiece with the tip of the electrode. The welder guides the electrode by hand in welding a joint, and controls its direction and traveling speed. The welder maintains arc voltage by controlling arc length (the distance between the end of the electrode and the work surface). Because an electric arc is one of the hottest sources of heat, melting occurs instantaneously as the arc touches the metal. Arc welding is a highly popular process because of its flexibility and relatively low cost (ASM, 1993).

In shielded-metal arc welding, an arc is struck between the workpiece and a covered (or coated) metal electrode. Filler metal is provided by the consumable electrode. Combustion and decomposition of the electrode covering from the heat of the welding arc produce a gaseous shield that excludes the oxygen and nitrogen in the atmosphere from the weld area; these gases would otherwise cause excessive porosity and poor ductility in the weld. Welds by this method are of very high quality (Juers, 1993). Figure 13.2.15 depicts the components of the shielded-metal arc welding process.

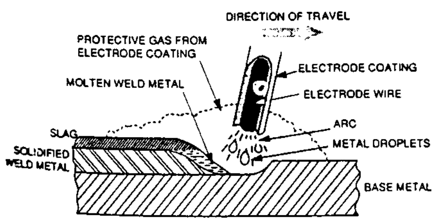

FIGURE 13.2.15 Shielded-metal arc-welding process. (From *ASM Handbook, Welding, Brazing, and Soldering,* Vol. 6, 10th ed., ASM International, Metals Park, OH, 1993, 175. With permission.)

Key System Components. For shielded-metal arc welding, the metallurgical properties of a weld depend greatly on the type of electrode and its covering. The *electrode coverings* contain shielding gas formers that exclude atmospheric gases from the weld area. Electrode coverings offer additional capabilities (Juers, 1993):

- Deoxidizers and nitrogen absorbers to purify the depositing metal
- Slag formers to protect the weld from oxidation
- Ionizing elements to stabilize the arc
- Alloying elements to produce higher-strength welds
- Iron powder to increase metal deposition rate

Selection of the proper electrode is based on many considerations (Juers, 1993):

- Base metal strength
- Base metal composition
- Welding attitude (position)
- Welding current
- Joint design and fit
- Base metal thickness and shape
- Service conditions
- Production efficiency and conditions

Many types and sizes of *power supplies* are used. Supplies can be either direct current (DC) or alternating current (AC) types; combination AC/DC power supplies are widely used. In general, power supplies are required that produce controllable levels of constant-current output. The rate of metal deposition is determined by the output current from the power supply.

Capabilities. Shielded-metal arc welding is the most widely used welding process for joining metal parts, principally because of its versatility. Also, the welding equipment is less complex, more portable, and less costly than for other arc-welding processes.

Shielded-metal arc welding is generally very useful in joining components of complex structural assemblies. Joints in virtually any position that can be reached with an electrode can be welded, even if directly overhead. Joints in blind areas can be welded using bent electrodes. Welding in positions other than flat require the use of manipulative techniques and electrodes that cause faster freezing of the molten metal to counteract gravity. Shielded-metal arc welding can be done indoors or outdoors.

Metals welded most easily by the shielded-metal arc process are carbon and low-alloy steels, stainless steels, and heat-resistant alloys. Cast iron and high-strength steels can also be welded but preheating and postheating may be required. Shielded-metal arc-welding electrode materials are available for matching the properties of most base metals; thus, the properties of a joint can match those of the metals joined.

Design Considerations. Joint design (shape and dimension) is determined by the design of the workpiece, metallurgical considerations, and established codes or specifications.

Welds should preferably be located away from areas of maximum stress. Poorly placed welds can result in undesirable, and unplanned, stress concentrations that can cause early failure of the joint.

Poor joint fit-up increases welding time and is often the cause of poor welds.

Process Limitations. Metals with a low melting point, such as zinc, lead, and tin, cannot be welded by electric arc methods.

Limitations of shielded-metal arc welding compared with other arc-welding methods are related to metal deposition rate and deposition efficiency. Consumable electrodes have a fixed length, usually 18 in. (460 mm), and hence welding must be stopped periodically to replace the electrode. Another limitation is the requirement to remove the slag covering that forms on the weld after each welding pass.

There is a minimum gauge of sheet that can be successfully welded without burn-through. Generally 0.060 in. (1.5 mm) is the minimum practical sheet thickness for low-carbon steel sheet that can be welded by a welder possessing average skill (Juers, 1993).

Special techniques are required when welding pieces of unequal thickness because of their different heat dissipation characteristics. Solutions include

- Placing a copper backing plate against the thinner section to match the heat dissipation from the thick section
- Redesigning the component so that the thick and thin sections taper at the joint to approximately the same size

Distortion is unavoidable in welding because of residual stresses that arise from nonuniform heating and cooling. Various procedures can be used to minimize distortion, such as clamping the workpieces. But straightening of the workpiece may be required to achieve the required dimensional accuracy.

References

Advani, S.G., Ed. 1994. *Flow and Rheology in Polymer Composites Manufacturing,* Composite Metals Series, Vol. 10, Elsevier, New York.

ASM. 1984. *Powder Metallurgy,* 1984. *ASM Handbook,* Vol. 7, 9th ed., ASM International, Metals Park, OH.

ASM. 1993. *Welding, Brazing, and Soldering,* ASM Handbook, Vol. 6, 10th ed., ASM International, Metals Park, OH.

David, S.A. and Vitek, J.M., Eds. 1993. *International Trends in Welding Science and Technology,* ASM International, Metals Park, OH.

Eagar, T.W. 1993. Energy sources used for fusion welding, in *Welding, Brazing, and Soldering, ASM Handbook,* Vol. 6, 10th ed., ASM International, Metals Park, OH, 2–6.

Froes, F. 1996, *Hot Isostatic Pressing,* ASM International, Metals Park, OH.

Humpston, G. and Jacobson, D.M. 1993. *Principles of Soldering and Brazing,* ASM International, Metals Park, OH.

Jenkins, I. and Wood, J. V., Eds. 1991. *Powder Metallurgy: An Overview.* The Institute of Metals, London.

Juers, R.H. 1993. Shielded metal arc welding, in *Welding, Brazing, and Soldering, ASM Handbook,* Vol. 6, 10th ed., ASM International, Metals Park, OH, 175–180.

Linnert, G., 1994. Fundamentals, in *Welding Metallurgy: Carbon and Alloy Steels,* Vol. 1, 14th ed., American Welding Society, Miami, FL.

MPIF. 1995. *Powder Metallurgy Design Manual,* 2nd ed., Metal Powder Industries Federation, Princeton, NJ.

National Research Council (NRC). 1995. Intelligent processing of advanced materials, in *Expanding the Vision of Sensor Materials,* NMAB-470, National Academy Press, Washington, D.C., 34–40.

Schwartz, M.M. 1994. *Joining of Composite Matrix Materials,* ASM International, Metals Park, OH.

Woishnis, W.A. 1993. *Engineering Plastics and Composites,* 2nd ed., ASM International, Metals Park, OH.

Mechanical Assembly

S. H. Cho

Total labor involved in the assembly processes in the U.S. varies from 20% (farm machinery) to almost 60% (telecommunications equipment). On average, assembly tasks occupy 53% of manufacturing time, and 10 to 30% of total production cost of most industrial products. Use of improved assembly methods and technologies is essential to reduce overall manufacturing costs.

Assembly Methods and Systems

Assembly systems are classified in several different ways (Table 13.2.18). Figure 13.2.16 indicates the types of automated systems that could be cost-effective based on assembly part count and production volume.

Generally, *automatic assembly systems* consist of three major components:

- *Transfer system* to move work carriers with in-process subassemblies between workstations;
- *Parts feeding device* to supply parts to be assembled into the appropriate position, where the parts are loaded by the handling/placing mechanism;
- *Parts handling/placing mechanisms* to pick parts and perform assigned assembly tasks such as placing, inserting, and screwing.

TABLE 13.2.18 Classification of Assembly Systems

Type of Assembly System	Classification Basis	Description
Manual	Level of automation	Assembly tasks completed manually
Automatic	Level of automation	Adopts mechanized devices or industrial robots with supplementary equipment for handling and assembling parts
Semiautomatic system	Level of automation	Manual workers and mechanized devices cooperate to complete assembly tasks
Cell-type system	Configuration	Very flexible integrated assembly workstation; assembly completed by various equipment, such as robots or pick-and-place units, parts feeders, parts tray, magazines, automatic tool changer, and auxiliary jig/fixtures
Line-type system	Configuration	Assembly tasks divided into subtasks which are completed at workstations connected by transfer systems; handles large parts, cycle-time variation, and gripper change
Dedicated system	Degree of flexibility	Not flexible — can assemble only one product of a single model; generally economical for large production volumes
Flexible system	Degree of flexibility	Accommodates different products; economical for medium-size and mixed-model production

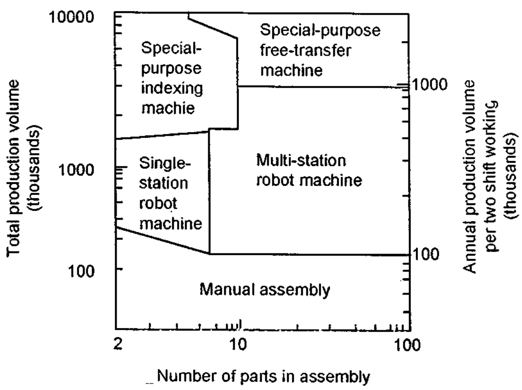

FIGURE 13.2.16 Conditions for economic application of various assemby systems. (Courtesy of S. H. Cho.)

The *transfer system* can be continuous or intermittent according to the transfer method. In the *intermittent transfer system,* assembly tasks in a workstation are performed during a stationary period of the work carriers, which are transferred to the subsequent workstation after completion of the assembly tasks. *Continuous transfer systems* have a problem of not being able to assure positioning accuracy between an in-process subassembly and the tools of a workstation; thus, the intermittent type is usually used (Boothroyd, 1992).

There are two modes of *intermittent transfer:* in-line and rotary. For the *in-line* mechanisms, the walking beam, the shunting work carrier, and the chain-driven work carrier are commonly used. The *rotary* type employs mechanisms such as rack and pinion, ratchet and pawl, Geneva mechanism, and mechanical cam drives. Currently, free-transfer conveyors with stopping and positioning mechanisms are widely used in flexible assembly lines.

Most *feeding systems* have devices to orient parts supplied by the following means (Yeong and Vries, 1994):

- *Bulk supply,* for parts that are easily separated, fed, and oriented automatically. Bulk supply usually adopts various part feeders, e.g., vibratory bowl feeders and various nonvibratory mechanical feeders for small parts. These feeders usually only handle one type of part and cannot be applied to assembly systems which require flexible part-feeding devices.

- *Organized supply* uses special pallets such as a kit or a magazine for parts that cannot easily be separated, fed, and oriented.

Parts-handling/placing mechanisms include pick-and-place units and various types of industrial robots. To load and assemble parts, the mechanism usually employs various assembly wrists: jaw-type gripper, vacuum suction pad, magnetic chuck, screwdriver, nut runner, and others. A number of different wrists are required when automatically assembling different parts in a workstation. To accommodate this assembly situation, a multifunctional gripper, a tool-changing system, and a universal gripper have been developed and widely used in robotic assembly systems.

Selection of Assembly Systems

The placement of a part in its assembled position and part mating impose tight constraints on the positioning mechanism and on part properties such as clearances and geometry. These constraints are more severe for assembly systems mating a variety of precision products that must adapt to frequent design changes. The assembly systems of this type must possess the adaptability to changing assembly environments, thus requiring *flexible automatic assembly* (Boothroyd, 1984).

A typical *flexible robotic assembly system* uses an industrial robot for part handling, part positioning, and part mating. These systems are limited by positioning and orientation misalignment caused by the low positioning accuracy of robots, uncertainty in part handling, and variation in the location of parts.

Various approaches are available that take into consideration uncertainty in orientation and parts properties variation. Figure 13.2.17 depicts various types of assembly wrists that are in use. In general, wrists can be classified into three basic configurations:

- Passive accommodation;
- Active accommodation;
- Passive-active accommodation.

There are two types of *passive wrist methods:*

- Wrist accommodates misalignments by deforming its structure elastically under the influence of the contact forces generated during the assembly of the misaligned parts. A *mechanically compliant structure* is needed for either the robot wrist or the assembly worktable, which can be deformable according to the reaction force acting on the mating parts. Remote center compliant wrist is one of such typical wrists; this method usually requires part chamfering (Cho et al., 1987).
- Wrist corrects misalignment by *applying external forces* or torques to the misaligned parts in a prescribed manner or a random way. For instance, a vibratory wrist utilizes pneumatic actuators controlled by a pulse width modulation controller to generate desirable vibration; this method does not require part chamfering.

Active wrist methods employ sensor-controlled wrists and compensate misalignments by controlling the fine motion of the assembly wrist or the work table based upon sensory feedback. Advances have been made in the area of sensing technique, gripper and actuating mechanism design, and the related control algorithms. Sensors for these wrists are needed prior to contact (vision, range, displacement, proximity), during contact (touch, slip), and after contact (force, moment). Based upon the force sensor signals and the associated algorithms, the wrist motion can be corrected to reduce misalignment.

The *passive–active accommodation method* is achieved by combination of the "passive and active" techniques. The basic strategy is that the part mating is continued within some allowable forced moments, while beyond this the insertion method is switched from the passive to active to reduce the mating force by using sensors with compliant structures.

The *sensor-based assembly* is similar to the method employing the active wrist. Both rely on sensory information for fine-motion control. The principal difference is that the former utilizes robot motion, while the latter relies on wrist motion.

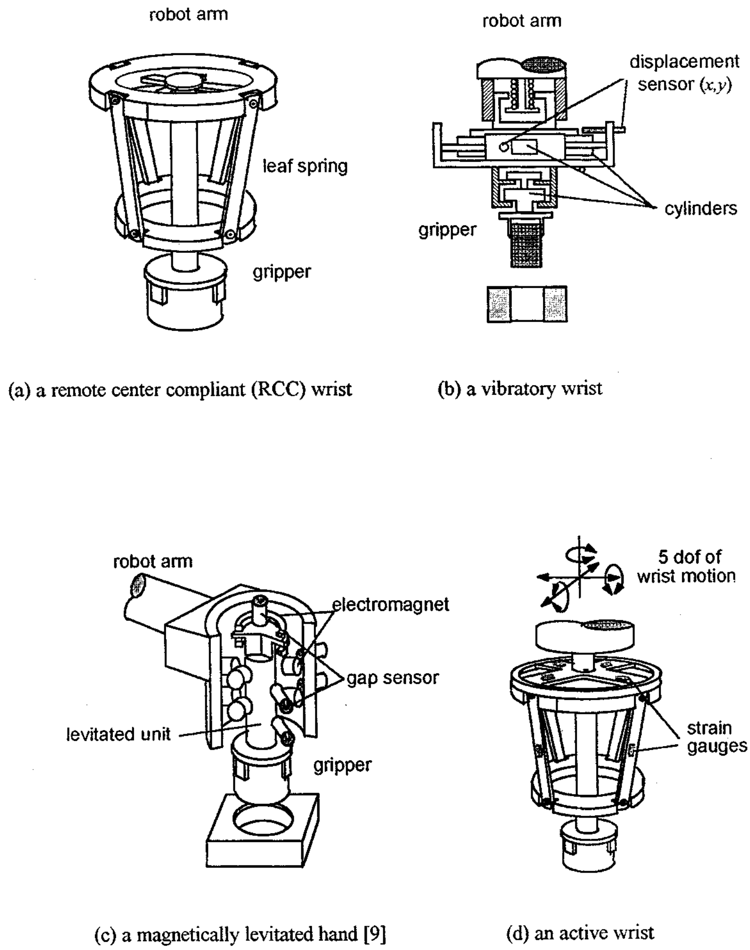

FIGURE 13.2.17 Various assembly wrists. (Courtesy of S. H. Cho.)

Reaction forces (*force/torque information*) that occur during part mating can represent the status of the contact between mating parts. Among the approaches are pattern classifiers that determine the contact state at which the assembly parts are contacting each other, position error recovery via fuzzy logic, heuristic search with fuzzy pattern matching, and learning of nonlinear compliance using neural network.

Visual/optical information is critical to compensate for positioning and orientation error occurring due to misalignment. Visual information is often combined with other data, such as force/torque, pressure, and displacement, because a rather longer time is required for image processing, object recognition and error calculation and because visual information is sensitive to external environmental conditions such as illumination.

Assembly Line

An assembly line usually consists of a set of workstations that perform distinct tasks linked together by a transfer mechanism. Each task is an assembly operation, while each workstation represents a location along the line where the tasks are processed. A buffer storage is placed between workstations for reducing the effect of a workstation failure on the throughput.

Line balancing is essential for designing a cost-effective assembly system (Groover, 1980). The time required for the completion of a task is known as the process time, while the sum of the process times of the tasks assigned to a station is the station time. The total task processing times for assembling all the parts is the work content. When *assembly sequences* are generated without considering line balancing, the sequences may not guarantee the minimum number of workstations. Therefore, line balancing must be concurrently considered in defining the assembly sequences.

Design for Assembly

Approximately 80% of manufacturing cost is determined at the conceptual design stage. Design for assembly (DFA) is crucial during early design. The objective of the DFA is to facilitate the manufacturing and assembly of a product. DFA applies to all the assembly operations, such as parts feeding, separating, orienting, handling, and insertion for automatic or manual assembly (Ghosh and Gagnon, 1989).

DFA is directed toward:

- Reducing the number of parts by modularization;
- Easing feeding and minimizing reorientation;
- Easing insertion by self-aligning, self-locating, elimination of part interference, and efficient fastening.

Axiomatic DFA uses design guidelines based on experience of product designs and assembly operations. *Procedural DFA* evaluates the design efficiency based upon the production cost.

Figure 13.2.18 shows results that were obtained using DFA rules to evaluate part designs in consideration of feeding and insertion. Such analysis has great potential in improving assembly operations.

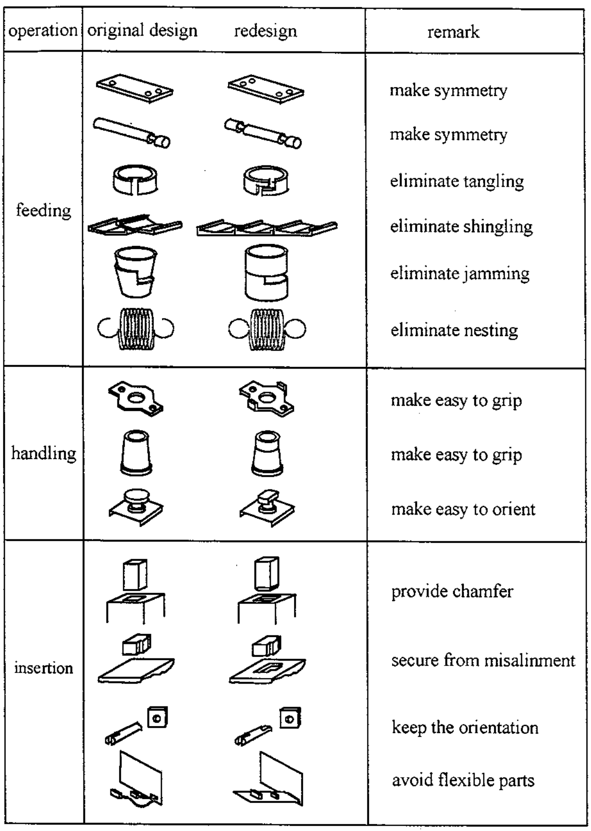

FIGURE 13.2.18 Results of DFA analysis. (Courtesy of S. H. Cho.)

References

Boothroyd, G. 1984. *Economics of General-Purpose Assembly Robots,* CIRP General Assembly, Madison, WI.

Boothroyd, G. 1992. *Assembly Automation and Product Design,* Marcel Dekker, New York.

Cho, H.S., Warnecke, H.J., and Gweon, D.G. 1987. Robotic assembly: a synthesizing overview, *Robotica,* 5, 153–165.

Ghosh, S. and Gagnon, R.J. 1989. A comprehensive literature review and analysis of the design, balancing and scheduling of assembly systems, *Int. J. Prod. Res.*, 27(4), 637–670.

Groover, M.P. 1980. *Automation, Production Systems, and Computer-Aided Manufacturing,* Prentice-Hall, Englewood Cliffs, NJ.

Yeong, M.Y. and Vries, W.R. 1994. A methodology for part feeder design, *Ann. CIRP,* 43(1), 19–22.

Material Handling

Ira Pence

Material handling provides the right amount of all the required materials at the right place and time to support manufacturing. Properly designed, the material-handling system provides for the acquisition, transportation, and delivery of material such that the minimum cost is incurred considering capital, labor, and expenses. It focuses on obtaining material and supplies, moving them between process steps, and delivering the finished product to the customer. Figure 13.2.19 depicts the elements of a modern materials-handling system which reaches beyond the factory floor, serving as an integrating force for production operations.

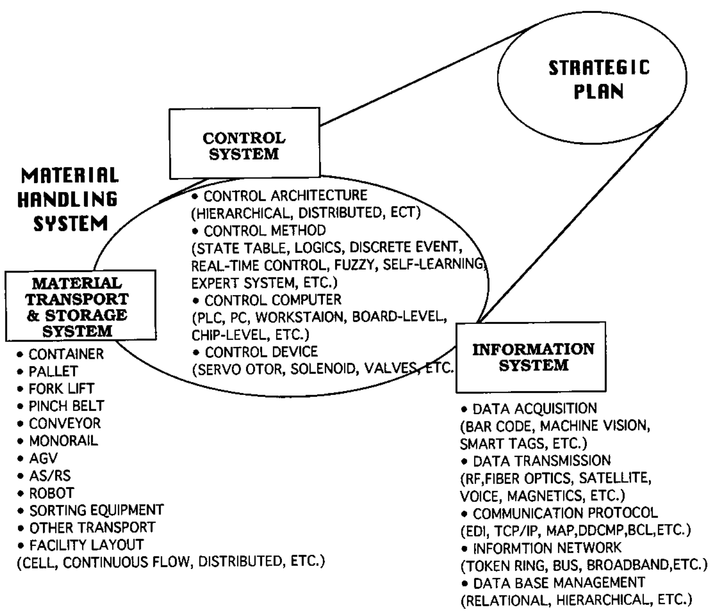

FIGURE 13.2.19 Modern material-handling system. (Courtesy of J. Lee.)

The material-handling system should be analyzed as a single entity so trade-offs in expense in various elements can be made to minimize the total cost. Often, however, the activities are broken into purchasing, transportation, warehousing, and distribution functions which operate independently. Such divisions are arbitrary because the material-handling system is an unbroken chain of activities that start when the suppliers complete their last value-added operation and extend until the product is delivered. Minimum-cost material handling will not be achieved if the partitioning of material flow into subelements leads to optimization of each subelement at the expense of the entire system.

Planning is crucial to the design of a smoothly functioning materials operation. Planning should be consistent with the strategic plan for the entire manufacturing operation. In formulating the plan, a realistic assessment of the problems and opportunities associated with arranging, controlling, and implementing the material flow must be used. Specific tactics, such as just-in-time delivery, should support the strategic plan. Expensive automation should not be used unless the volume and stability of the product justify its use.

Logistics

The movement of material in a manufacturing enterprise is usually broken into two broad categories, inside the plant and external to it. External movement is generally referred to as "logistics," while the internal flow is known by many different names. Material movement considered logistics is generally marked by a wide geographic scope, diversity of equipment and technology, and some uncertainty in delivery time. But the basic goals are identical to in-plant movement and most of the analysis tools are applicable to both categories.

For large manufacturing enterprises the number of individual components involved in the production of the product, often referred to as SKUs (stock keeping units), can be very large. In the past, it was helpful to treat items of similar size, weight, storage requirements, and delivery times as a single commodity. This simplified manual analysis and was sometimes done by computer programs that analyze logistical information. However, current computer systems make it practical to treat each item individually.

Basic Elements

The basic activities associated with material handling are *moving, storing,* and *controlling* materiel. These activities are interrelated with production scheduling and information must flow both ways.

Movement may be over short distances, as from one machining center to another, or long, from one plant to another. In all cases the basic information that must accompany the move includes the part identity, timing, quantity, source, and destination. Each move should be planned and scheduled, taking into account the speed of the basic mechanism as well as allowance for loading, unloading, logging, counting, etc.

Movement can be continuous or in batches, synchronous or asynchronous, horizontal or vertical. Each move should be examined for the characteristics of urgency, safety, size, weight, etc. before selecting the technology to be used to perform the move. Singular instances should be accommodated in the most expeditious manner, and repetitive moves must be made in the most efficient and effective manner.

One of the basic tenets of material handling is to retain *control* of the material. Inventory control generally maintains up-to-date records on the quantity and location of material on hand. However, in a *virtual warehouse*, it is important to know what material has been ordered and the status. Control includes procedures and equipment to properly handle the material.

The *storage* of material should be minimized. Materials are typically stored to compensate for uncertainty in delivery systems and to allow ordering of economic quantities.

Occasionally unexpected changes in production will result in delaying the release of material to the factory floor and that material must be stored until needed. *Just-in-time delivery,* where parts are delivered directly from local suppliers to the assembly line several times a day, has been proved in several industries. Thus, the need for storage due to uncertainty of delivery has been reduced. At the same time, global sourcing has increased. As supply lines lengthen, the uncertainty of delivery increases. The more complex the supply system, the more likely the occurrence of an unexpected delay. In designing a storage system, the trade-off of higher transportation cost but less inventory vs. volume discounts on purchases and transportation, but with storage costs, must be evaluated.

Further Information

The Material Handling Industry produces two catalogs each year. One provides information on the publications available from the MHI, including all the educational material; standards and specifications; operating, maintenance, and safety manuals; and reference works. The other provides a directory of member companies and the products they manufacture. Both are available from the MHI Literature Department (704) 522-8644; FAX (704) 522-7826.

Case Study: Manufacturing and Inspection of Precision Recirculating Ballscrews

Toskiaki Yamaguchi, Yashitsugu Taketomi, and Carl J. Kempf

The precision and quality of the components of mechanical devices used in both industrial and consumer products must meet high performance and durability requirements. The manufacturing challenges that apply to ballscrews, in general, are representative of other components such as gears, shafts, and bearings. This case study on precision ballscrews illustrates how the different unit manufacturing processes presented in this chapter apply to a particular application, and how design and manufacturing engineering decisions are affected by quality and cost considerations.

Many of the processing steps and fundamental techniques discussed in this case apply to the manufacture of many other precision components. The case study also illustrates the rationale for continuously improving production processes and discusses strategies for improvement that benefit from past experience.

Overview of Ballscrew Design and Manufacturing Considerations

Ballscrews convert rotary motion into linear motion. Ballscrews have low friction compared with standard leadscrews, and have enabled precise control of mechanical systems at a relatively low cost. They are used extensively in production machinery, such as milling machines, and are being applied in other fields, such as robotics, inspection equipment, and office automation equipment.

The key components of a ballscrew system consist of a screw shaft with a spiral groove, a nut with a corresponding spiral groove that rides along the shaft, and balls which are captured between the shaft and the nut. A recircultation tube provides a return path for the balls from the end of the nut. Components in this assembly are typically made from steel alloys, chosen to provide a good combination of toughness, surface hardness, and ease of manufacture.

Ballscrews are applied in a variety of ways, but the most common configuration is one in which a rotational input of the shaft imparts a translational motion to the nut. The shaft is normally supported by rotating bearings at both ends, and the translating element attached to the nut is supported by linear guide bearings. Ballscrews range in size from very small units with a shaft diameter on the order of 0.08 in. (2 mm) and a length of about 4 in. (100 mm) to very large units with a shaft diameter on the order of 12 in. (300 mm) and a length of up to 50 ft (16 m). Key dimensional features which are controlled to a high degree of precision for an assembled ballscrew system are shown in Figure 13.2.20.

The main factors in ballscrew performance are accuracy and lifetime. *Accuracy* is determined primarily by the precision of the screw lead, i.e., the linear displacement of the nut which is produced by rotary displacement of the shaft. The measurements used to assess the *precision* of the lead are discussed below. To obtain a *long lifetime*, the shaft must have high surface hardness in order to withstand loads at the shaft/ball and nut/ball interfaces. Since ballscrews may be subject to a variety of loading conditions, the ballscrew must possess good impact strength and toughness.

Secondary design factors include low audible noise, low static and dynamic friction, low friction variation, minimal backlash, high mechanical stiffness, and resistance to dirt and contaminants. Ballscrews are often used in specialized applications, such as operation in vacuum or ultraclean environments, in corrosive or dirty environments, in a thermally controlled environment, and in environments where vibrations must be minimized. To meet these demands, basic ballscrew designs are adapted to

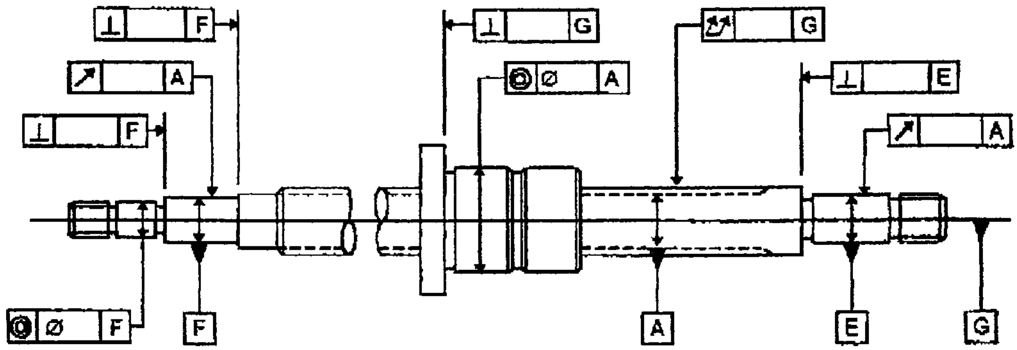

FIGURE 13.2.20 Key ballscrew dimensions. (Courtesy of T. Yamaguchi, Y. Taketomi, and C. J. Kempf.)

use special materials, surface treatments, and to provide features such as hollow shafts for coolant flow or damping materials.

Major steps in production and assembly of traditional ground ballscrews are summarized as a flowchart in Figure 13.2.21.

Initial Machining Operations

A lathe is used to establish the outer diameter of the shaft and cut the groove. In both operations, an allowance is made for material that will be removed in the final grinding operations. For short shafts, the support is provided from the shaft centers and a single cutting tool can be used with multiple passes to produce the desired outside diameter and groove. For long ballscrews, the lateral deflection of the shaft during cutting operations is significant and workpiece supports are necessary to prevent deflection of the shaft.

For long ballscrews, the total processing time can become extremely long using multiple passes with a single-point cutting tool. In such cases, a multiple-point tool that can remove more material on a single cutting pass is more efficient to use. But a multipoint cutter requires more time for setup and adjustment. Thus, there is an economic trade-off between single-point and multipoint cutting techniques.

Because of subsequent heat-treat operations, the shaft will undergo dimensional changes. The design must allow for these dimensional changes. This is an area in which production experience is critical. Statistical analysis of previous manufacturing results, together with modeling of material behavior, is critical to continued design and proceess improvement.

Surface Treatments

In order to withstand the loads at contact points, a very high surface hardness is necessary. A surface hardness of Rockwell C 58-62 with a depth on the order of 0.8 to 1.2 mm is required. For short ballscrews, *carburization* is used to develop the necessary surface hardness. For longer shafts, carburization is impractical because of the size of carburizing furnaces. In this case, an induction hardening process is used.

Electrically heated gas furnaces are used in the carburizing process. The immersion in the carburizing furnace is followed by quenching and tempering. During carburizing, key parameters such as temperature and gas concentrations are continuously monitored and adjusted by a process control system. Because high hardness is needed only in the ball groove, areas of the shaft and nut that will be subject to subsequent machining operations are coated before carburizing. This coating, which is applied like paint, prevents the diffusion of carbon into the material surface.

For longer shafts, a different steel is chosen and an induction hardening process is used. The general trend within the industry is toward induction hardening since it has the advantage of being a continuous process. When using induction hardening, single shafts can be processed immediately after machining. Induction hardening machines require less initial capital than for a carburizing system.

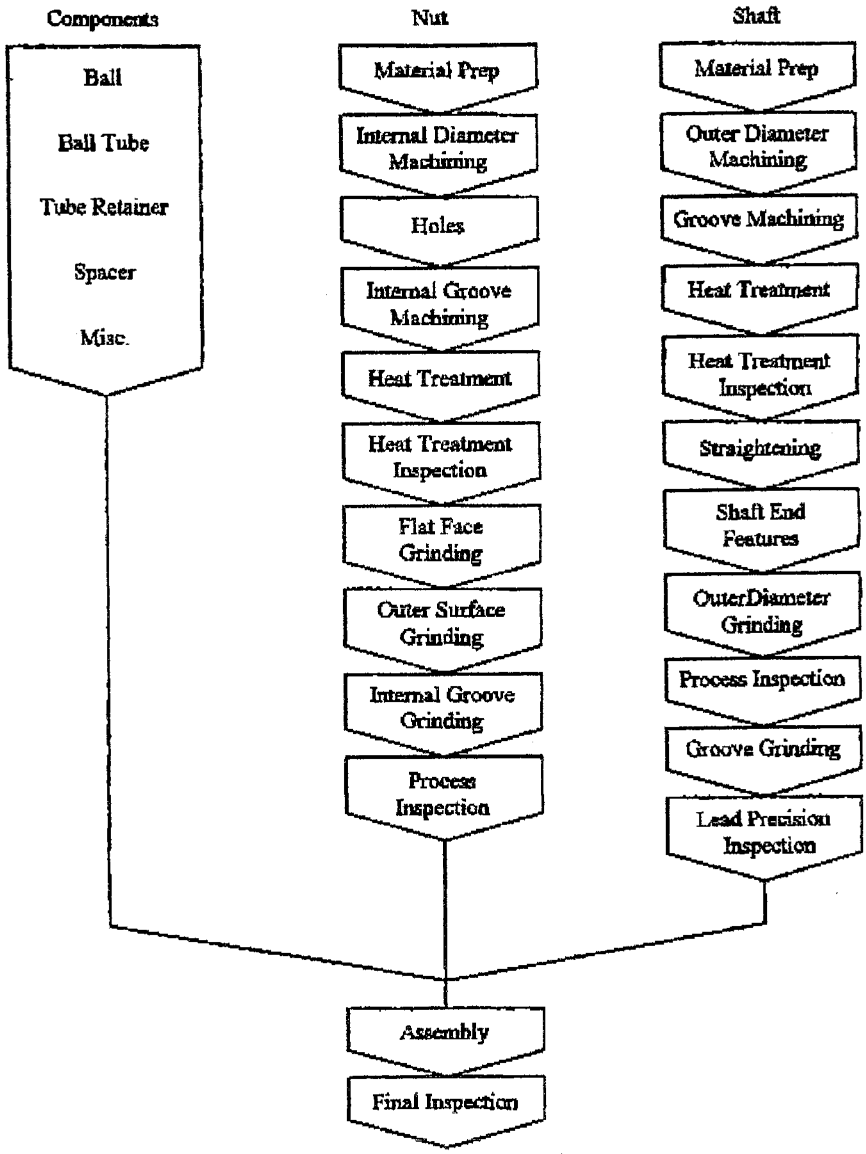

FIGURE 13.2.21 Ballscrew processing flowchart. (Courtesy of T. Yamaguchi, Y. Taketomi, and C. J. Kempf.)

The shaft undergoes dimensional changes as a result of the surface heating, and a *straightening* operation is necessary. Prior to straightening, the shaft is chilled in a carefully controlled manner to sub-zero temperatures to ensure that the solid-state transformations are completed; the result is a more stable microstructure. For straightening, a *specialized press* is used to measure automatically the deviation from perfect straightness; it automatically applies loads along the shaft as necessary to restore the shaft to near-perfect straightness.

Grinding and Finishing Operations

After surface treatment, chilling, and straightening, the outside diameter of the shaft is ground to the final dimensions. Except for very short ballscrews, the grinding cutting forces will cause lateral deflection of the shaft, and thus the centers in the shaft ends cannot be used as datums for final machining operations. Consequently, the outside diameter of the shaft is ground to precise tolerances. Although the outer diameter of the shaft is not a functional part of the finished ballscrew, it is used as a datum surface when the final grinding of the shaft groove is performed.

Grinding of the shaft can be done on either centerless grinders or cylindrical grinders. Since centerless grinders take a long time to set up, they are generally used for large production runs of shafts having the same diameter. For the case of cylindrical grinders, the shaft is supported at the centers in the shaft ends as well as at workpiece supports along the length of the shaft. By adjusting the work rests, minor variations in taper and bending of the shafts can be corrected. This processing step requires considerable operator skill and experience in order to minimize variations in the outer diameter and residual bending of the shaft.

To maintain precision and finish of the outer diameter, automatic balancing of the grinding wheel is necessary. CBN grinding wheels are used although they are more expensive than traditional abrasive wheels; their long life and low wear rates make them economical.

After turning the external diameter to final size, shaft end features are produced. These features include bearing seats, keyways, flats, and locknut threads. The final finishing operation for the ballscrew generates the ball groove. This is the key processing step in assuring that the ballscrew possesses the required lead accuracy. If this process is not carefully controlled, there will be variation in the lead, depth of the groove, and smoothness of the groove, causing subsequent problems with accuracy, stiffness, running force and noise, and lifetime.

Removal of the outer layer of hardened material during outer diameter grinding can cause minor bending of the shaft as the net residual stresses in the material change. Prior to final cutting of the ball groove, the straightness of the shafts is checked and minor corrections made.

To maintain very high lead precision, the final groove grinding is done in a specially temperature-controlled environment of 68 ± 2°F (20 ± 1°C). Cutting oil is applied liberally to the shaft to minimize thermal effects due to cutting and deviations between the grinding machine and shaft temperature. As in the case of external diameter grinding, automatic balancing is used to minimize vibrations. To further reduce vibrations, each grinding machine is mounted on an individual base to minimize vibration coupling between various machines. The isolation properties of the machine bases are adjusted when a new machine is placed in service and undergo periodic inspection and adjustment during operation.

Since the size and lead of the groove varies from ballscrew to ballscrew, the grinding wheel must be matched to the groove shape. Thus, an inventory of several types of wheels is necessary; traditional abrasives are used since the costs of CBN wheels would quickly become prohibitive.

In general, accuracy in the groove-grinding process requires a combination of modeling and statistical analysis. As in the case of surface treatment, the collection and analysis of past production data allows continuous refinement of the manufacturing process.

Assembly and Inspection

The main factor in ballscrew accuracy is the lead. To facilitate a quantitative measurement of the lead, four fundamental parameters are used. To measure the lead error, precision measurements over a long travel range are made using computer-controlled laser interferometry.

Unacceptable variations in friction, increased running noise, and a reduction in life result from poorly formed grooves. Thus, the depth and cross-sectional profile of the groove must fall within allowable tolerances. Direct measurement of groove cross section is quickly performed on an optical profile projector. Precise measurement of groove depth can be made on selected samples. The screw shaft is supported between two centers and rotated at a fixed speed. A table carrying a contact probe moves along the screw shaft in the axial direction synchronously with the screw rotation. The probe is placed in contact with the groove to measure variations in the groove depth over the entire length of the shaft. By performing a frequency analysis on the groove depth errors and accounting for the rotational speed of the ballscrew during the measurement process, frequencies of unwanted vibrations occurring in the production machinery can be detected and the source of the anomaly eliminated to improve the production process.

Even with good control of the screw shaft and nut groove diameters during production, additional steps are necessary to obtain the desired amount of axial play or preload in the assembled ballscrew. By selecting slightly different ball sizes, required axial play or preload can be achieved. For a given nominal ball diameter, balls usually are grouped in steps of 20 to 40 μin. (0.5 or 1.0 μm). To increase production efficiency, assembly jigs are used.

The majority of precision ballscrews are preloaded in order to remove backlash and achieve the desired axial stiffnes. Two methods of achieving the desired preload are used. The first preloading method increases the ball size until the desired preload is achieved. The second method uses double-nut preloading in which a spacer is inserted between the two nuts to take up the axial play and achieve the desired preload.

The preloaded ballscrew has some running torque when the screw is rotated. The relation between this running torque and the preload has been determined based on both theoretical and experimental studies. Since preload cannot be measured directly, it is estimated based on measurements of the running torque in a specialized torque-measuring machine.

A special machine is used for direct measurement of axial stiffness. In this process, the screw shaft is clamped and an axial load is applied to the nut. A displacement sensor is fixed to the shaft close to the nut and the relative displacement of the nut can be measured when force is applied. The measurements of force and displacement can be plotted on an X-Y recorder to depict the stiffness characteristics of the ballscrew assembly.

References for Case Study

Oberg, E., Ed. 1971. *Machinery's Handbook,* 19th ed., Industrial Press, New York 2044–2068.

Yamaguchi, T., 1983. Ballscrew manufacturing, and inspection, *Tool Eng. Mag.*, June, 92–99 (in Japanese).

13.3 Essential Elements in Manufacturing Processes and Equipment

Sensors for Manufacturing

John Fildes

Introduction

A good modern definition for a sensor must capture the diversity of these devices. A sensor is a device that detects or measures the state or value of a physical or chemical variable and provides the result in a useful way. At the minimum, a sensor contains a transducer that converts the detected or measured quantity to another form of representation. For example, a very simple sensor is an indicator whose color changes upon reaction with a minimum amount of a chemical species. Nonetheless, the sensors that are normally encountered are more complex, containing a transducer, an output display, and possibly supporting electronics for signal conditioning, communications, and logic functions.

Sensor technology is undergoing rapid change because of three developments. One development is the emergence of integrated and smart sensors, wherein transducers have been miniaturized, usually through the use of silicon micromachining, and integrated with electronics for signal conditioning, communications, and logic functions. The second development is the ongoing adaptation of nondestructive evaluation (NDE) measurements and laboratory measurements for on-line use in supervisory and intelligent process control systems. These NDE and laboratory-type measurements require rather complex systems, with extensive signal conditioning and data analysis. The third development is also related to the emergence of supervisory control systems. The data from multiple sensors is comparatively analyzed in a process called data fusion to better identify the state of the system and the occurrence of process faults. Thus, the topic of sensors and sensory systems now encompasses transducers, integration with supporting electronics for communications and logic functions, data analysis techniques, and data fusion methods.

Classification of Sensors

A good taxonomy for sensors is provided by the requirements for different degrees of process control, which is shown in Figure 13.3.1. This taxonomy contains three classes: sensors used in regulatory feedback control loops, process analyzers, and product quality analyzers. In Figure 13.3.2, the basic element of control is the regulatory feedback loop that maintains controllable processing parameters at the desired values. These types of sensors, which usually provide a single value and are relatively generic in their applicability, are used for monitoring variables such as temperature, pressure, flow, level, displacement, proximity, and velocity. For use of sensors in regulatory control, sensitivity, selectivity, simplicity, speed, reliability, and low cost are the critical attributes. Historically, sensors were almost solely used in this function, but this is no longer the case. Sensors, or more properly sensory systems, are now also used as process analyzers and product quality monitors. Sensors of this type, which are used for feedback in supervisory control, are more complex and application specific, and their output tends to be a matrix of values (e.g., a video image or a spectrum). Sensors of this type provide a representation of the process or product that has greater information content, but is more abstract than the representation provided by sensors used in regulatory control. Thus, extensive computations and modeling are needed for process analyzers and product quality sensors, but there is also more ability to correct for deficiencies in sensitivity and selectivity through computational means. In this case, measurement speed is usually less demanding. These differences in the three types of sensors are summarized in Table 13.3.1.

The typical regulatory variables are temperature, pressure, flow, level, humidity, position, and motion. These variables form the basis for many process factors as summarized in Table 13.3.2 and the major sensor technologies for these variables are summarized in Table 13.3.3.

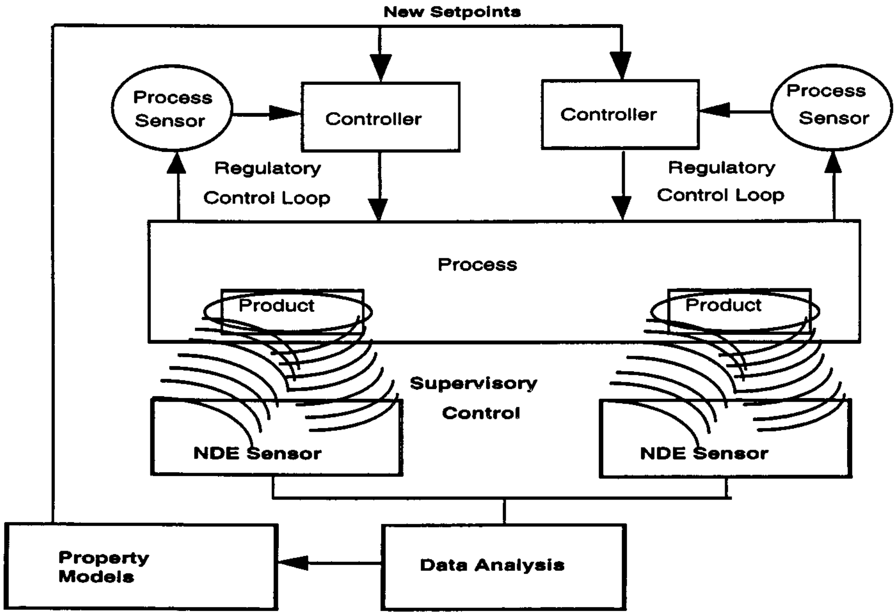

FIGURE 13.3.1 Regulatory and intelligent control.

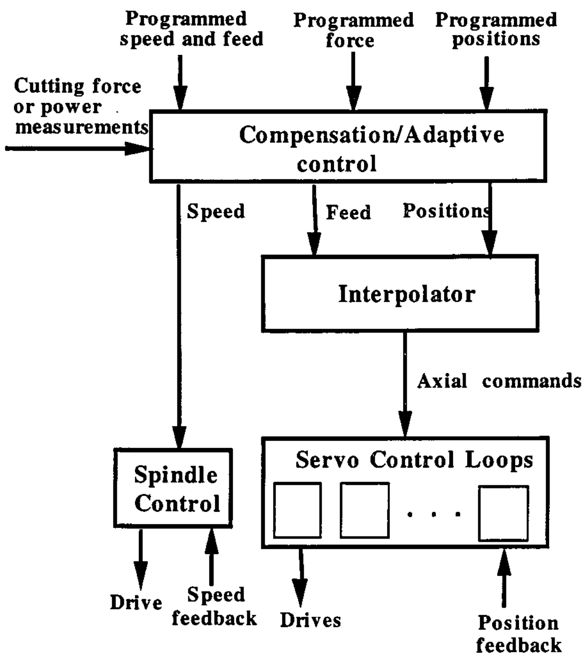

FIGURE 13.3.2 Typical CNC architecture.

Integration refers to inclusion of signal conditioning on the same substrate as the transducer. One of the major advantages of integration is that the integrated sensor can operate with a very small transducer output, which results in significantly smaller sensors. With the advent in telecommunications and the Microelectricalmechanical systems (MEMS) technologies, we can integrate software, hardware, power

TABLE 13.3.1 Characteristics of Regulatory and Intelligent Process Control Sensors

Characteristic	Regulatory Feedback Sensors	Process and Product Sensors
Speed of response	Very fast	Slow
Type of output	Single value	Matrix of values
Relationship of output to system parameter	Simple and direct	Abstract representation
Amount of computation	Little or none	Extensive
Sensitivity	Inherently high	Can be improved by computation and modeling
Selectivity	Inherently high	Can be improved by computation and modeling
Cost	Very low to moderate	High
Size	Small	Large
Applicability	Broad	Application specific

TABLE 13.3.2 Regulatory Processing Variables

Controllable Variable	Influenced Processing Factors
Temperature	Rates of chemical reactions; degree of cure of polymers; degree of softening and viscosity; cooking times; drying times; annealing times; degree of carborization
Pressure	Consolidation density; porosity; concentration of gaseous chemical reactants; forging; bending; injection times in molding
Flow	Mix ratios; quenching time; concentration of chemical reactants; color; pH; humidity
Level	Capacity; mix ratio; concentration of chemical reagents
Humidity	Drying of paper and grains; cooking
Position	Parts handling and placement; machining accuracy
Motion	Assembly; machining accuracy

TABLE 13.3.3 Sensor Technologies for Regulatory Control

Variable	Major Sensor Technologies
Temperature	Thermocouples, thermistors, resistance thermometers, infrared pyrometers, acoustic pyrometers
Pressure	Manometers, bellows and diaphragms, strain gauges, piezoresistive, piezoelectric, capacitive, thermocouple, ionization
Flow	Differential pressure (orifice plate, Venturi, pitot tube); velocity (turbine, vortex, electromagnetic, Doppler ultrasonic, time-of-flight ultrasonic); positive displacement (rotary sliding vane, gear, impeller); mass flow (thermal, coriolis); variable area (rotameters)
Level	Floats, pressure, radio frequency, ultrasonic, microwave, resistance tape, optical
Humidity	Dew point, length of hair, conductivity, capacitive, resonate
Position	Rotary encoder, linear variable differential transformer, potentiometers, magnetostrictive, magnetic, inductive proximity, magnetic, ultrasonic
Motion	Tachometers, pitot tube, anemometers

source, and communications on the same chip. Table 13.3.4 lists the advantages and disadvantages of integrated sensors.

TABLE 13.3.4 Characteristics of Integrated Microsensors

Characteristic	Advantage/Disadvantage
Batch fabrication	Excellent control, low cost, sensor arrays
Loss of modularity	Difficult to package and limitations on materials
On-chip amplification	Better signal-to-noise ratio
On-chip compensation	Better accuracy through compensation for interferences
On-chip feedback	Better linearity
On-chip scaling and conversion	Standardized output
On-chip multiplexing	Sensor arrays, bus addressable

Smart sensors are defined as those that also include logic functions, rather than just signal-conditioning electronics. Smart sensors may also include communications circuitry, diagnostics, and sometimes control outputs so that they can be directly connected to actuators, and documentation and trending functions. Conventional pressure sensors have rather broad accuracy limits, such as 0.25% of span. High-end smart pressure sensors have an accuracy of less than 0.1% of span, and mid-range smart pressure sensors approach this accuracy. Smart pressure transmitters are already available with built-in proportional-integral-derivative control functions, and fuzzy logic capabilities will soon be available. Multivariate measurement is not yet available, but will become so soon. Most likely, pressure and temperature will be the two measurements. Smart sensors are also available for presence detection, positioning, infrared photodetection with triangulation, and flowmeters, specifically magnetic meters, Coriolis mass flowmeters, and ultrasonic flowmeters. Integrated and smart sensor technology is also improving accelerometers, proximity detectors, and tactile sensors. The use of integrated and smart sensors will increase because of the overloading of shared resources and the use of distributed control schemes to solve this problem.

Use of Sensors in Supervisory and Intelligent Control Systems

As shown in Figure 13.3.1, supervisory control augments regulatory control by using process analyzers to better characterize the state of the process and sensors of product characteristics to assess the outcome of the process and to determine corrections for unacceptable deviations. These corrections are then implemented by the regulatory controllers. There are many ways in which supervisory control can be implemented. A common way is to measure important aspects of the process and the product performance, and to use a model to relate variations in processing parameters to variations in product performance.

Process analyzers are central to enabling supervisory control because they allow a more accurate, but abstract, representation of the state of the process. These devices augment regulatory control by monitoring other important variables in the process, but ones which are generally not directly controllable. Use of process analyzers provides improved reliability, flexibility, predictive diagnostics, ease of use, and central data collection with process documentation, trending, recipe handling, and statistical quality assurance. Process analyzers are now used for monitoring important process gases such as oxygen, carbon monoxide and dioxide, oxides of nitrogen and sulfur, and hydrocarbons. For liquids, process analyzers are available for pH, conductivity, redox potential, dissolved oxygen, ozone, turbidity, specific ions, and many organic compounds. The measurement techniques are summarized in Table 13.3.5.

TABLE 13.3.5 Process Analyzers

Process Analyzer Technology	Typical Uses
Electrochemical (potentiometric, amperometric)	Gases such as carbon monoxide and dioxide, oxides of nitrogen and sulfur, and hydrocarbons; species in liquids such as dissolved oxygen, pH, redox potential, specific ions, organic compounds, inorganic compounds
Chromatography (gas and liquid)	Gases such as CO and CO_2; species in liquids such as alcohols, flavors, lipids, polymers, and other organic compounds
Infrared spectroscopy	Near infrared — web processes for thickness, composition, solvent, coating Mid-infrared — chemical and petrochemical processes, polymers, food processes
Ultraviolet/visible	Gases and liquids including all elemental halogens, other inorganics, aromatics, carbonyls, many salts of transition metals

Traditionally, process analyzers have been stand-alone devices with a single sensor and a dedicated operator interface. The trend is toward modularity, multiple sensors, and digital communications so that process analyzers can be incorporated into distributed control systems in an open architecture control environment. Smart sensor technology is also turning conventional sensors into process analyzers. A good example is provided by smart infrared temperature sensors. These sensors provide sample and hold, correction for reflected radiation when the emissivity of the target is less than one and its temperature is lower than ambient, analog outputs for control, digital outputs, trend analysis, and area sampling with line scanners. This functionality has made temperature mapping of surfaces in furnaces practical and is used in annealing of aluminum, steel reheating, and oven drying of webs such as paper.

Another trend is the use of sensor fusion, where information from multiple sensors is combined to improve the representation of the process. Sensor fusion techniques can use mechanistic models, statistical models, or artificial intelligence techniques, such as neural networks, fuzzy logic, and expert systems. Sensor fusion provides more reliability because the validity of the data of each sensor can be assessed from the other sensors, and if a sensor is faulty, its value can be predicted from the other sensors. Sensor fusion provides a better representation of the process because it captures the interrelationships that often exist between processing variables, which are treated independently without sensor fusion.

As shown in Figure 13.3.1, intelligent control involves the use of product quality sensors. Of the three classes of sensors discussed in this section, these sensors tend to be the most complex and provide the most complete, but abstract, representation of the state of the product. In some cases, the same technologies are used for product quality sensors as for process analyzers. An obvious example would be where chemicals are reacted without a change of state. In other cases, specialized product quality sensors have emerged. Table 13.3.6 lists some of the product quality sensors that are being used or are being experimentally evaluated.

TABLE 13.3.6 Product Quality Sensors

Sensor	Applications
Time-of-flight ultrasound	Either thickness or elastic constants, if the other is known
Electromagnetic acoustic transducer (EMAT)	Either temperature or elastic constants, if the other is known
Eddy currents	Electrical conductivity, magnetic permeability, thickness, temperature, presence of flaws
NMR (high resolution for liquids, low resolution for solids)	Composition of chemicals, petroleum, foods, polymers, fibers
Infrared and Raman spectroscopy	Degree of cure of polymers

Computer Control and Motion Control in Manufacturing

Yoram Koren and M. Tomizuka

Computerized Numerical Control Architecture

A typical architecture of a CNC system consists of three levels, as shown in Figure 13.3.2. At the lowest level are the axial servocontrol loops and the spindle controller. These servoloops are closed at high sampling rate. The interpolator that supplies the axial position commands to the control loops is at the intermediate level of this architecture. At the highest level are the compensation algorithms for the errors of mechanical hardware deficiencies, such as machine geometry errors and thermal deformation of the machine structure. This level also includes adaptive control algorithms that adapt the machine feed and speed to the cutting tool and workpiece material to maximize machine productivity at rough cutting and maintain precision at fine cutting.

CNC Part Programs

The CNC software consists of a control program and part programs. The numerical data which are required for producing a specific part by a CNC machine is called the *part program*. The part program is arranged in the form of *blocks* of information, where each block contains the numerical data required to produce one segment of the workpiece. Each block contains, in coded form, all the information needed for processing a segment of the workpiece: the segment shape and length, its cutting speed, feed, etc. Dimensional information (length, width, and radii of circles) and the contour shape (linear, circular, or other) are taken from an engineering drawing. In NC, dimensions are given separately for each axis of motion (*X*, *Y*, etc.). Cutting conditions such as cutting speed, feed rate, and auxiliary functions (coolant on and off, spindle direction, clamp, gear changes, etc.) are programmed according to surface finish and tolerance requirements.

The part program contains the required positions of each axis, its direction of motion and velocity, and auxiliary control signals to relays. The controller generates an internal signal indicating that the previous segment is completed and that the new block of the part program should be read. The controller also operates the drives attached to the machine leadscrews and receives feedback signals on the actual position and velocity of each one of the axes.

In CNC systems the part dimensions are expressed in the part programs by integers. Each unit corresponds to the position resolution of the axes of motion and is referred to as the *basic length-unit* (BLU). The BLU is also known as the "increment size" or "bit-weight," and in practice it corresponds approximately to the accuracy of the system. To calculate the position command that the computer sends to the CNC machine, the actual length is divided by the BLU value. For example, in order to move 0.7 in. in the positive *X* direction in a system with BLU = 0.001 in., the position command is $X + 700$.

In the first generations of CNC systems, dimension were given in part programs by BLUs, as in NC. In new CNCs, however, dimensions, or desired cutter positions, are given in a normal way, as to a regular computer. The command $X - 0.705$, for example, will move the *X* axis the negative direction by 0.705 in. The resolution by which the dimension commands are given depends on the system BLU.

In addition to cutter positions, the part programmer must program the machining parameters such as tool diameter, cutting speed (n), feed (s), and depth of cut (d). The task of the part programmer is to convert the machining parameters n and s to NC control variables — spindle speed (N) and feed rate (f), as shown in Figure 13.3.3.

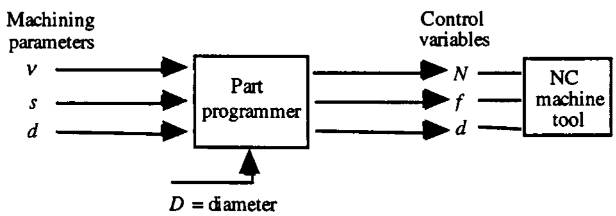

FIGURE 13.3.3 Conversion of machining parameters to control variables.

Point-to-Point and Contouring Axes of Motion

CNC systems consist of two functional types: point to point (e.g., drilling machine) and contouring or continuous path (e.g., milling machine).

Point-to-Point Systems. The simplest example of a point-to-point (PTP) CNC machine tool is a drilling machine. In a drilling machine the workpiece is moved along the axes of motion until the center of the hole to be drilled is exactly beneath the drill. Then the drill is automatically moved toward the workpiece (with a spindle speed and feed which can be controlled or fixed), the hole is drilled, and the drill moves out in a rapid traverse feed. The workpiece moves to a new point, and the above sequence of actions is repeated.

In a PTP system, this system requires only position counters for controlling the final position of the tool upon reaching the point to be drilled. The path from the starting point to the final position is not controlled. The data for each desired position is given by coordinate values. However, in high-speed drilling applications, such as the task of single-spindle drilling of an engine block, a control loop is needed to control the acceleration and deceleration of the motion. The digital signal processing technique has been used to control the settling time of the motion system for very fast PTP motion profile (1 to 4 g) so that the spindle could perform drilling operations without breaking the drill. A linear motor-based machine tool has been built and demonstrated by Anorad Corp. (Hauppauge, NY) for Ford Motor's engine machining line.

Contouring Systems. In contouring, or continuous-path, systems, the tool is cutting while the axes of motion are moving, as, for example, in a milling machine. All axes of motion might move simultaneously, each at a different velocity. When a nonlinear path is required, the axial velocity changes, even within the segment. For example, cutting a circular contour requires a sine-rate velocity change in one axis, while the velocity of the other axis is changed at a cosine rate.

In contouring machines, the position of the cutting tool at the end of each segment together with the ratio between the axial velocities determines the desired contour of the part, and at the same time the resultant feed also affects the surface finish. Since, in this case, a velocity error in one axis causes a cutter path position error, the system has to contain continuous-position control loops in addition to the end point position counters. Consequently, each axis of motion is equipped with both a position loop and a position counter. Dimensional information is given in the part program separately for each axis and is fed to the appropriate position counter. Then, an *interpolator,* in the controller, determines the proper velocity commands for each axis in order to obtain the desired tool feed rate.

Interpolators. In contouring systems the machining path is usually constructed from a combination of linear and circular segments. It is only necessary to specify in the part program the coordinates of the initial and final points of each segment and the feed rate. The operation of producing the required shape based on this information is termed *interpolation,* and the corresponding software algorithm in CNC is the *interpolator.* The interpolator coordinates the motion along the machine axes, which are separately driven, to generate the required machining path. The two most common types of interpolators are the linear and circular. Parabolic interpolators are also available in a few CNC systems which are used in the aircraft industry.

Linear interpolator: The ability to control the movement along a straight line between given initial and final coordinates is termed *linear interpolation*. Linear interpolation can be performed in a plane (two dimensional), using two axes of motion, or in space (three dimensional), where the combined motion of three axes is required. In this chapter only two-dimensional linear interpolators are discussed. To illustrate the interpolator function, consider a two-axis system, where a straight cut is to be made. Assume that the X axis must move p units at the same time that the Y axis moves q units. The contour formed by the axis movement has to be cut with a feed rate of V length-units per second (e.g., mm/sec). The numerical data of p, q, and V are contained in the part program and are fed into the interpolator. The interpolator then generates two velocity signals V_x and V_y, where

$$V_x = \frac{pV}{\sqrt{p^2 + q^2}}$$

and

$$V_y = \frac{qV}{\sqrt{p^2 + q^2}}$$

The position reference inputs to the axial control loops are

$$R_x = V_x t$$

and

$$R_y = V_y t$$

As seen, the two-dimensional linear interpolator supplies velocity commands simultaneously to the two machine axes and maintains the ratio between the required incremental distances.

Circular interpolator: The two most common interpolators in CNC systems are linear and circular. The *circular interpolator* eliminates the need to define many points along a circular arc. Only the initial and final points and the radius are required to generate the arc. The circular interpolator divides the arc into straight lines with a contour error smaller than one BLU. It operates on an iterative basis, where

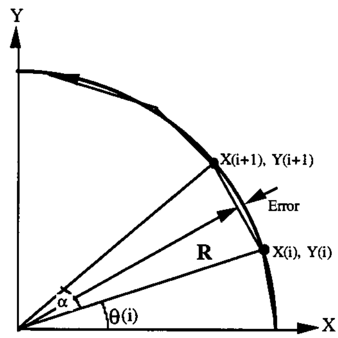

FIGURE 13.3.4 A circular interpolator divides the arc to straight lines.

for a given point, $x(i)$,$Y(i)$, a small incremental angle is added to calculate the next point, as shown in Figure 13.3.4.

Both the linear and circular interpolators are based on generating incremental positions every T seconds, where T is the sampling period of the CNC system. Typically, CNC systems operate on one common sampling period for both the interpolator and the control loops. In these systems a typical T may be between 1 and 10 msec. Some other CNC systems utilize a separate computer for the interpolation and another one or several microprocessors for closing the control loops.

Motion Control Systems

Control Loops. A typical closed loop of a CNC machine is shown in Figure 13.3.5. The computer compares the command and the feedback signals and gives, by means of a digital-to-analog converter, a signal representing the position error of the system, which is used to drive the DC servomotor. The feedback device, which is an incremental encoder in Figure 13.3.4, is mounted on the leadscrew and supplies a pulsating output. The incremental encoder consists of a rotating disk divided into segments, which are alternately opaque and transparent. A photocell and a lamp are placed on both sides of the disk. When the disk rotates, each change in light intensity falling on the photocell provides an output pulse. The rate of pulses per minute provided by the encoder is proportional to the revolutions per minute of the leadscrew.

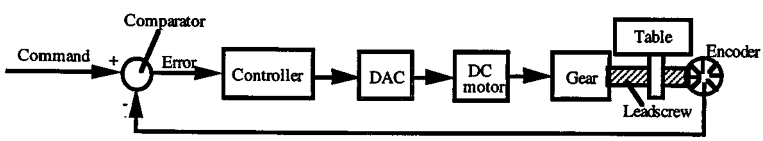

FIGURE 13.3.5 Closed loop of CNC.

Normally, the motor is mechanically coupled to the load via a drive mechanism. For example, robot links and positioning tables are loads. It is usually a good assumption to ignore the dynamics in the current feedback loop and regard the command current as the actual current, which sets the torque input to the mechanical portion of the system. Notice that the inertia and bearing friction of the motor must be considered as a part of the mechanical portion of the motion control system, which strongly influences the design of the velocity and position loop feedback controllers. These controllers can be implemented in either analog or digital form. In recent years, digital implementation on microprocessors and/or DSPs has become popular, and digital velocity and position controls are often referred to as *digital servo.* The current feedback is usually built into the "drive" (power amplification system). While feedback signals for the velocity and position controllers are usually obtained from the motor, the velocity and position of the load, e.g., a positioning table, at the opposite end of the drive mechanism are the quantities of our ultimate concern. Closed-loop control schemes based on the motor velocity and position are sometimes called semi-closed-loop control schemes. The velocity and position of the load must be fed back for full closed-loop control. Typical drive mechanisms are ballscrews and various types of gears. Several manufacturers provide so-called direct drive (DD) motors, which are capable of delivering large torques but with a significantly reduced maximum speed. DD motors may eliminate drive mechanisms. However, they are heavy and may not always be the best solution depending on applications. Common sensors for positions are potentiometers and shaft encoders, and those for velocities are tachogenerators and frequency-to-voltage convertors (FVC), the input to which is encoder pulses. In digital servos, encoders are popular for measuring positions, and velocities are either estimated from encoder pulses or are obtained from FVCs.

Traditional velocity and position loop feedback controllers are of PID (proportional plus integral plus derivative) type. The output of a PID controller is

$$u(t) = k_p e(t) + k_i \int_0^t e(\tau)\, d\tau + k_d \frac{de(t)}{dt}$$

where e(t) is the error, $u(t)$ is the controlling input, and k_p, k_i, and k_d are, respectively, proportional, integral, and derivative control gains. The above equation represents the PID control law in the continuous-time (analog) form. In digital control, the PID control law is implemented in a discretized form. A typical discrete-time (digital) PID control law is

$$u(k) = k_p e(k) + k_i T \sum_{j=0}^{k} e(j) + \left(\frac{k_d}{T}\right)\left[e(k) - e(k-1)\right]$$

where k denotes the kth sampling instance and T is the sampling period. In position loop feedback control, e and u correspond to the positioning error and the velocity command, respectively. For the velocity loop controller, they are the velocity error and the current command. Another popular linear controller is the lead/lag compensator. Input-output (I/O) interfaces include analog-to-digital convertors (A/D), decoders for processing encoder pulses, and digital-to-analog convertors (D/A). Motion control systems can be built from components and programmed with custom software, they can be purchased as plug-in boards for various buses, or they can be purchased as packaged systems.

PID and lead/lag compensators are simple and utilized in many applications. However, they alone might not be adequate for problems where the performance requirements are stringent. Extreme care must be taken during the design of a closed-loop control system. By increasing the magnitude of the feedback signal (e.g., more pulses per one revolution of the leadscrew), the loop is made more sensitive. That is known as increasing the open-loop gain. Increasing the open-loop gain excessively may cause the closed-loop system to become unstable, which obviously should be avoided.

The basic nature of "feedback" control is that the control action is based on the error. When the reference input for the position loop is fixed, the integral action may assure zero error at the steady state. In tracking control, however, the position command is continuously varying, which combined with the dynamics of the closed-loop system makes tracking errors always remain. For example, in contouring of a circular arc in machining, the command position signal for each motion control axis is sinusoidal. Such an operation is essentially a test of the frequency response of the motion control axis. The controller is normally tuned so that the frequency response gain is close to but is not exactly 1 in the operating range. In high-speed contouring operations, corresponding frequencies are high, and the gain is normally below unity. Then, the actual diameter is slightly smaller than the desired diameter. This consequence is often called the radial reduction error. Such errors may be reduced by applying the disturbance observer scheme to the position loop. However, the disturbance observer is still a feedback controller, and one sampling time delay in digital implementation further diminishes its effectiveness.

The following discussion describes some principles of motion control systems for PTP machines and contouring machines.

Control Loops for Point-to-Point Systems. The control loops of PTP systems are designed to control the position of the machine tool axes. Each axis is *separately* driven and should follow the command signal. The system design starts by selecting the type of control: open loop or closed loop, a decision which depends on the required specifications of the NC system and economy. Open-loop controls use stepping motors as the drive devices of the machine table. The drive units of the stepping motors are directly fed by the controller output pulses. The selection of the appropriate motor depends on the maximum torque, required velocity, and step size in the system. Stepping motors can be implemented on small-sized PTP systems in which the load torque is small and constant. Larger PTP machines and contouring systems utilize closed-loop control systems.

In PTP systems each axis is driven separately at the maximum allowable velocity. This velocity depends on the drive type and on the mechanical structure of the particular machined or manufacturing system. In order to avoid large overshoots the velocity is decelerated before the target point in which the tool starts to operate (e.g., to drill). Since the path between the points is insignificant, *the deceleration is accomplished in each axis separately.*

In practical systems the deceleration is accomplished by three stages. A typical three-stage deceleration diagram of one axis of the table is given in Figure 13.3.6. The table moves at rapid velocity V until reaching a distance L_1 from the target point, where the table is instructed to move at smaller velocity V_1. After a time delay, which depends on the system inertia, the table moves at a new velocity V_1 until reaching a distance of L_2 units from the target point, where again the velocity is reduced to V_2. When the table is at a distance of L_3 units before the target point, the velocity is reduced once more and the table "creeps" toward the final point at very low velocity V_3, and subsequently stops.

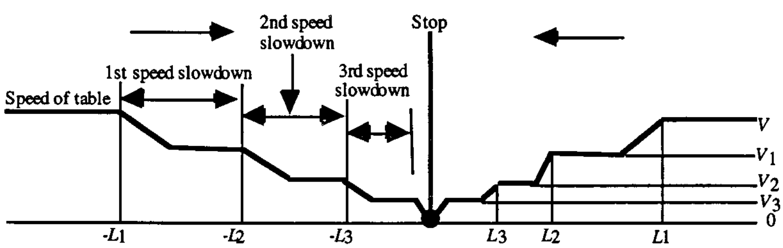

FIGURE 13.3.6 Deceleration procedure in a PTP system.

Control Loops for Contouring Systems. The control in CNC contouring systems operates in closed loops, which compare the command pulses from the interpolator with the feedback signal from the encoder or the resolver. The most sophisticated design applies to the closed-loop control of contouring systems. In the design of these loops, the transfer function of each element must first be determined; the system is then set up in block diagram form, and finally the loop gain is established based on performance analysis. The transfer function of each element is based upon its mathematical model. In establishing the mathematical model the engineer is faced with a compromise between *accuracy* and *complexity,* on one hand, and *approximation* and *simplicity,* on the other. In this section we shall discuss simple models, in which the principles of design can be readily illustrated.

The control loops of contouring systems are usually of the closed-loop type as shown in Figure 13.3.7. They use two feedback devices: a tachometer that measures the motor speed and is included in the drive unit and a position feedback transducer which is capable of also measuring the axis velocity (such as an encoder, resolver, or inductosyn). In encoder-based systems the encoder is mounted on the leadscrew and emits pulses; each pulse indicates a motion of 1 BLU of axis travel. Therefore, the number of pulses represents position and the encoder pulse frequency is proportional to the axis velocity.

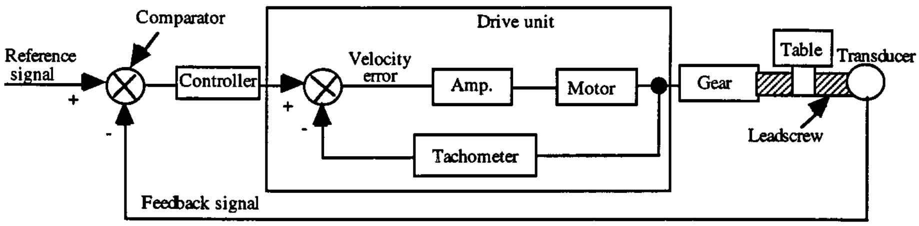

FIGURE 13.3.7 Control loop of contouring system.

Error Sources in CNC Systems

Despite the high precision of the CNC equipment, it still has position and contour errors. Error sources in CNC machines are classified into three categories:

1. Mechanical Hardware Deficiencies
    - Orthogonality of machine axes
    - Straightness of the machine axes
    - Thermal deformation of the machine structure
    - Backlash in the gears and leadscrew nuts
    - Uneven leadscrew pitch
    - Friction in moving components (leadscrews, guideways, etc.)
2. Cutting Process Effects
    - Tool deflection due to cutting forces
    - Large vibrations and chatter
    - Tool wear
    - Workpiece thermal deformations
    - Workpiece deformations due to cutting forces
3. Controller and Drive Dynamics
    - Disturbances due to cutting forces
    - Disturbance due to friction forces in the machine guideways
    - Contour errors caused by tracking errors of nonlinear contours (e.g., circles)
    - Contour errors caused by mismatch of equivalent parameters in the axial controllers
    - Corner errors in contouring operations

In general, errors of the first type, such as machine geometry errors and thermal deformation of the machine structure, can be compensated for by the machine controller. To further understand the measurement of errors, the next section will examine the metrology for precision engineering.

References

Ohnishi, K., Matsu, N., and Hori, Y. 1994. Estimation, identification, and sensorless control in motion control system, *Proc. IEEE,* 82(8), 1253–1265.

Tomizuka, M. 1993. On the design of digital tracking controllers, *ASME J. Dyn. Syst. Meas. Control,* 115(2), 412–418.

Metrology and Precision Engineering

Kam Lau

Introduction

Precision engineering in manufacturing generally refers to the engineering processes of achieving tighter tolerances or dimensional accuracy of a design. The process of precision engineering begins from the part *design*, part *fabrication*, and, finally, part *inspection*. This section addresses some of today's precision engineering considerations in fabrication and inspection.

Factors in Precision Engineering

The fabrication process is a material transformation process that determines the final part dimensions through material forming, removal, or insertion. Precision in the fabrication generally refers to the dimensional "repeatability and accuracy" of a part transformed under such a process. The factors to be considered to ensure good "precision" are (1) errors of the *machine* system(s), (2) the *environment* in which the machining is being performed, (3) the *manufacturing process*, and (4) the *instrumentation* to verify the performances of the factors.

Machine Errors. Machine errors can be classified into geometric, thermal-induced, dynamic, and structural errors. Geometric errors are errors related to the undesirable machine geometry caused by nonorthogonality (squareness error) of axes, linear positioning (or scale) error, reversal (or backlash) error, straightness error, pitch, yaw, and roll errors of a machine axis during a linear move. If the machine is equipped with a rotary table, geometric error would include axis wobble, rotational positioning error, eccentricity, and parallelism errors.

Geometric Errors: Geometric errors are much better understood in the machine industry than the thermal, dynamic, and structural errors. The techniques and instruments to check for the geometric error are very well established and are commonly applied in the machine tool industry. However, for small- to medium-size machines, geometric errors constitute only about 25% or less of the total manufacturing error. For larger machines, the percentage can go as high as 50%. As such, just knowing the geometric errors of a machine system may not be adequate in the realization of the total manufacturing error.

Traditional devices for measuring different types of geometric errors are the mechanical square, straightedge, autocollimator, electronic level, step gauge, optical polygon, and dial indicator. In some instances, the combined use of these devices is needed. Newer instruments that can offer faster and more precise measurements are the laser interferometer system, the 5-D laser interferometer system, and the telescopic ballbar.

Thermal Errors: Thermally induced machine error is considered one of the key factors affecting the accuracy of a machine tool. Thermally induced errors arise as a result of nonuniform heat generation within the machine such as in motors, bearings, and guideways; heat generated during the cutting process; and the coolant effect as well as the environmental effect resulting in the uneven growth of the machine structure. Thermal error can contribute as much as 50% of the total manufacturing error in small- to medium-size machines. However, this effect has been largely unrecognized or often ignored by the machine tool industry until recently.

There are basically two alternatives to monitor the thermally induced errors — intermittent and continuous monitorings. Intermittent monitoring generally involves the use of a touch probe or dial indicator(s) to periodically measure against one or multiple fixed positions as the machine is going through a thermal exercise. In the case of the dynamic spindle thermal study, an artifact such as a sphere or a rod can be mounted on the spindle while it is running at a certain speed. At some elapsed time (e.g., 1-min interval), the spindle will stop momentarily and the machine will reposition the artifact to the dial indicator(s) to check for the repeatability. Any deviations from the initial position (i.e., cold-start position) are considered as thermal growth due to the spindle warm-up. A typical spindle thermal test takes about 4 to 8 hr to complete. A similar procedure can be applied for the environmental and servomotion-related thermal growth measurements.

The continuous thermal growth monitoring generally involves the use of some noncontact sensors such as capacitance gauges, optical sensors, or inductance sensors arranged in the similar fashion as above. The benefits of continuous monitoring are that there is no interruption of the spindle dynamic and that the machine positioning repeatability does not necessarily interfere with the results. Furthermore, the noncontact nature generally allows the spindle to operate at any speeds, thus giving a much broader range of thermal assessment.

Laser interferometer systems are also used occasionally to monitor thermal growth effects. Because of its ability to identify the growth along the entire axis, the laser measurement is better suited for monitoring the growth of a linear-scale system caused by internal or external heat sources. Linear-scale error caused by axis movement is about one fifth that of the spindle thermal in medium to small machines. It is even less in larger machine since the heat dissipation effect in larger machine is much more effective.

Dynamic Errors: Dynamic errors here refer to those caused by the motions of a CNC machining center. These includes the servo-gain mismatch, servo-stick-slip, servo-oscillation, controller error, etc. Errors pertaining to tool chattering, structural deformation (caused by machine carriage acceleration and deceleration), machine/part deadweight, etc. are considered structural errors and are discussed in the later section.

Servo-gain mismatch is often caused by the electrical (or computer) gain setting of one axis not matching the other. The problem is not severe when the machine is primarily used for static positioning purposes, such as drill, boring, etc.; however, it can be a problem in precision contour milling when two or more axes are to be used in synchronization with each other. The magnified elliptical error is caused by one axis responding faster than the other in reaching the commanded positions. Servo-gain mismatch can easily be corrected in a routine machine maintenance.

Structural Errors: Structural errors include tool-chattering error and structural deformation error (due to acceleration and deceleration of the machine carriage, the deadweight distribution, the cutting force, etc.).

Tool chattering generally affects the machinability and surface finish of the workpiece, not the dimensional accuracy. Structural deformation caused by the acceleration and deceleration of the machine carriage and workpiece is rather insignificant for quasi-static positioning and slow-speed contouring. However, the error can be significant if the contouring speed is high. Another major contributor to the structural error is the cutting force. An excessive amount of cutting force can cause the spindle axis, the tool, the fixture, and the part to deform.

A good device to gauge the potential structural error is a compliance system. A basic compliance system consists of a load cell and a dial indicator, which are set up between the machine table and the spindle. The table is programmed to move in small increments (e.g., 5 mm) in either directions of the load cell. The readings from the load cell (force) and the indicator (actual displacement of the table) are then recorded at every increment. A compliance chart can then be obtained by plotting the forces (F) against the differences between the actual and commanded displacements (DD) of the table (i.e., F vs. DD).

Another type of structural error for large machines is from the machine foundation error. Most large machine bases are built in sections. These sections are then aligned, assembled, and anchored together to a common reinforced concrete foundation. As such, the foundation becomes part of the machine structure and the accuracy and repeatability of the machine are therefore heavily dependent on the stability of the foundation. It is not unusual to find the performance of the machine degraded as a result of floods and earthquakes, and loosening of anchor supports due to prolong use or lack of maintenance.

Errors Introduced by Environmental Effects. For the manufacturing plants that have little or no control of the plant environment, environmental effects can be very significant sources of errors. Two of the most dominant environmental effects are thermal and vibration effects.

Thermal Errors: For many nontemperature-controlled manufacturing plants, it is not unusual to observe a total temperature swing of 20°F throughout a day of operation. This wide fluctuation of environmental temperature can cause significant accuracy and repeatability problems to the machining systems, as well as the workpieces. Even in a somewhat controlled environment, thermal errors can still be a major problem. For instance, if the machine is located in the vicinity of a frequently operated bay door, where there is a substantial temperature difference between the plant temperature and the outdoor temperature, or if it is placed next to a heat source such as a welder, an electric blower or exhaust, a hydraulic pump, or under direct sunlight, the heat source can still cause tremendous localized thermal distortion of the machining system.

In most cases, one can reduce the localized environmental thermal effect by isolating the heat sources with simple panel shielding or by relocating the machines. If an overall plant temperature control is unachievable or impractical for economic reasons, one may consider (1) applying localized thermal control of certain key machining systems or (2) implementing computer thermal compensation techniques.

Vibration Errors: Environmental vibration error is a result of one or more external vibration sources affecting the structural stability of the machine system. This type of error can be significant if the machining system is located next to a heavy punch press operation or where frequent forklifting operation is present. Using a spindle analyzer or a laser interferometer system to access the amount of the environmental vibration error is common.

Other Errors: Other types of environmental errors are the interference or instability of the electric power source, the factory-supplied pneumatic and hydraulic pressures, air pressure, and humidity, etc. These types of errors may be of lesser magnitude than the above; however, it is always good practice not to underestimate their potential effects in any precision manufacturing considerations.

Errors Introduced by the Manufacturing Process .

The magnitudes of these types of errors are very much dependent on the process control and manufacturing practices implemented by each individual plant. In general, the concerns in this area are the effects of the coolant on the workpiece and the machine structure, the pallet and the fixture, the repeatability of the pallet and tool changer, and the tool deflection in manufacturing.

Coolant Effects: Eighty percent of machining uses coolant, sometimes referred to as *wet machining.* In most cases, the coolant temperature is not controlled. For smaller machining systems, a small coolant tank can be found next to the machine. For larger machines, a large coolant tank can be found underneath the ground. The main purposes of the coolant in machining are (1) for lowering the cutting temperature and (2) for chip removal. As the coolant is being recycled during the machining process, its temperature gradually warms up. This significant increase in temperature affects the workpiece dimensions since the workpiece temperature before machining is generally at room temperature. Similarly, the pallet dimensions are also affected.

Tool and Fixture: The conditions of the pallets, tools, and fixtures often govern the repeatability of a manufacturing system. These components should be checked routinely for wear and chipping. Fixtures and clamping devices should be routinely checked to ensure that proper clamping forces can be applied and that contact surfaces are in good condition. Many of these checks can be accomplished visually or by performing dial indicator repeatability checks.

When excessive cutting force is expected, it is necessary to consider the maximum possible amount of the tool, workpiece, and fixture deflections resulting from the force. The amount, if it exceeds the manufacturing tolerance, should be reduced by either reducing the depth of cut or the feed rate or by strengthening the tool and fixture. A compliance system is a good qualifier to measure the tool deflection under load.

Instrumentation and Inspection in Precision Engineering

This section introduces the instrumentation often used in precision engineering.

CNC Machining Performance Evaluation

The American National Standards Institute developed *ANSI B5.54 Standards for CNC Machining Center Performance Evaluation* for the purpose of providing detailed guidelines for machine tool users and developers to evaluate and compare the performances of CNC machining systems. The *Standards* delineates procedures for the measurements of the machine and environmental errors as stated in the above, as well as choices of conventional and state-of-the-art instruments used in doing the measurements. Experience indicates that these measurements are useful not only for performance evaluation, but also for better understanding of the sources of errors. This information is crucial for preventive maintenance and for accuracy enhancement of CNC machine systems.

Some of the key techniques introduced by the *B5.54 Standards* are the telescopic ballbar measurement, spindle dynamic and thermal drift measurement, laser diagonal measurement, and the 1-day test.

Instrumentation and Metrology

Telescopic Ballbar. The telescopic ballbar test is gradually becoming one of the most powerful and convenient tests for CNC machining center evaluations. As is shown in Figure 13.3.8, a telescopic ballbar consists of a spring-suspended reed having two spheres attached to both ends. The spheres are allowed to move relative to each other with a limited travel, e.g., 3 to 4 mm. Inside the reed is a displacement sensor, known as an LVDT (linearly variable displacement transducer), that measures the relative move of the spheres. The output of the LVDT is connected to a computer.

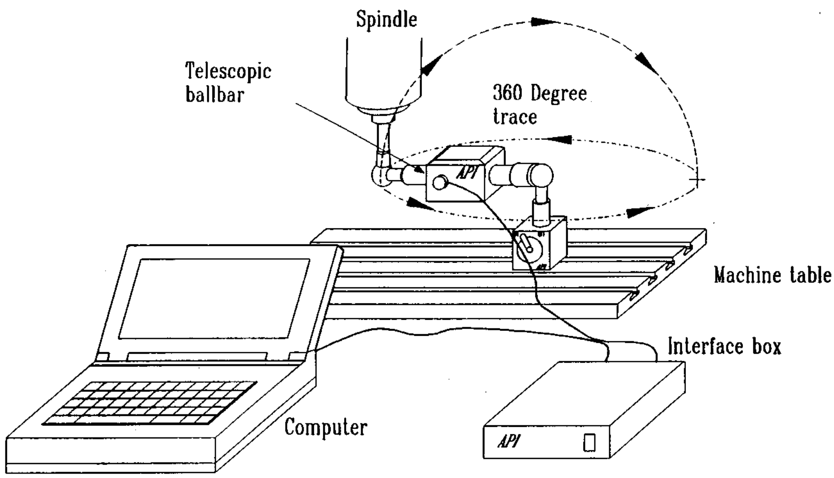

FIGURE 13.3.8 The telescopic ballbar.

To use the telescopic ballbar, one of spheres is magnetically attached to the magnetic socket which is installed on the machine spindle; the other is magnetically attached to a second magnetic socket mounted on the table. The machine is then programmed to contour in a circle in the *X-Y*, *X-Z*, or *Y-Z* plane, as shown in Figure 13.3.8. As the contouring proceeds, any errors of the machine in contouring will be translated into relative moves picked up by the LVDT and instantaneously recorded by the computer. After the machine has stopped, the computer then plots the errors on a chart for analysis. Errors such as backlash, axes nonsquareness, scale mismatch, servo-gain mismatch, cyclic motion, stick-slip, and certain straightness errors can be readily detected by using a telescopic ballbar.

Laser Interferometer. A laser interferometer system can produce measurements pertaining to the linear, angular, straightness, and squareness errors of a machine system. To many in the manufacturing industry, laser interferometer measurements are considered as the primary calibration standards. However, it is important that when using a laser interferometer, the proper procedures are followed. Precautions should be taken when using a laser interferometer system: avoid setting up the laser in areas where large air turbulence is present; set up the laser to minimize the cosine error and Abbe error; beware of the environmental effects (i.e., temperature, pressure, and humidity) to the laser accuracy; and understand the material parameters (i.e., material temperature, coefficient of expansion, etc.) to the measuring accuracy. In the event that these effects are significant, it is necessary to compensate for them in the final results.

Setting up a laser interferometer system for machine measurement can be very tedious and sometimes frustrating. It is not unusual to take 2 to 3 days to measure all linear, straightness, angular, and squareness of a mid-size three-axis CNC machine (referred to as 21 parameters measurement). A new laser system known as the 5-D laser interferometer system, developed by Automated Precision, Inc. (Gaithersburg, MD), has the capability of measuring five degrees of freedom (i.e., *X*-, *dY*, *dZ*, pitch, and yaw) simultaneously in one single setup. In contrast, the 5-D system is able to cut down the measuring time to 3 to 4 hr.

A laser interferometer system should not be confused with an alignment laser system. A laser interferometer system works on the polarized light interference principle and can produce precision to within a tenth (or better) of a wavelength of light (i.e., 630 nm for an HeNe laser). By using the same measuring principle, a laser interferometer system can also generate angle measurements to within 0.1 arc-sec. An alignment laser works on beam-pointing effects and combines the measurements with a photodetector for straightness and angle measurements (not linear measurement). In this case, the laser beam is used as a straightedge. Unless the electro-optics elements are used carefully, the accuracy is much lower than that of a laser interferometer system. Although alignment lasers are commonly used in machine alignment, it should not be construed as a precision calibration standard.

Spindle Analyzer. A spindle analyzer can be used to measure several important parameters of a machining system. These include thermal growth of a machine resulting from environmental variations (i.e., temperature, vibration, etc.) and internal heat sources (i.e., motors, spindle bearings, hydraulic systems, etc.), spindle dynamic errors resulting from a worn or contaminated bearing, and machine axis repeatability.

For advanced applications, it is recommended that the sensors be of noncontact nature since contact sensors can create problem when measuring a high-speed spindle in motion. The noncontact sensors can be optical, capacitive, or inductive. They all have their advantages and disadvantages dependent on their applications. Users should consult with the manufacturers before making their selections.

Other Metrology Instruments. The above-mentioned metrology instruments represent some of the latest and most commonly used instruments in the manufacturing industry. Other instruments that are also used frequently are electronic autocollimators, electronic levels, force gauges, temperature sensors, vibration sensors, dial indicators, proximity sensors, mechanical straightedges, step gauges, precision-indexing tables, etc. One should not, however, overlook the benefits offered by some of the latest data acquisition and data analysis software packages such as the SPC (statistical process control). Of course, the proper selection and use of any instrument are keys to a better understanding of the factors in precision manufacturing.

Inspection System and Metrology

The direct-computer-controlled coordinate measuring machine (DCC-CMM) has been the dominating means for final workpiece dimensional inspection in the manufacturing industry since the early 1980s. Since then, several advanced dimensional measuring devices have also been developed and are gaining wide acceptance in industry. These are the high-speed laser tracking systems and the manually operated measuring robots. There is also a growing interest in the manufacturing industry to reduce the off-line inspection process (i.e., CMM-type applications) by performing some of the inspections on the CNC machining system. This is referred to as on-machine gauging. This section will discuss some of the key considerations related to the use of these advanced inspection systems.

Laser Tracking Interferometer System. The laser tracking interferometer system (LTS) was developed at the National Institute of Standards and Technology (NIST) in the mid-1980s for medium- to large-dimensional inspections. It was then commercialized and introduced to the industry in late 1989. Since then, the LTS has been gaining popularity and is becoming one of the dimensional measuring standards.

Through the combination of a precision dual-axis gimbal, a laser interferometer system and an optical target (i.e., a retroreflector), the laser beam is precisely directed to the target through the manipulation of the gimballed mirror via the control computer. As the beam is sent back to the laser by the target, it is partially deflected to a dual-axis photodetector. The beam position is then interpreted by the computer. As the target moves, the photodetector immediately generates an error signal to the computer which drives the mirror to ensure the beam stays locked onto the target. While it is tracking, the computer acquires the laser measurement and the two angle measurements (a and b angles) of the mirror and computes for the three-dimensional position of the target.

The advantages of the LTS are that (1) it is considered one of the most accurate large-dimensional coordinate measuring devices with an accuracy of better than 10 ppm (i.e., 100 mm at 10 m); (2) it has a very measurable envelope — 25 m $\times$ 360°; (3) it can track a target moving at a speed of 4 m/sec; (4) it can sample up to 1000 samples/sec; and (5) it is compact, portable, and fully automatic. It is well suited for rapid surface scanning, jigs and fixture alignment, replacing conventional CMMs for large structural measurements, and for large CMMs, robotic devices, or CNC machining center calibrations.

The disadvantages are that (1) when used in areas where significant air turbulence is present, the system becomes less reliable; (2) the accuracy may be reduced when used in areas where heavy forklifting activities are present (since the floor foundation is part of the measuring frame); and (3) the system needs zero referencing if the interferometer beam is interrupted during measurement.

On-Machine Gauging. On-machine gauging (or in-process gauging) refers to the dimensional inspection process implemented during or after a machining cycle right on the machining system. In other words, the machining system, retrofitted with a sensor (i.e., a touch probe), is used to serve as a coordinate measuring machine while the workpiece is still on the machine. This concept certainly has merit since most CNC machining systems have control, scale, and structural integrity equal to or better than many DCC-CMMs.

The advantages of performing dimensional inspections on the same machining system are obvious: (1) it eliminates the need of moving the workpiece to a CMM; (2) it reduces the inspection cycle time; and (3) it eliminates realignment error should reworking the workpiece become necessary. The disadvantages, however, are (1) since the same machine frame is used for machining and inspection, if the machining system has any inherent errors (such as geometric errors) which affect the workpiece accuracy, it is incapable of detecting those errors in the inspection and (2) the machine will experience a large degree of thermal distortion caused by the internal heats.

In order to implement on-machine inspection, it is therefore necessary to first perform geometric accuracy evaluation of the machining system according to B5.54 Standards. All efforts should be made to ensure that the accuracy is maintained. Second, a machine thermal growth analysis should be performed to ascertain the limitation of the thermal distortion in inspection. In either case, a minimum rule of thumb of four times the accuracy tolerance should be applied. New techniques of thermal and geometric modeling and compensation of the machining system can be considered in order to achieve the inspection goals.

Other Inspection Systems. Other emerging inspection machines, such as the manually operated robotic measuring device, stereotriangulation measuring systems, and photogrammetry systems, are also gaining popularity in manufacturing. Although they may not offer the same types of accuracy and versatility as the previously mentioned system, they feature a new trend of inspection requirements — portability, agility, shop-floor hardiness, high-speed data acquisition, low cost, and powerful software capability.

References

ANSI. 1992a. Performance Evaluation of Computer Numerically Controlled Machining Centers, ANSI/ASME B5.54-1992, ASME, New York.

ANSI. 1992b. Axes of Rotation, Methods for Specifying and Testing, ANSI/ASME B89.3.4M-1985 (R1995), ASME, New York.

ANSI. 1995. Temperature and Humidity Environment for Dimensional Measurement, ANSI/ASME B89.6.2-1973 (R1995), ASME, New York.

ANSI. 1997. Methods for Performance Evaluation of Coordinate Measuring Machines, ANSI/ASME B89.4.1-1997, ASME, New York.

Slocum, A. 1992. *Precision Machine Design,* Prentice-Hall, Englewood Cliffs, NJ.

Technical Manual, Automated Precision, Inc. (API), Gaithersburg, MD.

Mechatronics in Manufacturing

Tai-Ran Hsu

Introduction

Mechatronics can be defined as: "A technology that involves the design, manufacture and production of *intelligent* products or engineering systems involving *mechanical* and *electronic* functions." Mechatronics is a melding of two English words, *mechanical* and *electronics*. This terminology was first used by Yaskawa Electronic Corp. in Japan during the 1970s. The original notion of mechatronics involved the development and automated production of consumer products such as the Canon SLR autofocus camera (Gilbert, 1992). The application of this technology was soon extended to many other consumer electronic products that included video cassette recorders and the well-known Sony Walkman radios. The rapid advances of microprocessor and microcomputer technologies in the 1980s have broadened the applications of mechatronics to all smart products and systems, ranging from common consumer products to highly sophisticated space engineering equipment. In general, mechatronics systems engineering comprises a number of engineering disciplines. Figure 13.3.9 illustrates interactions of the mechatronics systems engineering with other engineering fields.

Elements of Mechatronic Systems Engineering

Most mechatronic products or systems consist of the following three modules:

1. *The sensing and control module:* This module involves hardware that includes electronic elements, electronic circuits, and the software that provides the commands and operation logics to all components in an automatic control systems. Basic hardware components include *sensors* (position, velocity, acceleration) including tactile, optical, and voice; *actuators* (linear, rotary, voice) driven by electrical, pneumatic, or hydraulic means; *drives* (DC/AC stepper or servomotors); *encoders* (optical or magnetic) and time counters. Other major components include signal processors, amplifiers, power suppliers, and A/D or D/A converters. Control software such as PLC, PID, and other application software are written in C, C^{++}, LISP, FORTH, and RPL programming languages.
2. *Mechanical systems and kinematic linkages module:* The mechanical components in a mechatronic system provide the means to achieve the desired objectives of the product or the system. Principal mechanical components include gears, cams, chains, levers, shafts, pulleys, couplers, bearings, joints, and fasteners. These components are synthesized to construct mechanisms that perform the desired motion of end effects with accurate paths, positions, velocity, and accelerations.
3. *The microprocessors and computer interface module:* Microprocessors act as the brain of the mechatronic product or system. Principal functions of a microprocessor are to manage various instructions to all components such as transducers and actuators in the system, as well as I/O

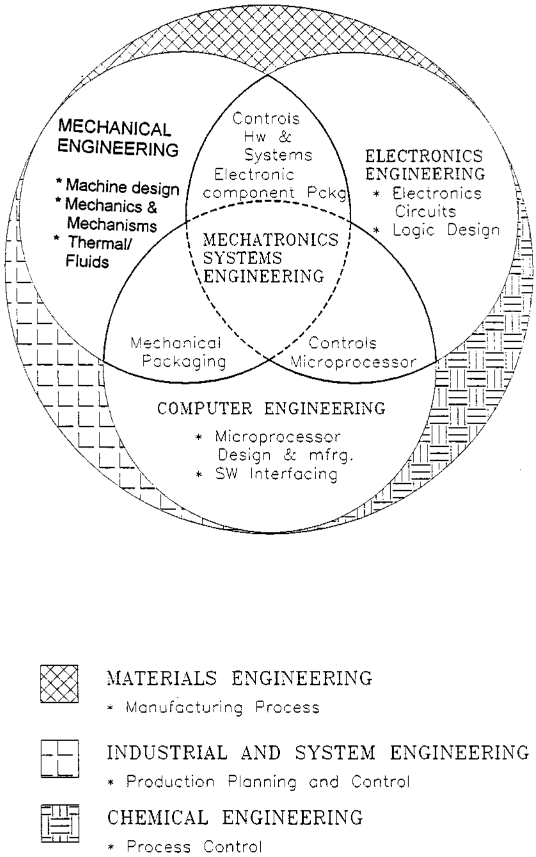

FIGURE 13.3.9 The relationship of various modules in a typical mechatronic device.

interfacing. These processors can also accept function commands from a microcomputer with specific control and applications software.

A block diagram that illustrates the relationship of various modules in a typical mechatronic system is presented in Figure 13.3.10. One may imagine that the microprocessor and microcontroller module functions as the brain, and the sensing/control module and the mechanical systems module each function like the respective sensor-nerve system and the body and limbs in the human anatomy.

References

Gilbert, M.M. 1992. Camera design is something to shoot for, *Mach. Des.*, March.
Sze, S.M. 1984. *VLSI Technology*, McGraw-Hill International Book Company, Singapore.
White, R.M. 1985. *Introduction to Magnetic Recording*, IEEE Press, New York, 72–79.
White, R. 1994. *How Computers Work*, Ziff-Davis Press, Emeryille, CA, 1994.

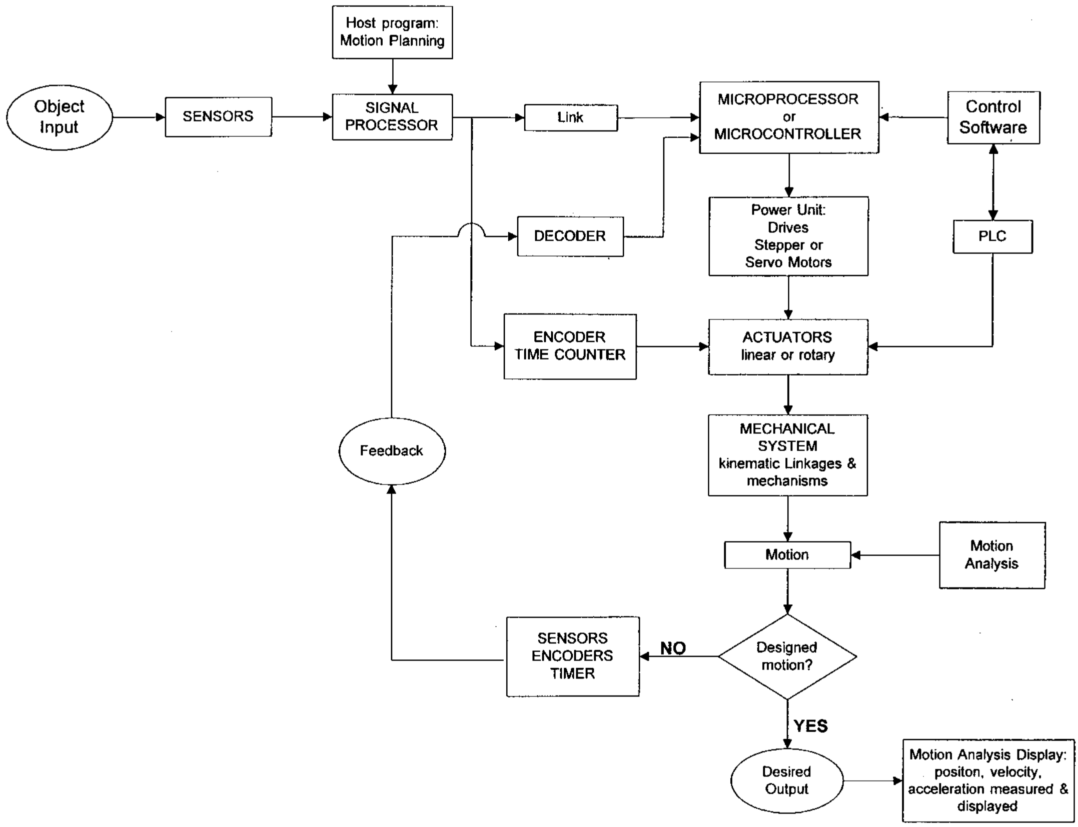

FIGURE 13.3.10 Relationship among modules of a typical mechatronic system.

13.4 Design and Analysis Tools in Manufacturing

Computer-Aided Design Tools for Manufacturing

David C. Anderson

Introduction

Computer-aided design (CAD) tools for manufacturing are computer programs that evaluate producibility of a product under development using computer models of the product and simulation models for manufacturing processes. Examples of such processes are assembly, casting, molding, stamping, forming, and inspection. The early stages of design are critical to the success of the final product since most of the final cost of a product is committed through decisions about the product's geometry and the materials that lead to the selection and planning of manufacturing processes and production facilities.

Concurrent engineering (CE) is a methodology for product development processes where the tasks are performed simultaneously based on "what if" decision-making processes. CE is a driving force behind the development of CAD tools for manufacturing.

CAD System and Manufacturing

One of the first and most prominent manufacturing applications of CAD system was the automated programming of numerically controlled (NC) machine tools. NC machines were developed in the 1950s, but their widespread use was hindered by the difficulty in writing the requisite NC programs, lists of coded cutting tool motions that directed the machine to cut the desired part. Computer programs that facilitated the generation of NC programs were among the first CAD tools for manufacturing. Early systems focused on profile cutting, generating cutter location data that described the 2D coordinates of the paths of cutting tools that would remove the material to create the final part. Three-dimensional NC programming capabilities are available in most CAD systems today. The user guides the generation of NC data by interactively selecting each surface to be machined and answering questions about tools, approach directions, and machining preferences.

Solid Models and Manufacturing

A key element of virtually all CAD tools for manufacturing is the need to interpret the CAD data according to manufacturing capabilities, requirements, and constraints. Generally, this interpretation involves determining the characteristic shapes in the CAD data related to the manufacturing process of interest and applying knowledge about the process to determine the manufacturing operations and parameters. Solid models are rigorous computer data structures that contain a complete, unambiguous representation of the nominal geometry of an object. Solid models and solid modeling operations, like the Boolean union, difference, and intersection of solids, enabled geometric computations about the design that were not possible with earlier CAD data. With solid modeling, the emphasis in CAD shifted from data for visual communication to product data models that could be used for more sophisticated computer analyses with more automation.

In solid models, geometric features can be described in a form more suitable for engineering and manufacturing applications. As a result, feature-recognition algorithms have become an important element of many CAD tools for manufacturing for translating CAD models into usable geometric data for design evaluations.

However, feature recognition has limitations because it is not possible to translate every solid model into a given set of features. Also, some applications of feature recognition require additional, nongeometric information. This information must be added to the CAD data before, during, or after the feature-recognition process. For example, part tolerance data are required for machining process planning but are not available in current CAD data. Feature-based design (FBD) was developed to overcome some of the limitations of feature recognition. Instead of deriving a features model from a solid model, FBD systems create the product model with features during the design process. Many FBD applications have been demonstrated in the area of machining, and some commercial CAD vendors have incorporated these features into their systems.

FBD also has limitations. The features that are useful for one manufacturing process may not be useful for another. For example, a model made with cavity features, such as holes, slots, and pockets, provides ready-to-use data for machining planning. However, this model is inappropriate for sheet metal bending or welding processes that require an understanding of the bends and protrusions on a part, not its cavities.

Product Data Standards and Manufacturing

The International Standards Organization (ISO) has developed the Standard for the Exchange of Product model data, STEP. In the U.S., the Integrated Graphics Exchange Specification (IGES)/PDES Organization (IPO) developed the Product Data Exchange Specification, PDES. These efforts merged and PDES was renamed Product Data Exchange using STEP, which is now the American National Standard for STEP. STEP became the international standard (ISO 10303) in March 1994.

STEP is organized as a series of "parts" that are developed and published separately. The parts are organized into numerical series: description methods (parts 11 to 20), integrated resources (parts 41 to 200), application protocols (parts 201 to 1200), abstract test suites (parts 1201 to 2200), implementation methods (parts 21 to 30), and conformance testing (parts 31 to 40). The product information is specified in a formal specification language, EXPRESS, which is documented in Part 11. Part 1 is an overview of the entire standard.

The STEP application protocols (APs) are important to CAD tools for manufacturing. These are the implementable portion of the STEP standard. Each AP draws upon integrated resources and adds specific constraints, relationships, and attributes to meet the information requirements of a particular application. It is expected that several hundred APs may be developed for many different industrial applications.

Design for "X" Tools

Many CAD tools for manufacturing belong to a class of programs described as "design for x," where "x" signifies an application area, such as design for assembly or design for castability. The phrase "design for manufacturability" (DFM) can be considered as a specialization of design for x in which all the applications are manufacturing. Generally, DFM applications are computer programs that perform computations to analyze the producibility of a product with respect to a specific manufacturing process or set of processes. The format of the design data representing the product required by the program varies greatly. The program then provides an evaluation of the suitability of the design according to this domain. The form of this evaluation also varies with each program.

DFM programs act as "manufacturing experts" that provide qualitative, and perhaps quantitative, information about potential problems in a design based on predefined knowledge about a manufacturing process. The programs emulate the process done by human experts, examining the design data and reporting any problems based on experience. Many efforts are based on expert systems technology from the field of artificial intelligence (AI). The searching is performed through the facilities of an AI language, such as PROLOG or LISP, or using an expert system "shell," a preprogrammed generic expert system. Design data are first translated into a knowledge base, the AI version of a database, containing facts and rules. In a sense, the detailed design data are made into logical data that can be processed using AI methods. The program computationally searches the design data for data patterns that match problem conditions represented in the knowledge base of the program. The problem conditions are computer representations of design data that are known to cause manufacturing difficulties for a given process. For example, a "design for injection molding" program may report that an internal corner radius in a geometric model of a part is too small and may cause problems in the mold. In some cases, the program provides a quantitative evaluation of the design, or producibility index. This provides a convenient numerical comparison between two competing designs.

References

Amirouche, F.M.L., 1993. *Computer-Aided Design and Manufacturing,* Prentice-Hall, Englewood Cliffs, NJ.

Boothroyd, G. 1994. Product design for manufacture and assembly, *Comput. Aided Des.,* 26(7), 505–520.

Laurance, N. 1994. A high-level view of STEP, *Manuf. Rev.,* 7(1), 39–46.

The National Product Data Exchange Resource Center, U.S. Product Data Association (US PRO) National Institute of Standards and Technology, "http://elib.cme.nist.gov/nipde/" (world-wide web document), 1995.

Whitney, D.E., Nevins, J.L., and De Fazio, T.L. 1989. *Concurrent Design of Products and Processes: A Strategy for the Next Generation in Manufacturing,* McGraw-Hill, New York.

Zeid, I. 1991. *CAD/CAM Theory and Practice,* McGraw-Hill, New York.

Tools for Manufacturing Process Planning

Tien-Chien Chang

Introduction

Process planning prepares a production documentation which specifies the operations and operation sequence necessary to manufacture a product. Process planning is defined as an act that determines the manufacturing operations, operation sequence, and resources required to make a product. In the process domain, there are machining process planning, welding process planning, EDM process planning, forming process planning, etc. In the product domain, there are mechanical part process planning, mechanical assembly process planning, and electronics assembly process planning. The input to a process planning system (or a human process planner) can be an engineering drawing, a CAD model, or a three-dimensional solid model. While most human planners prefer an engineering drawing on paper or in electronic form, process planning systems usually use CAD models.

The result of the process-planning activity is a "process plan" (see Figure 13.4.1), also called route sheet, operation sheet, or operation planning summary. It can be as aggregate as a list of work center identification numbers or as elaborate as a 50-page document with setup drawings, tool specifications, operation time estimates, etc.

PROCESS PLAN — ACE Inc.

Part No. S0125-F — Material: steel 4340Si
Part Name: Housing
Original: S.D. Smart Date: 1/1/89 — Changes: ________ Date: ______
Checked: C.S. Good Date: 2/1/89 — Approved: T.C. Chang Date: 2/14/89

No.	Operation Description	Workstation	Setup	Tool	Time (Min)
10	Mill bottom surface1	MILL01	see attach#1 for illustration	Face mill 6 teeth/4" dia	3 setup 5 machining
20	Mill top surface	MILL01	see attach#1	Face mill 6 teeth/4" dia	2 setup 6 machining
30	Drill 4 holes	DRL02	set on surface1	twist drill 1/2" dia 2" long	2 setup 3 machining

FIGURE 13.4.1 A process plan.

Since the information on the process plan is used in scheduling the production and controlling the machine, the production efficiency and the product quality are affected.

Manual Process Planning

Process planning involves several or all of the following activities:

- Selection of machining operations
- Sequencing of machining operations
- Selection of cutting tools
- Selection of machine tools
- Determination of setup requirements
- Calculations of cutting parameters
- Planning tool path and generation of NC part programs
- Design of jigs and fixtures

The details incorporated in a typical process plan usually vary from industry to industry. It depends on the type of parts, production methods, and documentation needs. A process plan for a tool room–type manufacturing environment typically relies on the experience of the machinist and does not have to be written in any great detail. In fact, the instruction "make as per part print" may suffice. In typical mass-production-type industries, the process-planning activity is embodied in the transfer and flow lines used for manufacturing component parts and assembly. For metal-forming-type manufacturing activities, such as forging, stamping, die casting, sand casting, injection molding, etc., the process-planning requirements are embedded directly into the design of the die/mold used, where most process-planning activity is fairly simple. A process planner must

- Be able to understand and analyze part requirements
- Have extensive knowledge of machine tools, cutting tools, and their capabilities
- Understand the interactions between the part, manufacturing, quality, and cost
- Possess analytical capabilities

Tolerance Charting

During process planning it is important to ensure that the setup and operation sequence will yield a satisfactory part. Tolerancing charting (Figure 13.4.2) has been used to help in allocating process tolerances and verifying the operation sequence. A tolerance chart analyzes one dimension at a time. In a tolerance chart, the top is the part drawing. Dimensions and tolerances are presented with the geometry. Dashed lines show the stock boundary. From the features of the part and the stock, extension lines are drawn to the body of the chart. The section below the drawing shows the critical dimensions and tolerances. These dimensions and tolerances must be satisfied after the processes are complete. Following the process sequence, each operation is listed in the third section of the chart. A line is drawn from the reference surface of a setup to the cut surface. For example, in operation 10, the raw stock boundary at the left is the reference surface. The second surface from the right-hand side is created by this operation. From the chart, one can calculate the resultant tolerances. The results are compared with the blueprint tolerance.

Although traditionally a tolerance chart is implemented on paper and through a fixed procedure, it can also be implemented in a computer. The process tolerance stack-up may be used to verify the design specification and select the appropriate processes and sequences.

Computer-Aided Process Planning

There are two basic approaches to computer-aided process planning — variant and generative. The variant approach is signified by the terminology used by the computer to retrieve plans for similar

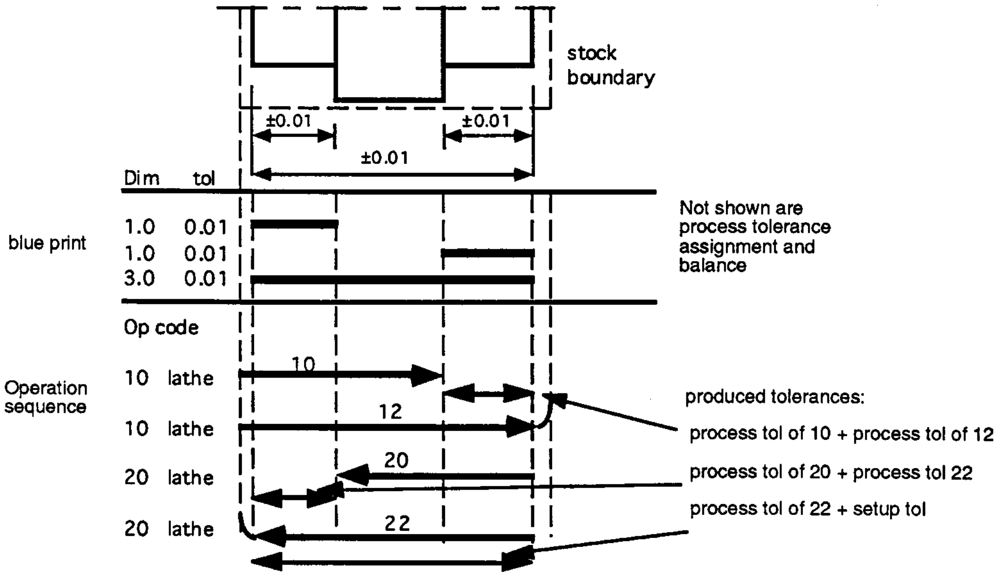

FIGURE 13.4.2 Tolerance chart.

components using table lookup procedures. The human process planner then edits the plan to create a "variant" to suit the specific requirements of the component being planned. Creation and modification of standard plans are the process planner's responsibility. The generative approach generates a plan for each component without referring to existing plans. Generative-type systems can perform many functions in a generative manner, while the remaining functions are performed with the use of humans in the planning loop.

Variant Process Planning

The variant approach to process planning was the first approach used to computerize the planning techniques. It is based on the concept that similar parts will have similar process plans. The computer can be used as a tool to assist in the identification of similar plans, retrieving them and editing the plans to suit the requirements for specific parts.

In order to implement such a concept, part coding and classification based on group technology is used as a foundation. Individual parts are coded based upon several characteristics and attributes. Part families are created of "like" parts having sufficiently common attributes to group them into a family. This family formation is determined by analyzing the codes of the part spectrum. A "standard" plan consisting of a process plan to manufacture the entire family is created and stored for each part family. The development of a variant-process-planning system has two stages: the preparatory stage and the production stage (Figure 13.4.3).

During the preparatory stage, existing components are coded, classified, and later grouped into families. The part family formation can be performed in several ways. Families can be formed based on geometric shapes or process similarities. Several methods can be used to form these groupings. A simple approach would be to compare the similarity of the part code with other part codes. Since similar parts will have similar code characteristics, a logic which compares part of the code or the entire code can be used to determine similarity between parts.

Families can often be described by a set of family matrices. Each family has a binary matrix with a column for each digit in the code and a row for each value a code digit can have. A nonzero entry in the matrix indicates that the particular digit can have the value of that row, e.g., entry (3,2) equals one

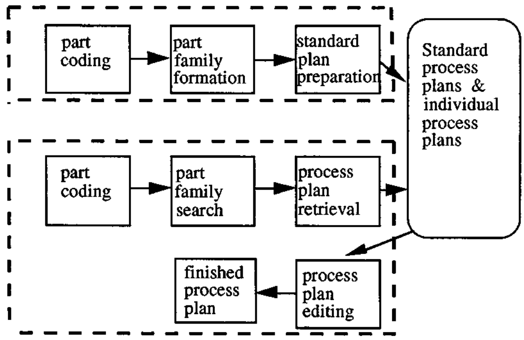

FIGURE 13.4.3 Variant-process-planning approach.

implies that a code x3xxx can be a member of the family. Since the processes of all family members are similar, a standard plan can be assigned to the family. The standard plan is structured and stored in a coded manner using operation codes (OP-codes). An OP-code represents a series of operations on one machine/workstation. For example, an OP-code DRL10 may represent the sequence center drill, change drill, drill hole, change to reamer, and ream hole. A series of OP-codes constitute the representation of the standard process plan.

Before the system can be of any use, coding, classification, family formation, and standard plan preparation must be completed. The effectiveness and performance of the variant-process-planning system depend to a very large extent on the effort put forth at this stage. The preparatory stage is a very time-consuming process.

The production stage occurs when the system is ready for production. New components can be planned in this stage. An incoming component is first coded. The code is then sent to a part family search routine to find the family to which it belongs. Since the standard plan is indexed by family number, the standard plan can be easily retrieved from the database. The standard plan is designed for the entire family rather than for a specific component; thus, editing the plan is unavoidable.

Variant-process-planning systems are relatively easy to build. However, several problems are associated with them., e.g.,

1. The components to be planned are limited to previously planned similar components.
2. Experienced process planners are still required to modify the standard plan for the specific component.
3. Details of the plan cannot be generated.
4. Variant planning cannot be used in an entirely automated manufacturing system, without additional process planning.

Despite these problems, the variant approach is an effective method, especially when the primary objective is to improve the current practice of process planning. In most batch-manufacturing industries, where similar components are produced repetitively, a variant system can improve the planning efficiency dramatically. Some other advantages of variant process planning are

1. Once a standard plan has been written, a variety of components can be planned.
2. Programming and installation are comparatively simple.
3. The system is understandable, and the planner has control of the final plan.
4. It is easy to learn and easy to use.

Generative Approach

Generative process planning is the second type of computer-aided process planning. It can be concisely defined as a system which automatically synthesizes a process plan for a new component. The generative approach envisions the creation of a process plan from information available in a manufacturing database without human intervention. Upon receiving the design model, the system is able to generate the required operations and operation sequence for the component.

Knowledge of manufacturing has to be captured and encoded into computer programs. By applying decision logic, a process planner's decision-making process can be imitated. Other planning functions, such as machine selection, tool selection, process optimization, etc., can also be automated using generative planning techniques.

A generative-process-planning system comprises three main components:

1. Part description
2. Manufacturing databases
3. Decision-making logic and algorithms

The definition of generative process planning used in industry today is somewhat relaxed. Thus, systems which contain some decision-making capability on process selection are called generative systems. Some of the so-called generative systems use a decision tree to retrieve a standard plan. Generative process planning is regarded as more advanced than variant process planning. Ideally, a generative-process-planning system is a turnkey system with all the decision logic built in. However, due to the differences among manufacturing shops, decision logics have to be customized for each shop.

The generative-process-planning approach has the following advantages:

1. Consistent process plans can be generated rapidly.
2. New components can be planned as easily as existing components.
3. It has potential for integrating with an automated manufacturing facility to provide detailed control information.

There is no fixed representation or procedure that can be identified with generative process planning. The general trend is to use a solid model CAD-based input and expert system or an object-oriented planner construct. Most of the research systems are of this type. A few commercial products can also be classified as generative.

Conclusions

Process planning is a critical function in design and manufacturing. The quality of the product and the cost of production are affected by the process plan. A process plan incorporates information on the shop capability, resource requirement, best routing, etc. In order to produce a good process plan, a planner must be knowledgeable in both the manufacturing practices and the current shop status and capabilities. At this moment most computerized planners are still based on data retrieval and database lookup. As the information technology is further developed and we have better understanding of process capabilities, integrated planning systems will evolve.

References

Chang, T.C. and Wysk, R.A. 1995. *An Introduction to Computer-Aided Process Planning Systems,* Prentice-Hall, Englewood Cliffs, NJ.

Curtis, M.A. 1988. *Process Planning,* John Wiley & Sons, New York.

Halevi, G. and Weill, R.D. 1995. *Principles of Process Planning,* Chapman & Hall, New York.

Kambhampati, S., Cutkosky, M.R., Tennenbaum, J.M., and Lee, S. 1993. Integrating general purpose planners and special reasoners: case study of a hybrid planning architecture, *IEEE Trans. Syst., Man Cybernetics*, 23(6), 1503–1518.

van 't Erve, A.H., 1992. Generative Computer Aided Process Planning for Part Manufacturing, An Expert System Approach, Ph.D. Thesis, Department of Mechanical Engineering, University of Twente, Netherlands.

van Houten, F.J.A.M. 1991. PART: A Computer Aided Process Planning System, Ph.D. Thesis, Department of Mechanical Engineering, University of Twente.

Wang, H.-P. and Li, J.K. 1991. *Computer-Aided Process Planning*, Elsevier, New York.
Zhang, H.-C. and Alting, L. 1994. *Computerized Manufacturing Process Planning Systems,* Chapman and Hall, New York.

Simulation Tools for Manufacturing

Hank Grant

Introduction

Digital simulation uses a mathematical model to represent a real or hypothetical physical system. A computer simulation model of a physical system provides a laboratory in which alternative designs can be explored and analyzed. The model, executed on the computer, is a software replica of the manufacturing system and is controlled so that the behavior of the system can be studied and analyzed. Decisions can be made concerning production alternatives. For example, adding a new lathe can be considered without disrupting the actual physical system.

Simulation depends on describing a system in terms acceptable to the computer language used. To do this, it is necessary to have a *system-state description,* which is typically characterized by a set of state variables included in the computer program that make up the simulation model.

Types of Simulation Models

Simulation models are typically classified into three types: discrete event, continuous, and combined. Simulation software has been designed to address each of these types of models.

Discrete Event. Discrete event simulation is used to model systems when there are specific events in time when the variables of the system may change in values. The mechanics of those changes must be well known and easily characterized. The behavior of the system is represented by the behavior of individual objects of interest called *entities*. The simulation model characterizes the behavior of these entities as they move through the system in simulated time.

Discrete events are points in time where the characteristics of an entity may change and where the state variables of the system may change. For example, when a customer arrives for service, the state of the system may change (number in the system, status of the server, etc.). The modeling of systems using this approach consists of developing descriptions of the events and how they cause the state variables to change and the entities to be manipulated.

The individual events may not always be predictable, and stochastic elements may be present in the operation of the system. For example, the time between arrivals of customers to the system may be a random variable. Simulation languages have many tools to support random variation in models.

A special kind of discrete event model is a *network model*. Network models use a standard set of symbols to represent the flow of entities in the system. They are graphical in nature, and are very useful communication vehicles as well as very powerful in building simulation models quickly and easily. There are several languages available that include network modeling capabilities and they are described below.

Continuous. *Continuous simulation* is an approach that is popular among engineers and economists. The main building blocks of this approach are as follows (Pidd, 1994).

Aggregated variables: Instead of a concern with individual entities, the main concern is with the aggregated behavior of populations. For example, the changing sales of a product through time.

Smooth changes in continuous time rather than focusing on individual events, where the stress is on the gradual changes which happen as time progresses. Thus, just as the graph of a variable might be smooth, the aim is to model the smooth changes of the variable by developing the suitable continuous equations.

Differential or difference equations: The model consists mainly of a set of equations which define how behavior varies through time; thus, these tend to be differential equations or, in simpler cases such as system dynamics, difference equations.

Nature does not present itself labeled neatly as discrete or continuous; both elements occur in reality. Modeling, however, as mentioned above, involves approximation, and the modeler must decide which of these approaches is most useful in achieving the desired aim of the simulation.

Combined Discrete Event/Continuous. In some cases, both approaches are needed and the result is a mixed discrete-continuous simulation. An example of this might be a factory in which there is a cooking process controlled by known physics which is modeled by continuous equations. Also in the factory is a packing line from which discrete pallets of products emerge. To model the factory will require a mixed approach.

Modeling Languages

Specifically designed computer simulation languages provide many features for managing the updating of the state variables and advancing time. They also provide features for recording system performance statistics and for generating random numbers to introduce system randomness.

The lowest level of computer language typically used is FORTRAN or BASIC. This requires that the entire simulation model be coded, which is labor-intensive. High-level languages, such as SLAM, SIMSCRIPT, and GPSS, facilitate simulation because they provide subroutines for time advancement, entity maintenance, and statistic collections. Higher-level simulation languages are designed for special purposes; MAP/1, SPEED, and MAST are three designed for the simulation of manufacturing systems.

Some simulation languages can produce animations. This permits the simulation to be illustrated graphically on a computer terminal so that the analyst can see the system in action and observe its interactions and behavior, a visual function beyond the scope of standard reporting technique. For example, TESS (a software program) provides animation, as well as model-building and output analysis capabilities, for the SLAM simulation language.

The following discussion by Banks (1994) provides an overview of the primary languages available.

Applications of simulation exist in many arenas such as manufacturing, material handling, health services, military decision support, natural resources, public services, transportation, and communications, to mention a few.

These simulation applications are usually accomplished with the use of specially developed simulation software. This tutorial describes the software in two categories. The first of these is software for general purposes. This type of software can solve almost any discrete simulation problem. In this section, five products, GPSS/H™, GPSS/World™, SIMAN V′, SIMSCRIPT II.5′, and SLAMSYSTEM′, will be discussed to provide a feel for this type of software.

GPSS/H. GPSS/H is a product of Wolverine Software Corporation, Annandale, VA (Smith and Crain, 1993). It is a flexible, yet powerful tool for simulation. It provides improvements over GPSS V that had been released by IBM many years earlier. These enhancements include built-in file and screen I/O, use of an arithmetic expression as a block operand, interactive debugger, faster execution, expanded control statement availability, and ampervariables that allow the arithmetic combinations of values used in the simulation. The latest release of GPSS/H is version 2.0. It added a floating point clock, built-in math functions, and built-in random variate generators. Options available include Student GPSS/H, Personal GPSS/H within the 640K memory limit, and GPSS/H 386 providing unlimited model size.

GPSS World. GPSS World, from Minuteman Software, is a complete redesign of GPSS/PC™ (Cox, 1992). It is designed as a high-power environment for simulation professionals. It includes both discrete and continuous simulation. Its features include interactivity, visualizability, and configuration flexibility. It utilizes 32-bit computing, virtual memory, preemptive multitasking, symmetric multiprocessing, and distributed simulation. Highlights include drag-and-drop model building, 512 megabytes of virtual memory for models, point-and-shoot debugging, an embedded programming language, built-in probability distributions, multiple data types, and many other improvements to GPSS/PC.

The GPSS World family is a set of three software products including:

1. GPSS World is the center of the family. This self-contained modeling environment includes local Simulation Server™ capabilities.
2. Simulation Server provides simulation services on a remote networked computer. It does not include a model-building user network.
3. Simulation Studio provides hierarchical modeling and user-drawn simulation capabilities.

There is an enhanced memory version of GPSS/PC that is also available. It allows access of up to 32 megabytes of memory.

SIMSCRIPT II.5. SIMSCRIPT II.5 from CACI Products Company is a language that allows models to be constructed that are either process oriented or event oriented (Russell, 1993). The microcomputer and workstation version include the SIMGRAPHICS animation and graphics package. SIMSCRIPT can be used to produce both dynamic and static presentation-quality graphics such as histograms, pie charts, bar charts, levels of meters and dials, and time plots of variables. Animation of the simulation output is also constructed using SIMGRAPHICS. SIMGRAPHICS can be used also to produce interactive graphical front ends or forms for entering model input data. An input form may include such graphical elements as menu bars with pull-down menus, text or data boxes, and buttons that are clicked on with a mouse to select an alternative. The graphical model front end allows for a certain set of modifications to the model to be made without programming, facilitating model use by those who are not programmers.

SIMAN V. SIMAN V from Systems Modeling Corporation is a general-purpose program for modeling discrete and/or continuous systems (Glavach and Sturrock, 1993; Banks et al., 1995). The program distinguishes between the system model and the experiment frame. The system model defines components of the environment such as machines, queues, and transporters and their interrelationships. The experiment frame describes the conditions under which the simulation is conducted, including machine capacities and speeds and types of statistics to be collected. "What-if" questions can usually be asked through changing the experiment frame rather than by changing the model definition. Some important aspects of SIMAN V are as follows:

1. Special features that are useful in modeling manufacturing systems include the ability to describe environments as work centers (stations) and the ability to define a sequence for moving entities through the system.
2. Constructs that enable the modeling of material-handling systems including accumulating and nonaccumulating conveyors, transporters, and guided vehicles.
3. An interactive run controller that permits break points, watches, and other execution control procedures.
4. The ARENA environment that includes menu-driven point-and-click procedures for constructing the SIMAN V model and experiment, animation of the model using Cinema, the input processor that assists in fitting distributions to data, and the output processor that can be used to obtain confidence intervals, histograms, correlograms, and so on. (More aspects of the ARENA environment are discussed later.)
5. Portability of the model to all types of computers.

SLAMSYSTEM. SLAMSYSTEM, from Pritsker Corporation, is an integrated simulation system for PCs based on Microsoft Windows™ (Pritsker, 1986; O'Reilly, 1993). All features are accessible through pull-down menus and dialog boxes and are selected from the SLAMSYSTEM Executive Window. A SLAMSYSTEM project consists of one or more scenarios, each of which represents an alternative system configuration. A project maintainer examines the components of the current scenario to determine if any of them have been modified, indicates whether or not tasks such as model translation should be performed, and allows the user to accomplish these tasks before the next function is requested. SLAMSYSTEM allows multiple tasks to be performed in parallel while the simulation is operating in the background.

Some of the features of SLAMSYSTEM are as follows:

1. Models may be built using a graphical network builder and a forms-oriented control builder, or text editor. When using the first method, a network symbol is selected with the mouse, then a form is completed specifying the parameters for that symbol. The clipboard allows many other operations such as grouping one or more symbols and placing them elsewhere on the network.
2. Output analysis includes a "report browser" that allows alternative text outputs to be compared side by side. Output may be viewed in the form of bar charts, histograms, pie charts, and plots. Output from multiple scenarios can be displayed at the same time in bar chart form. By using the Windows environment, multiple output windows can be opened simultaneously.
3. Animations are created under Windows using the facility builder to design the static background and the script builder to specify which animation actions should occur when a particular simulation event occurs. Animations can be performed either concurrently or in a postprocessing mode. Two screens can be updated simultaneously and up to 225 screens can be swapped into memory during an animation.
4. SLAMSYSTEM was designed to be used in an integrated manner. For example, historic data may be read to drive the simulation. CAD drawings may be loaded. Output charts and plots created by SLAMSYSTEM may be exported via the clipboard to other applications.

The newest release of SLAMSYSTEM is version 4.0. Some of its unique features include the following:

1. Multiple networks in a single scenario: Networks can be constructed in sections and combined at run time. The sections can be reused in future models.
2. New output graphics: These graphics support three-dimensional X-Y grids and displaying of point plot data.
3. Direct interface to SimStat (product of MC^2 Analysis Systems): These files may be loaded for advanced statistical analysis.
4. OS/2 metafiles for graphics: The OS/2 metafile format can be read for animation backgrounds or icons.

Conclusion

Simulation is a powerful approach to modeling manufacturing systems in that many complex and diverse systems can be represented. Simulation can predict system performance measures that are difficult to assess without a model. It is a proven, successful tool and has been in use since the 1950s. The current languages take advantage of the capabilities of today's microprocessors and provide the user with the needed on-line support for model development, management, and analysis.

References

Banks, J. 1994. Simulation software, paper presented as 1994 Winter Simulation Conference, Atlanta.

Banks, J., Burnette, B., Rose, J.D., and H. Kozloski. 1995. *SIMAN V and CINEMA V,* John Wiley & Sons, New York.

Cox, S.W. 1992. Simulation Studio™, in *Proceedings of the 1992 Winter Simulation Conference,* J. J. Swain, D. Goldman, R.C. Crain, and J.R. Wilson, Eds., Association for Computing Machinery, New York, 347–351.

Glavach, M.A. and Sturrock, D.T. 1993. Introduction to SIMAN/Cinema, in *Proceedings of the 1993 Winter Simulation Conference,* G.W. Evans, M. Mollaghasemi, E.C. Russell, and W.E. Biles, Eds., Association for Computing Machinery, New York, 190–192.

O'Reilly, J.J. 1993. Introduction to SLAM II and SLAMSYSTEM, in *Proceedings of the 1993 Winter Simulation Conference*, G.W. Evans, M. Mollaghasemi, E.C. Russell, and W.E. Biles, Eds., Association for Computing Machinery, New York, 179–183.

Pidd, M. 1994. An introduction to computer simulation, 1994 Winter Simulative Conference, The Management School, Lancaster University, U.K.

Pritsker, A.B. 1986. *Introduction to Simulation and SLAM II*, 3rd ed., John Wiley & Sons, New York.

Russell, E.C. 1993. SIMSCRIPT II.5 and SIMGRAPHICS tutorial, in *Proceedings of the 1993 Winter Simulation Conference*, G.W. Evans, M. Mollaghasemi, E.C. Russell, and W.E. Biles, Eds., Association for Computing Machinery, New York, 223–227.

Smith, D.S. and Crain, R.C. 1993. Industrial strength simulation using GPSS/H, in *Proceedings of the 1993 Winter Simulation Conference*, G.W. Evans, M. Mollaghasemi, E.C. Russell, and W.E. Biles, Eds., Association for Computing Machinery, New York, 218–222.

Tools for Intelligent Manufacturing Processes and Systems: Neural Networks, Fuzzy Logic, and Expert Systems

Tien-I. Liu

Introduction

Starting in the 1980s, researchers and practitioners became increasingly interested in intelligent machines and intelligent manufacturing. The goal is to model the skills and expertise of professionals so that machines and manufacturing systems can possess some of the characteristics of human intelligence. Three techniques, neural networks, fuzzy logic, and expert systems, have been widely used in manufacturing. This section describes the principles and functions of these tools. Examples in applying these tools to manufacturing applications are highlighted as well.

Neural Networks

Neural networks consist of a set of nodes which are nonlinear computational elements. The pattern of connectivity between nodes, known as weights, can be modified according to some preset learning rule. The knowledge of the networks is stored in their interconnections (Kohonen, 1986).

Since neural networks are parallel distributed processing, they have the following advantages:

1. They are adaptive and can learn from experience.
2. The network can be refined at any time with the addition of new training data.
3. Various model architectures can be used.
4. They can compute very quickly and thus they are very suitable for real-time applications.
5. They can be used for analyzing large amounts of data to determine patterns that may predict certain types of behavior.
6. They can capture the complexities of the process, including nonlinearities, even if the dynamics of the process is unknown.
7. They can make decisions based upon incomplete and noisy information.
8. They degrade gracefully even when parts of the structure have been destroyed.

Neural networks are best at performing the types of tasks which need human perception. These tasks do not have exact answers, e.g., classification and trend analysis. A typical example of such problems is machine diagnosis. An experienced mechanic can point out what is wrong with an automobile by standing beside the car and listening to the sound of the running engine. Another example is character recognition. Any person who is familiar with alphabets can easily identify the letter "A" in any of various typefaces or handwritten scripts. In both cases it is practically impossible to develop a set of if–then rules to let a computer to do the job. These tasks are also difficult to program using standard computer techniques.

Applications of Neural Networks. Many electronics and computer companies have put considerable effort into neural network development. IBM has announced a neural network development package for its computer; Intel Corporation has developed a microchip that supports this technology. Japanese companies such as Fujitsu, Hitachi, Mitsubishi, and Sumitomo Heavy Industries are also working in this field.

A bomb-detection machine which uses neural network technology to detect plastic explosives hidden in baggage has been developed. This machine has been installed in several airports. Neural networks have also been applied to ensure the operation of an industrial power distribution substation. It has replaced a conventional mechanical system and improved the performance by greater than an order of magnitude.

Integrating sensors with neural networks for monitoring and diagnostic purposes can enhance production reliability, prevent potential problems caused by abnormal conditions, and maintain high product quality in the factory (Liu and Iyer, 1993). The applications of neural networks for monitoring and diagnostic purposes have been used with ball-and-roller bearings, turning processes, milling processes, drilling processes, tapping processes, glass furnaces, etc. The results are very successful (Liu and Anatharaman, 1994).

Neural networks have also been used in the image processing for computer vision and speech-recognition systems (Badal, 1993). They also are used for the control of machines and processes. The neural network controller is capable of on-line learning of the system dynamics and then taking adequate action to achieve the predetermined goal. They also can be used to tune the gain of control systems.

Fuzzy Logic

The theory of fuzzy sets has been developed as a methodology for the formulation and solution of problems which are too complex or too ill defined to be solved by traditional techniques. In fuzzy logic, the membership in a set is not either 0 or 1; instead it is a value between 0 and 1. Membership functions span some problem domain, such as length or weight, and show the membership for each value of the problem domain. Membership functions are subjective evaluations and can be represented by many kinds of curve. However, the membership function cannot be assigned arbitrarily. The formulation of the membership function should be based upon the professional feeling and physical understanding of the problem. Let $S = \{s\}$ represent a space of objects. Then a fuzzy set X in S is a set of ordered pairs

$$X = \{s, f_x(s)\}, \quad s \in S \quad (1)$$

where $f_x(s)$ is the grade of membership of s in X and $f_x(s)$ is a number in the interval (0,1).

Fuzzy mathematics, which consists of precise rules to combine vague expressions, such as "very high" and "somewhat heavy," has been developed (Kandel, 1986). Generally speaking, fuzzy logic systems have the following advantages:

1. They are inherently flexible.
2. They are robust to noisy or missing data, unexpected disturbances, and errors in problem modeling.
3. They are suitable to deal with problems for which knowledge is approximate or problems which are so complex that it is difficult to develop an adequate mathematical model.
4. They usually are energy efficient.

Applications of Fuzzy Logic. Fuzzy mathematical techniques are very suitable for the control of machine tools, robots, and electronic systems (Mamdani, 1993). They are also applicable to image understanding for computer vision and pattern classification for the monitoring and diagnosis of manufacturing processes (Du et al., 1992).

The very first application of fuzzy logic was the control of the fuel-intake rate and gas flow of a rotating kiln used to produce cement. Since then, it has been used to control many automated manufacturing processes. Fuzzy logic has resulted in significant improvements to many commercial products, such as cameras and air conditioners. Although fuzzy logic was developed in the U.S., most of the action is in Japan. The most impressive application is a subway system operated by a fuzzy computer. It was installed in the 1980s by Hitachi at Sendai, about 200 mi north of Tokyo, Japan. This system is more than 10% energy efficient and is so smooth that passengers do not need to hang onto straps.

At AT&T Bell Laboratories, Dr. M. Togai and Dr. H. Watanabe developed the very first fuzzy logic processing chip in 1985. NASA is developing fuzzy controllers to help astronauts pilot the space shuttle in earth orbit. In the U.S. the interest in applying fuzzy logic is growing continuously.

Expert Systems

An expert system is a computer system which possesses the capability of a human expert. An expert system has the following five essential parts (Figure 13.4.4):

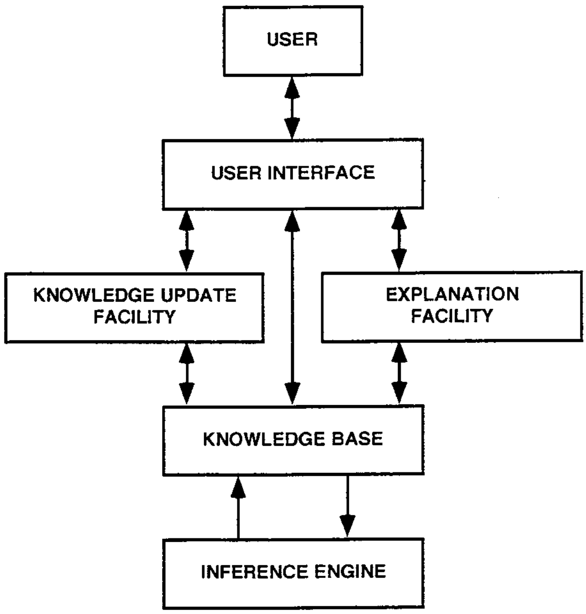

FIGURE 13.4.4 Typical expert system architecture.

1. *User Interface.* This module is the interface between the user and other parts of the expert system.
2. *Knowledge Base.* The knowledge about solving specific problems is stored in the knowledge base of an expert system. The process of building the knowledge base is called knowledge engineering and is done by a knowledge engineer (Figure 13.4.5). Knowledge engineering refers to the acquisition of knowledge from a human expert or other source and the coding of it into the expert system. The knowledge base usually consists of rules and facts. Rules are made up of English-like sentences or clauses. Rules are often defined using an if–then syntax that logically connects one or more antecedent clauses with one or more consequent clauses as follows:

   IF antecedent THEN consequent

   A rule says that if the antecedent is true, then the consequent is also true. The antecedent and consequent of rules refer to a specific fact that describes the state of the world. On the other hand, each fact is a single sentence that describes some aspect of the state of the world.
3. *Inference Engine.* The inference engine can infer new knowledge from existing knowledge stored in the knowledge base. Two general inference approaches are commonly used: forward chaining and backward chaining. Forward chaining is the reasoning from facts to conclusions resulting

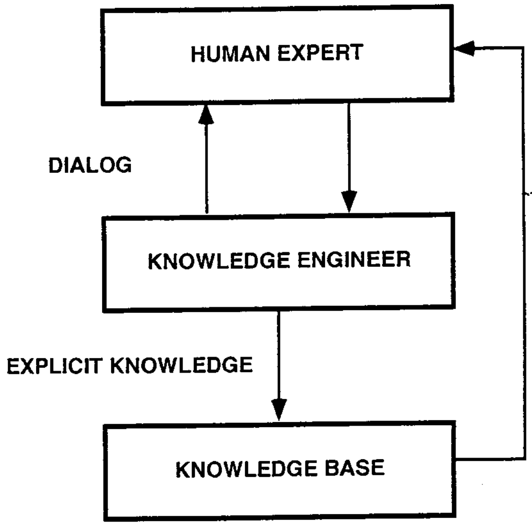

FIGURE 13.4.5 Knowledge engineering.

from those facts. Backward chaining involves reasoning in reverse from a hypothesis to the facts which support the hypothesis.

4. *Knowledge Update Facility.* The knowledge in many fields, including engineering and manufacturing, changes with time. The expert system can be updated through this facility.
5. *Explanation Facility.* Just like a human expert can explain how a specific conclusion has been drawn, the explanation facility can explain its reasoning to enhance the credibility of an expert system.

Expert systems are quite different from conventional programs because the problems usually have no algorithmic solution and rely on inferences to achieve a reasonable solution. In other words, the expert system is very suitable to solve problems which require heuristic rules. These heuristic rules can be stored in the knowledge base.

Expert systems have the following advantages as compared with the human expert:

1. They are steady and unemotional and have high reliability. Therefore, errors are reduced.
2. They have high availability. Users can get answers any time.
3. The expert system can be installed and used at multiple sites.
4. They can explain how a conclusion has been reached. Thus, the users feel comfortable working with expert systems.

Expert systems can help people solve problems. They can free the human expert from the routine job to do other work. Hence, they can increase efficiency and reduce cost.

Applications of Expert Systems. Expert systems have been used in many fields. There are many commercialized expert systems running on different computer platforms which help professionals in various fields, enhancing efficiency and productivity greatly.

Manufacturing, assembly, quality, reliability, and cost need to be taken into consideration in the early stage of product design. Expert systems have been developed for DFM, design for assembly (DFA), design for quality and reliability, product cost estimation, etc. These expert systems can be integrated with the existing CAD/CAM systems (Liu et al., 1995). Many expert systems have also been built and used in the areas of facility design, production planning and control, computer-aided process planning, material handling, quality control, equipment maintenance and repair, and real-time control (Alto et al., 1994). In simple words, design and manufacturing work can be upgraded from an experience-based to a science-based function by using expert systems.

Conclusion

Neural networks, fuzzy logic, and expert systems are the way of the future. They can make machines and manufacturing processes much smarter. Applying these techniques can lead to the realization of a fully automated factory in the future.

References

Alto, A., Dassisti, M., and Galantucci, L.M. 1994. An expert system for reliable tool-replacement policies in metal cutting, *ASME J. Eng. Ind.*, 116(3), 405–406.

Badal, D.Z. 1993. Neural network based object recognition in images, in *Proc. IEEE Int. Conf. Neural Networks,* San Francisco, CA, 1283–1288.

Du, R.X., Elbestawi, M.A., and Li, S. 1992. Tool condition monitoring in turning using fuzzy set theory, *Int. J. Mach. Tools Manuf.*, 32(6), 781–796.

Kandel, A. 1986. *Fuzzy Mathematical Techniques with Applications,* Addison-Wesley, Reading, MA.

Kohonen, T. 1986. An introduction to neural computing, *Neural Networks*, 1, 3–16.

Liu, T.I. and Anatharaman, K.S. 1994. Intelligent classification and measurement of drill wear, *ASME J. Eng. Ind.,* 116(3), 392–397.

Liu, T.I. and Iyer, N.R. 1993. Diagnosis of roller bearings using neural networks, *Int. J. Adv. Manuf. Technol.,* 8(2), 210–215.

Liu, T.I., Yang, X.M., and Kalambur, G.J. 1995. Design for machining using expert system and fuzzy logic approach, *ASM J. Mater. Eng. Performance,* 4(5), 599–609.

Mamdani, E.H. 1993. Twenty years of fuzzy control: experiences gained and lessons learnt, *Proc. IEEE 2nd Int. Conf. Fuzzy Systems,* San Francisco, CA, 339–344.

Tools for Manufacturing Facilities Planning

J. M. A. Tanchoco, Andrew C. Lee, and Su-Hsia Yang

Introduction

The main function of facility planning is the design of efficient flow of products from raw material to finished goods. It is one of the most important determinants of operating efficiency and production cost. Traditionally, the facility-planning problem is divided into three areas, namely, group technology (GT), material handling, and facility layout (see Figure 13.4.6). GT, which is closely related to cellular manufacturing, is usually defined as the grouping of dissimilar machines in close vicinity. Each group or cell is dedicated to the production of one or more parts families. The parts in the family are similar in their processing requirements (Wemmerlov and Hyer, 1989). Two of the most fundamental elements in facility planning are the facility layout and the material-handling system.

Facility layout positions the workstations around the fixed product based on the processing sequence. In a product layout, machines are arranged according to the processing sequence of the product, e.g., the assembly of automobiles, certain electronic products, etc. The machines are located so as to provide smooth and logical flow of material. In a group layout, also referred to as a cellular layout, products are grouped into logical product families. All machines which perform similar operations are grouped together in the process layout. A process layout is characterized by a high degree of flexibility and machine utilization. Regardless of the type of facility, a detailed layout should not be designed without giving serious consideration to material-handling requirements. The choice of material-handling methods and equipment is an integral part of the layout design. The facility layout design component performs two basic functions. The first function is to decide how to locate cells with respect to each other. The objective is to minimize either the total material flow distance or transportation time. The second function is to resolve the machine location and orientation relative to each other within the cell. The constraints in layout design or facility structure could prohibit the placement of the cells or machines in some

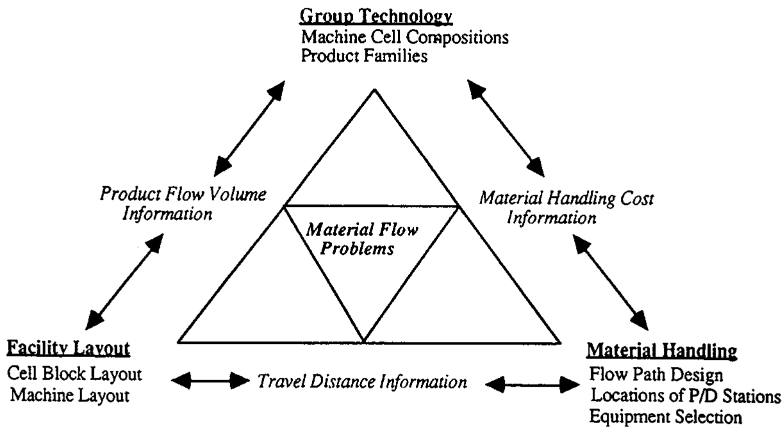

FIGURE 13.4.6 Facilities-planning framework components.

locations, thus changing the machine compositions of the cells. The resolution from this design problem has a significant impact on the distance that material has to be moved.

Material handling performs a critical function in modern dynamic manufacturing systems. The transportation of materials within a production system is often accomplished with limited resources, e.g., conveyors, forklift trucks, or automated guided vehicles (AGVs). Such transfer mechanisms not only deliver material in a timely fashion, but also provide temporary storage capacity. These limited resources require a capital investment, which increases the overall production costs. Excess transportation capability represents expenditures that are not accompanied by increased value. Insufficient capacity, on the other hand, can adversely affect production by delaying the delivery of material, thus reducing the total production volume. Material handling is generally considered as a non-value-adding activitiy. However, one can argue that material handling adds time and space value by making the material available and ready for processing. The material handling has three functions. First, the flow path design selects the route that material transfers from cell to cell or from machine to machine. Limited by the cost of remodeling, the existing plant configuration (such as the aisle width) might not be able to accommodate the traffic flow. Therefore, considerations must be given to the overall facility layout and building structure. The second function is to locate the pickup and delivery stations along the flow path. Along with the facility layout and flow path design, it is one of the most important determinants of the operating cost of material handling. Finally, the best combination of material-handling equipment is determined. Because of the complex, ill-structured, and experience-based nature of the problem, the decisions for choosing material-handling equipment tend not to be based on rigorous criteria. However, the designers must recognize that there is an appropriate level of technology for every application that will meet the combined need for maximum handling efficiency and acceptable cost.

Decision Factors for Facilities Planning

The effectiveness of a facility-planning system depends on the careful integration of GT, facility layout, and the material-handling system.

Group Technology. Mitrofanov (1959) is recognized as the first person to introduce the concept of GT and machine grouping. Since then, a number of researchers have developed techniques for machine grouping and part family formation. In general, these techniques can be categorized as follows:

1. *Machine-Component Matrix* (McAuley, 1972; King, 1980; McCormick et al., 1972): The main idea behind this approach is to delineate cells by grouping binary entries of a machine–component incidence matrix into fuzzy blocks along the matrix diagonal.

2. *Mathematical Programming Formulation* (Kusiak et al. 1986; Askin and Standridge, 1993): In this approach, a large, combinatorial mixed integer programming formulation is usually required to model the problem.
3. *Graph-Based Formulation/Partitioning* (Rajagopalan and Batra, 1975; Faber and Carter 1986): In this approach, the machine–part matrix is represented by a bipartite, transition, or boundary graph. Then, it is solved using graph theoretic clustering methodology.

Facility Layout. The facility layout problem has been formulated as a quadratic assignment problem (QAP). The objective function for this type of formulation is to minimize the total material-handling cost. Given the complexity of the QAP formulation, the size of problems which could be solved by optimal methods is very limited. Thus, heuristic algorithms are more suitable. These heuristics can be classified into two major groups. The first group is called construction algorithms (Seehof and Evans, 1967; Lee and Moore, 1967; Foulds and Giffin, 1985; Hassan and Hogg, 1991). The main idea behind these methods is to build the layout by adding one more block (a cell or a department) to a partial block layout until all blocks have been located. The methodology requires two steps: a selection step and a placement step. The selection step determines the order by which the blocks enter the layout, while the placement step selects the location of the new block to enter the layout relative to the blocks that are already in the layout. The objective is to maximize some kinds of performance criteria, e.g., total closeness rating. Graph theoretic methods have been applied as a solution methodology for this approach.

The second group of the layout heuristics is called improvement algorithms. This approach starts from an initial layout, and improvements are made by successive pairwise interchanges of blocks. The general form of an improvement algorithm consists of the following steps: (1) select a pair of activities, (2) estimate the cost of interchange, (3) exchange if the total cost is reduced, and (4) repeat until no more improvement can be made. This category of heuristics includes the computerized relative allocation of facilities technique (CRAFT) of Armour and Buffa (1963) and Buffa et al. (1964) and the methods of Hillier (1963), Fortenberry and Cox (1985), and Co et al. (1989).

Recognizing the weaknesses of both construction and improvement algorithms when applied separately, Golany and Rosenblatt (1989) proposed a hybrid method that takes advantages of both construction and improvement algorithms. They use the layout resulting from the construction algorithm as the initial layout and improve upon it using an improvement algorithm. Since most of the heuristics for facility layout are based on the "greedy approach," the solution is very sensitive to the initial layout. The final layout given by the algorithm may not be the best.

Material Handling. Material handling involves moving, storing, and controlling the flow of materials. There are several components, such as the flow path design, the locations of pickup and delivery stations, and the material-handling equipment selected, that have significant effects on the overall effectiveness of the material-handling system. The objective of flow path design is to determine the best "street network" that transporters pass through when parts are moved from one machine to the next. It is one of the major determinants in the calculations of travel times, operating expenses, and installation costs of the material-handling system. There are numerous types of flow path network configurations. The most widely adopted flow path design is the conventional flow network. It is usually a unidirectional flow network where any cell boundary is used as part of the flow path. In a unidirectional network, one has to determine the flow direction for each aisle. Its flexibility, reliability, and efficiency have made this type of network a popular choice especially when AGVs are used. In comparison, the potential of using a bidirectional flow network could make the system more efficient (Egbelu and Tanchoco, 1986). The same authors also developed guidelines for the use of a single-lane bidirectional flow network. However, the advanced hardware requirements and complicated system controllers are viewed negatively.

A single-loop flow path network can be found in many flexible manufacturing systems. The entire flow path design is made up of a single loop. This type of network can potentially minimize some of the problems associated with a conventional flow network. Congestion is inherently low and the operating rules are simple. It also has relatively low initial investment and maintenance costs.

Recently, a new flow path network configuration, a segmented flow topology (SFT), was developed by Sinriech and Tanchoco (1994). It is comprised of one or more zones, each of which is separated into nonoverlapping segments. Each segment is serviced by a single bidirectional material-handling device. Transfer buffers are located at both ends of each segment. The flow structure in each zone is determined by the logical flow requirements and by the existing aisle network. The SFT provides a simple flow structure and control system. The research also suggested that it can achieve a higher throughput capability compared to other material flow path network configurations.

The locations of pickup and delivery stations have generally been considered as a secondary issue in the design phase of facility layout and a material-handling system. Yet it can have detrimental effects on the costs of material handling and the a machine layout configuration. In a study by Warnecke et al. (1985), they confirmed that the actual pickup and delivery station flow distance is much more representative than taking the rectilinear distance between the centers of machine blocks. Since then, several design procedures have been proposed to find the optimal location of material-transfer stations. Montreuil and Ratliff (1988) suggested a systematic methodology for locating pickup and delivery stations within a facility layout using multifacility location theory. The objective function is to minimize the sum of the rectilinear distance traveled by all intercellular flows, given the boundary regions on station location. Luxhoj (1991) developed a two-phase design procedure that is suitable for the spine layout where the active flow lines are well defined.

In terms of material-handling equipment selection, there is a large variety of equipment types available with their own special functions and characteristics. Each equipment type has its own capability and limitations. Some of these characteristics are difficult to quantify. The integrated nature of the manufacturing systems complicates the material-handling selection problem. The problem was first addressed by Webster and Reed (1971). The procedure they suggested initially assigns material-handling equipment to departmental moves based on cost alone. Then, move assignments are interchanged to seek improvement in equipment utilization and total cost. Hassan et al. (1985) reformulated Webster and Reed's model as an integer programming model with the objective of minimizing the total operating and capital costs of the selected equipment. Due to the combinatorial nature of the problem, it is solved using a construction heuristic that exploits some similarities to both knapsack and the loading problems. Material Handling Equipment Selection System (MATHES) was developed by Fisher et al. (1988). MATHES is a rule-based system for the selection from 24 different types of material-handling equipment.

Conclusion

The framework discussed in this section provides an alternative perspective from the material flow viewpoint. It integrates all of the important design factors associated with facilities planning. It incorporates most of the desired properties with respect to the overall plant operations. At the same time, the framework also summarizes the difficulties and complexities which confront the facility designer. The description of the framework is intended as a general exposition of the fundamental concepts and a direction for future research in facilities planning.

References

Armour, G.C. and Buffa, E.S., 1963. A heuristic algorithm and similation approach to relative location of facilities, *Manage. Sci.*, 9(2), 294–300.

Askin, G. and Standridge, R. 1993. *Modeling and Analysis of Manufacturing Systems*, John Wiley, New York.

Buffa, E.S., Armour, G.C., and Vollman, T.E. 1964, Allocating facilities with CRAFT, *Harvard Business Rev.*, 42(2), 136–159.

Co, H., Wu, A., and Reisman, A., 1989. A throughput-maximizing facility planning and layout model, *Int. J. Prod. Res.*, 27(1), 1–12.

Egbelu, P. and Tanchoco, J.M.A. 1986. Potentials for bi-directional guide path for automated guided vehicle base systems, *Int. J. Prod. Res.*, 24(5), 1075–1097.

Faber, Z. and Carter, M.W. 1986. A new graph theory approach to forming machine cells in cellular production systems, *Flexible Manufacturing Systems: Methods and Studies*, North-Holland, Amsterdam, 301–318.

Fisher, E.L., Farber, J.B., and Kay, M.G. 1988. MATHES: an expert system for material handling equipment selection, *Eng. Costs Prod. Econ.* 14(4), 297–310.

Fortenberry, J.C. and Cox, J.F. 1985. Multiple criteria approach to the facilities layout problem, *Int. J. Prod. Res.*, 23(4)), 773–782.

Foulds, L.R. and Giffin, J.W. 1985. A graph-theoretic heuristic for minimzing total transport cost in facility layout, *Int. J. Prod. Res.*, 23(6), 1247–1257.

Golany, B. and Rosenblatt, M.J. 1989. A heuristic algorithm for the quadratic assignment formulation to the plant layout problem, *Int. J. Prod. Res.*, 27(2), 293–308.

Hassan, M.M.D. and Hogg, G.L. 1991. On constructing a block layout by graph theory, *Int. J. Prod. Res.*, 29(6). 1263–1278.

Hassan, M.M.D., Hogg, G., and Simth, D. 1985. Construction algorithm for the selection and assignment of materials handling equipment *Int. J. Prod. Res.*, (23(2), 381–392.

Hillier, F.S. 1963. Quantitative tools for plant layout analysis, *J. Ind. Eng.*, 14(1), 33–40.

Irani, S.A., Cavalier, T.M., and Cohen, P.H. 1993. Virtual manufacturing cells: exploring layout design and intercell flows for the machine sharing problem, *Int. J. Prod. Res.*, 31(4), 791–810.

King, J.R. 1980. Machine-component grouping in production flow analysis: an approach using rank order clustering algorithm, *Int. J. Prod. Res.*, 18(2), 213–232.

Kumar, K.R., Kisiak, A., and Vanelli, A. 1986. Grouping of parts and components in flexible manufacturing systems, *Eur. J. Operat. Res.*, 24, 387–397.

Kusiak, A., Vannelli, A., and Kummar, K.R. 1986. Clustering analysis: models and algorithms, *Control Cybernetics*, 15(2), 139–154.

Lee, R.C. and Moore, J.M. 1967. CORELAP: COmputerized RElationship LAyout Planning, *J. Ind. Eng.*, 18(1), 195–200.

Luxhoj, J.T. 1991. A methodology for the location of facility ingress/engress points, *Int. J. Oper. Prod. Manage.*, 11(5), 6–21.

McAuley, 1972. Machine grouping for efficient production, *The Prod. Eng.*, pp. 53–57.

McCormick, W.T., Schweitzer, P.J., and White, T.E. 1972. Problem decomposition and data reorganization by a clustering technique, *Operations Res.*, 20, 993–1009.

Mitrofanov, S.P. 1959. Nauchniye Osnovi Grouppovoi Technologi, Lenizdaz, Leningrad; translated into English, 1966, *Scientific Principle of Group Technology*, National Lending Library, England.

Rajagopalan R. and Batra, J.L. 1975. Design of cellular production system — a graph theoretic approach, *Int. J. Prod. Res.*, 13, 567–579.

Seehof, J.M. and Evans, W.O. 1967. ALDEP: Automated Layout DEsign Program, *J. Ind. Eng.*, 18(2), 690–695.

Sinriech, D. and Tanchoco, J.M.A. 1994. SFT — Segmented flow topology, *Material Flow Systems in Manufacturing*, J.M.A. Tanchoco, Ed., Chapman & Hall, London, 200–235.

Tompkins, J.A. 1993. *World Class Manufacturing*, IEEE, New York.

Warnecke, H.J., Dangelmier, W., and Kuhnle, H. 1985. Computer-aided layout planning, *Material Flow*, 1, 35–48.

Webster, D.B. and Reed, R. Jr. 1971. A material handling system selection model, *AIIE Trans.*, 3(1), 13–21.

Wemmerlov, U. and Hyer, N.L. 1989. Cellular manufacturing in the US industry: a survey of users, *Int. J. Prod. Res.*, 27(9), 1511–1530.

13.5 Rapid Prototyping

Takeo Nakagawa

Manufacturing Processes in Parts Production

The rapid progress of CAD technology for the design of machine parts has now made it easy to store three-dimensional shape data on computers. The application of this three-dimensional data has realized NC programming by CAD/CAM, resulting in the remarkable advance of automated production. The increase in highly functional machine parts and advanced designs has led to the design of more and more complicated surfaces using CAD. To reduce the lead time and costs for the development of new industrial products, "rapid prototyping" has been recognized as a unique, layered manufacturing technique for making prototypes.

With this rapid prototyping, shapes of machine parts are created by building up layers and layers of materials, unlike the material-removal technique which shapes by gradual machining using a cutting tool. In this sense, rapid prototyping resembles the joining technique of small particles or thin layers. Various different types of rapid prototyping methods have been born over the last couple of years. A common feature of these methods is that parts are directly shaped fully automatically according to CAD data. Specifically, in all of these methods, the three-dimensional CAD data is taken as composed of thin layers of two-dimensional data. The thin layers are formed using the two-dimensional data and built up to form an actual three-dimensional solid product. With the nature of the processing steps, this rapid prototyping technique is called *layered manufacturing.* As the method can also be used in other applications than prototyping, it is also referred to as *free-form fabrication.* Because three-dimensional objects can be made from three-dimensional CAD data, this new rapid-prototyping method is also called three-dimensional plotting. At the present time, laser stereolithography is the most widespread rapid-prototyping method.

Rapid Prototyping by Laser Stereolithography

Laser stereolithography involves the use of a liquid photocurable resin which cures instantaneously when scanned with a laser as a result of polymerization. As shown in Figure 13.5.1, the laser is scanned over this resin repeatedly to form thin layers of cured resin, which eventually build up to form a three-dimensional solid product. Specifically, the process involves first slicing a model based on three-dimensional CAD data stored on computer horizontally into equal thickness. Based on this slice data, the laser scans the thin liquid resin layer to form the first solid layer. Liquid resin is then poured over this cured layer of resin and again scanned by laser to form the next layer according to the next slice data. To ensure that the liquid resin on the cured resin is even, its surface is often swept by a blade. So, by repeating this process and forming layer over layer of cured resin, a solid object is formed.

Although the surface accuracy problems are resolved by making each layer extremely thin, a small amount of roughness of the surface eventually remains and subsequent polishing is often performed to achieve a smooth finish. For shapes which cannot be formed by building the layers upward, the resin is cured on a support. This support is also made of the same photocurable resin and is removed after the product has been formed. For some types of resins which do not cure completely with a laser alone, the whole product is cured by exposure to an ultraviolet lamp after it has been formed.

The photocurable resin is composed of photopolymerizing oligomer, reactive diluent, and photoinitiator. When a laser is irradiated onto this resin, the monomer undergoes a series of reactions to form a solid polymer which has a three-dimensional network structure. The resins used for laser stereolithography are the radical polymerization type, the cation polymerization type, or the hybrid type, which is a combination of the first two types. The curing properties and mechanical properties of the cured resin are important, because these affect the applications of the cured product formed by laser stereolithography. These resins are prepared minutely by adjusting the mixture rate of the resin components and the additives to suit laser stereolithography.

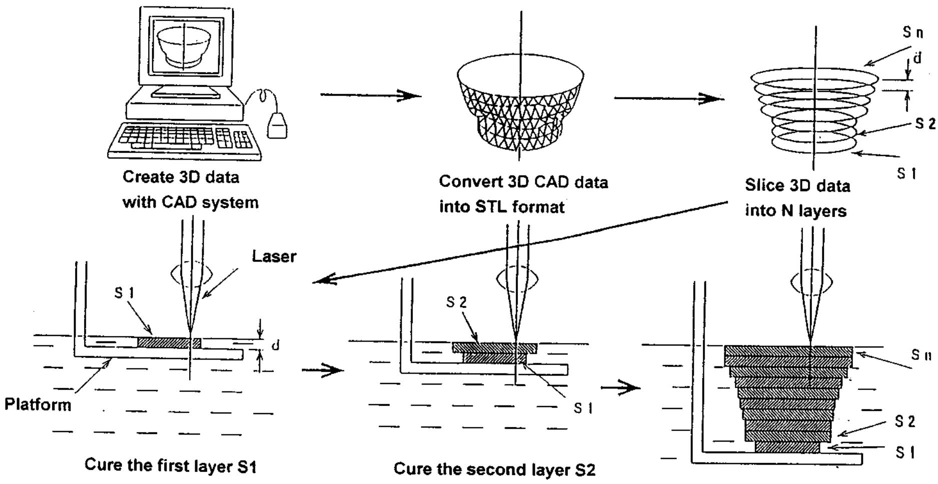

FIGURE 13.5.1 Principle of photocurable resin process.

In most cases, the He-Cd laser with 325-mm wavelength or Ar laser with 364-mm wavelength is used as the light source. Higher-power lasers like Ar perform higher-speed beam scanning, resulting in the increase of modeling speed. As shown in Figure 13.5.2, these lasers scan at very high speed in the same way as the laser printer by rotating galvanomirrors. In some special machines, the laser beam is exposed from the bottom as shown in Figure 13.5.3 by an *XY* plotter.

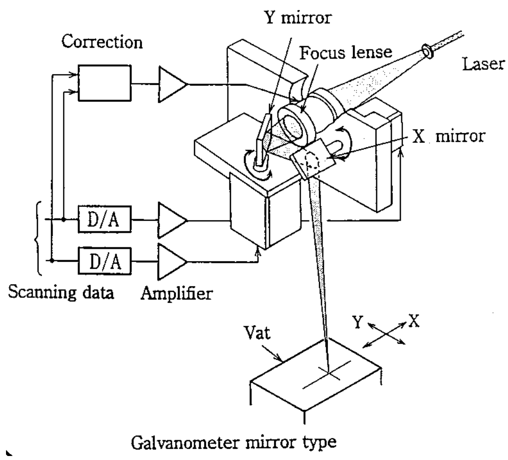

FIGURE 13.5.2 Scanning system of UV laser beam.

Laser stereolithography requires three-dimensional CAD data composed of surface or solid data in order to create solid models. Many types of CAD systems are now available on the market, and most of the CAD data they provide can be transmitted to laser stereolithography systems. At the same time, the laser stereolithography system is equipped with a scanning program, support and reinforcing rib design software, magnification/contraction functions, functions to determine the scanning and operating conditions, and simulation functions. Figure 13.5.4 shows a software flowchart in laser stereolithography.

The following shows the advantages of rapid-prototyping over material-removing processes such as machining.

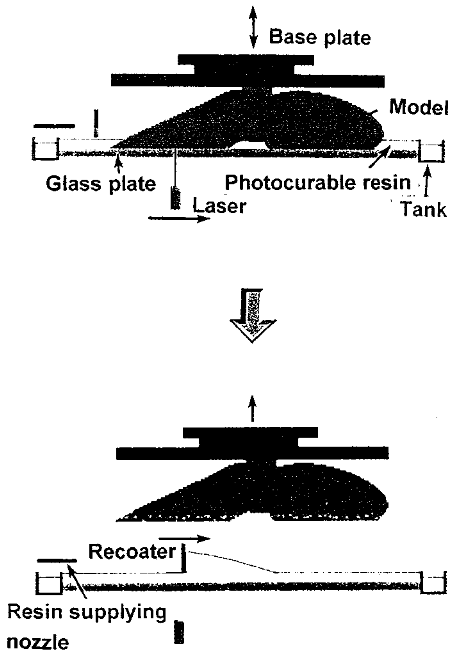

FIGURE 13.5.3 Laser beam radiation from below (Denken).

1. Deep holes and structures with complicated internal shapes which cannot be machined simply by cutting tools can be formed in a single process. Moreover, one rapid-prototyping machine is usually capable of fabricating any type of shape.
2. Rapid prototyping requires no complicated control programs such as tool path and repositioning of the workpiece. With the three-dimensional CAD data, there is no need for special knowledge of the cutting process, and operations from data input to actual fabrication are simple and short.
3. The rapid-prototyping systems produce no machining wastes. Because they do not vibrate and are silent, they can be used in offices like OA business machines. They can also be operated fully automatically even at night since there is no need for the management of tooling.

The major shortcomings of laser stereolithography are that only photocurable resins can be used and the material strength of these materials is slightly worse than the common polymer. In addition, metal products cannot be manufactured directly by laser stereolithography.

Other Rapid-Prototyping Methods

The laser stereolithography method was developed in an early stage and is currently applied extensively. Besides laser stereolithography, many different types of new rapid-prototyping methods have also emerged. As shown in Figure 13.5.5, rapid prototyping can broadly be classified into photopolymer, powder sintering, ink jetting, fused deposition, and sheet cutting. Figure 13.5.6 shows the history of these rapid prototyping systems. Most of the methods were developed in the U.S., but the photopolymer process and sheet lamination were first proposed in Japan.

Another photopolymer process is the mask pattern-curing method shown in Figure 13.5.7. Similar to the photocopying process, a master pattern based on slice data is created, this pattern on the glass sheet is placed over a photocurable resin layer, and this layer is exposed to ultraviolet light. Although the machine is large, the exposing speed is faster than the above-mentioned laser beam method, and the thickness of the product is very precise because each surface formed is cut by milling to obtain precise thin layers.

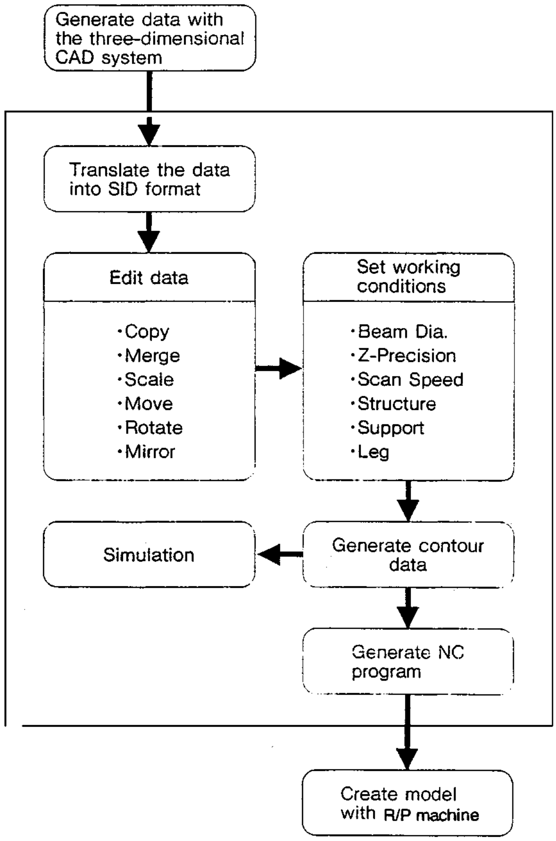

FIGURE 13.5.4 Software for R/P system (CMET).

Three-dimensional objects can also be formed by powder sintering. In the process shown in Figure 13.5.8, powder is used instead of liquid photocurable resin. The powder is evened out using a roller, a CO_2 laser is beamed, and the powder is bonded by heat fusion. In this case, powder is heated up beforehand to the temperature just below the melting point in the antioxidation environment using N_2 gas. It is possible to create high-density polymer solid models as well as porous models.

Porous polycarbonate models are quite suitable for the investment casting model. With this method, metal and ceramic powders can also be used. Metal and ceramic powders used are coated by resin and each metal or ceramic powder is bonded by the coated resin. Sintered porous ceramic molds can be used for casting molds. Powder binding can be performed by spraying binding material on the loose powder layer through the ink jet nozzle as shown in Figure 13.5.9. This is also used for making the sand mold for casting. When wax or resin is sprayed from the jet nozzle, wax or resin models can be fabricated as shown in Figure 13.5.10. In this case, the surface of the sprayed thin layer should be machined smoothly and flatly in order to obtain vertical accuracy.

Figure 13.5.11 shows the fused deposition method. In this method, a fine nozzle deposits a layer of resin or wax. Wax is normally used to form lost wax models. Fused deposition systems, in which material is supplied by the pellet or wire, enable materials to be formed very similarly to general injection mold materials like ABS and nylon. One of the two nozzles is used for making the support, where the support material is usually wax with lower melting points.

Figure 13.5.12 shows two methods which cut thin sheets according to slice data and laminate them to form three-dimentional objects. One method uses a laser to cut sheets applied with adhesive which are then laminated by hot toll pressing, while the other uses a knife to cut the sheets. In the latter case,

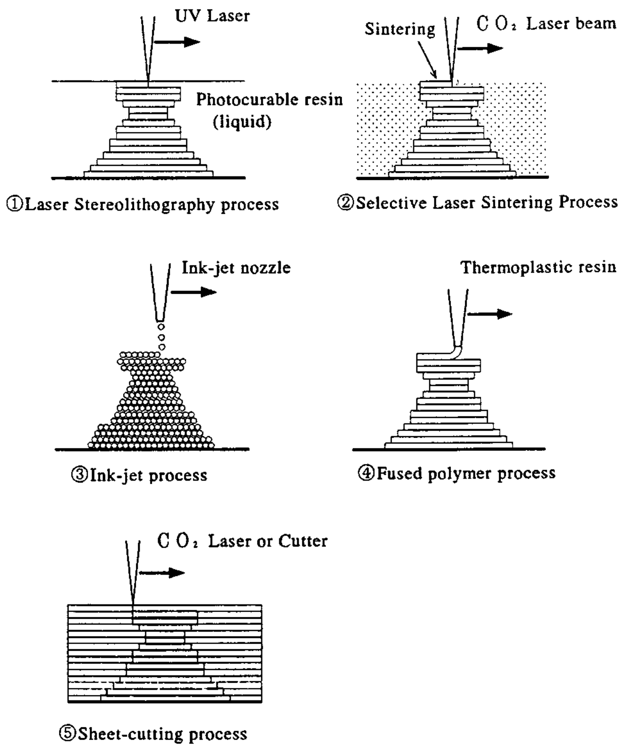

FIGURE 13.5.5 Schematic of various layer-additive fabrication processes.

adhesive is applied to sheets of paper according to the desired shape by spraying the toner using a dry Xerox-type copy machine. Due to the use of paper in these sheet lamination methods, the model formed should be immediately coated to prevent the absorption of moisture. Although there is a limit to the shapes that can be made, the method is nevertheless used for making wood models for casting, because it is inexpensive and enables large-sized models to be made and the model material is similar to wood.

The common feature of all of these methods is that slice data is obtained from three-dimentional CAD data and this slice data is used to laminate thin layers of material, which means that the same software can be used for all of these methods. Another feature is that all rapid-prototyping machines use the modified printing technology.

While many of these rapid-prototyping machines tend to be costly, inexpensive models are also now available. Machine cost reduction has been achieved by proper utilization of key parts which are used for the printer. New and improved methods should continue to be developed with the introduction of printing technology.

Application of Rapid Prototyping

Figure 13.5.13 shows the applications of three-dimentional models made by rapid prototyping. They are mainly intended for verifying CAD data, checking the designs, functional checks of prototypes, wax models for investment casting, master models for die and model making, mold making for prototype manufacturing, casting models, and medical use CT and MRI data. Although dimensional accuracy was given little importance in the verification stage of CAD data and design, high dimensional accuracy is now demanded of the functional check of prototypes. Because photocurable resins contract in the solidification process, slight distortions are generated in the fabricated product. Among the various rapid-prototyping machines, the photocurable resin process is most suitable for making very complicated shapes and obtaining the highest accuracy.

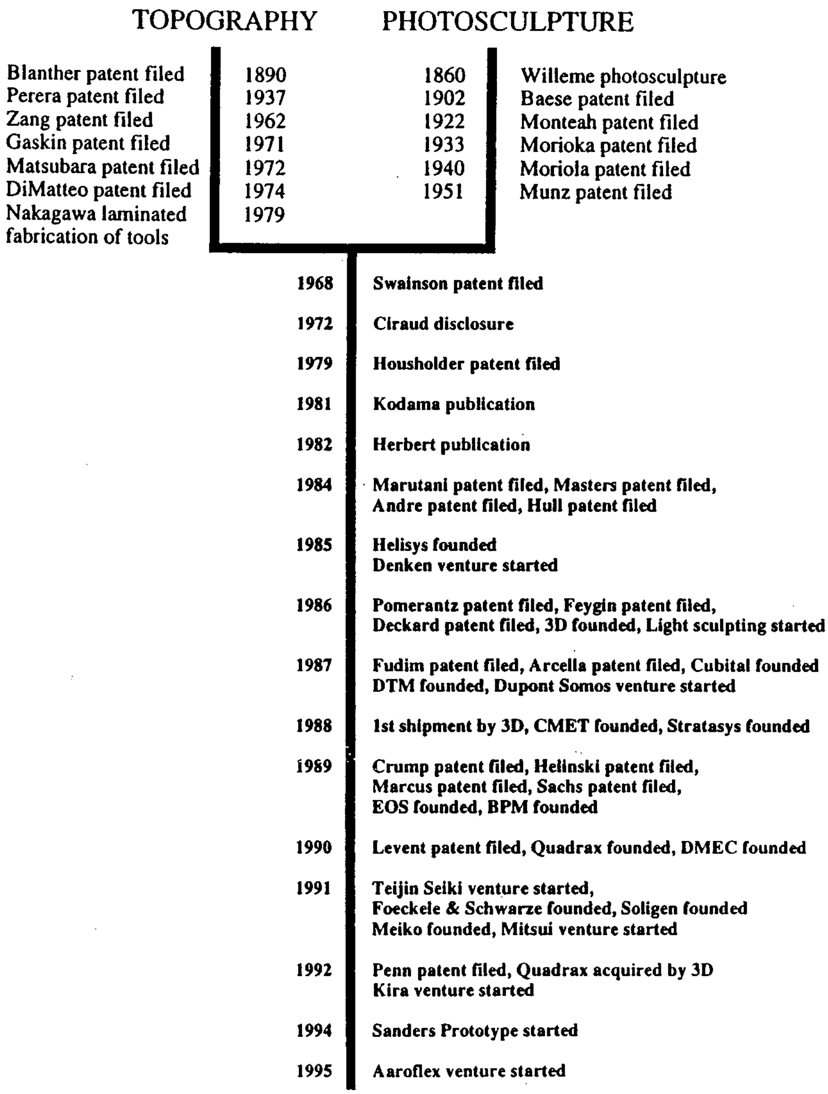

FIGURE 13.5.6 History of R/P (Joseph Beaman).

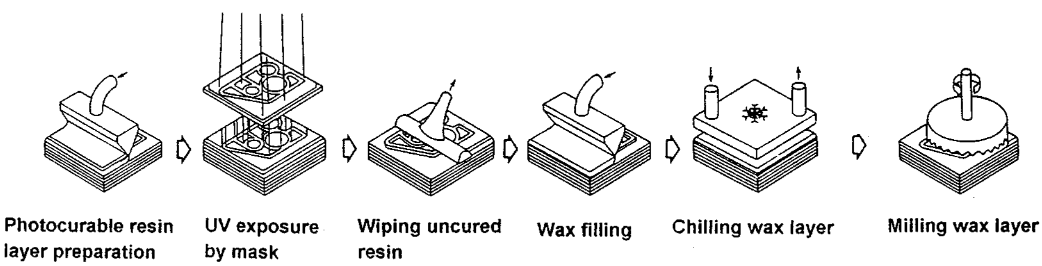

FIGURE 13.5.7 UV stereolithography (Cubital).

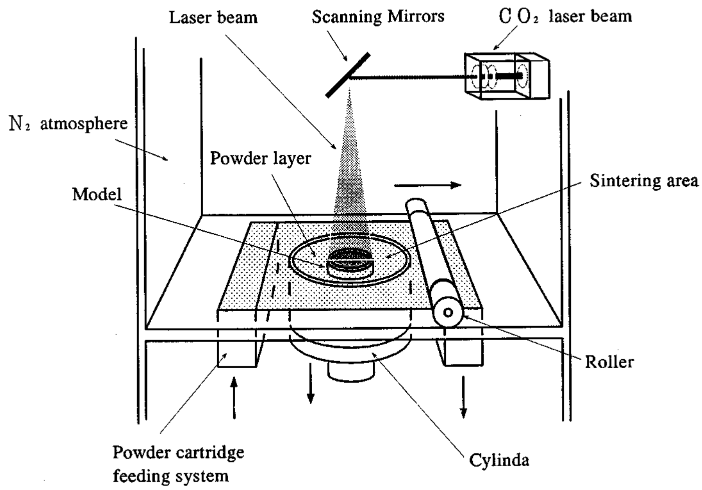

FIGURE 13.5.8 Selective laser sintering process (DTM, EOS).

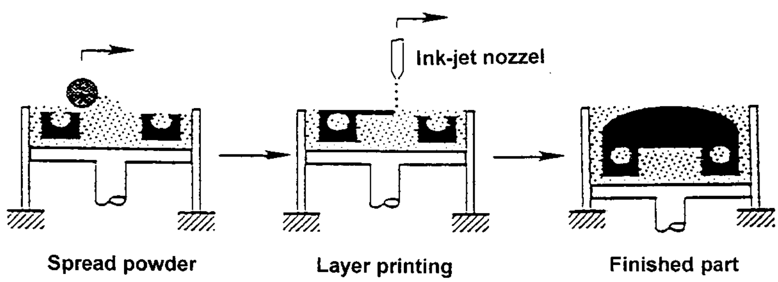

FIGURE 13.5.9 Ink-jet binding process (MIT, 3D printing).

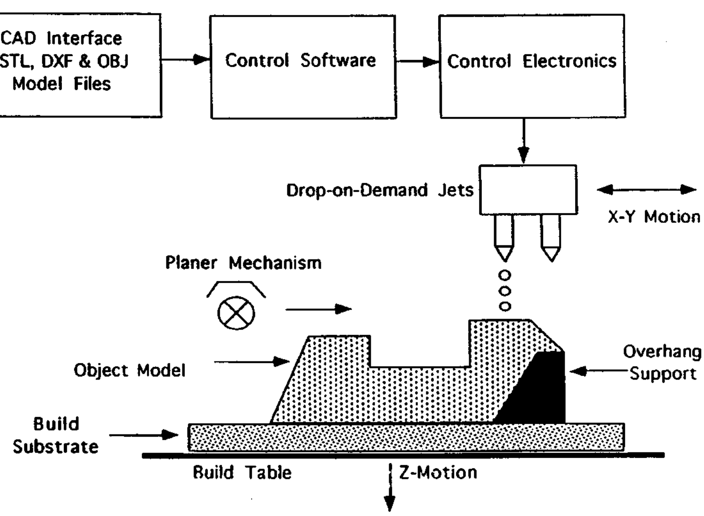

FIGURE 13.5.10 Ink-jet process (Sanders prototype).

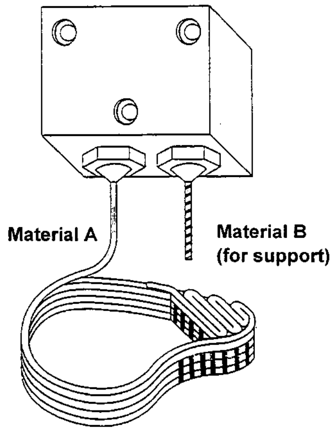

FIGURE 13.5.11 Fused deposition process (Stratasys).

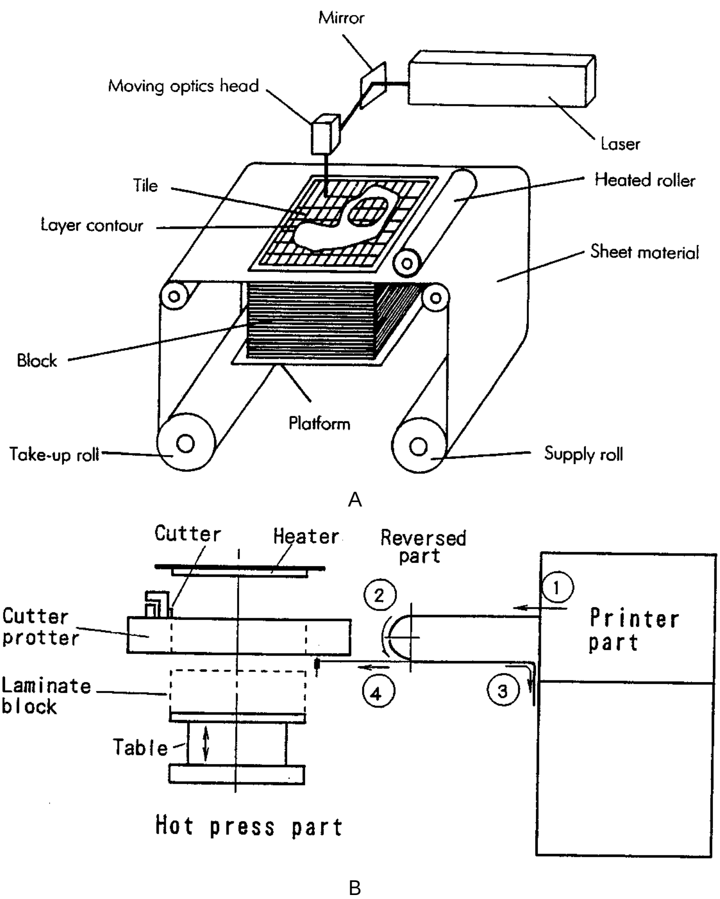

FIGURE 13.5.12 Sheet cutting process. (A) Laser (LOM). (B) Cutter (Kira).

Even in laser stereolithography, accuracy has improved to a considerable extent with the improvement of the resin and scanning method and the accumulation of know-how for positioning the reinforcing rib. It should also be possible to attain the same accuracy as injection molds by measuring formed products, correcting the data, or predicting errors.

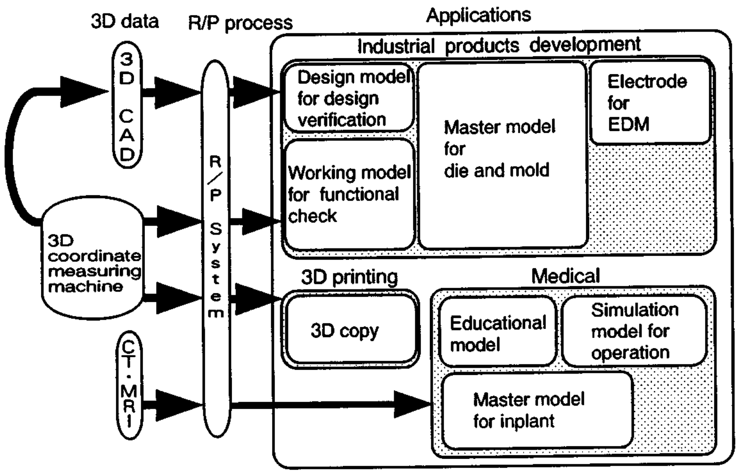

FIGURE 13.5.13 Application of rapid prototyping (source: CMET).

In general, photocurable resins are generally weak and brittle as compared with conventional polymer parts produced by injection molds. Urethane resin which is usually used for vacuum casting with a silicon rubber mold reversely copied from a rapid-prototyping model also lacks the required strength. In order to carry out the functional check of the prototypes created, other processes which can use normal thermoplastic should be used.

Figure 13.5.14 shows an intake manifold for car engines made by laser stereolithography. This serves as a test model for checking fluid performance of air. For such purposes, current photocurable resins available prove relatively satisfactory.

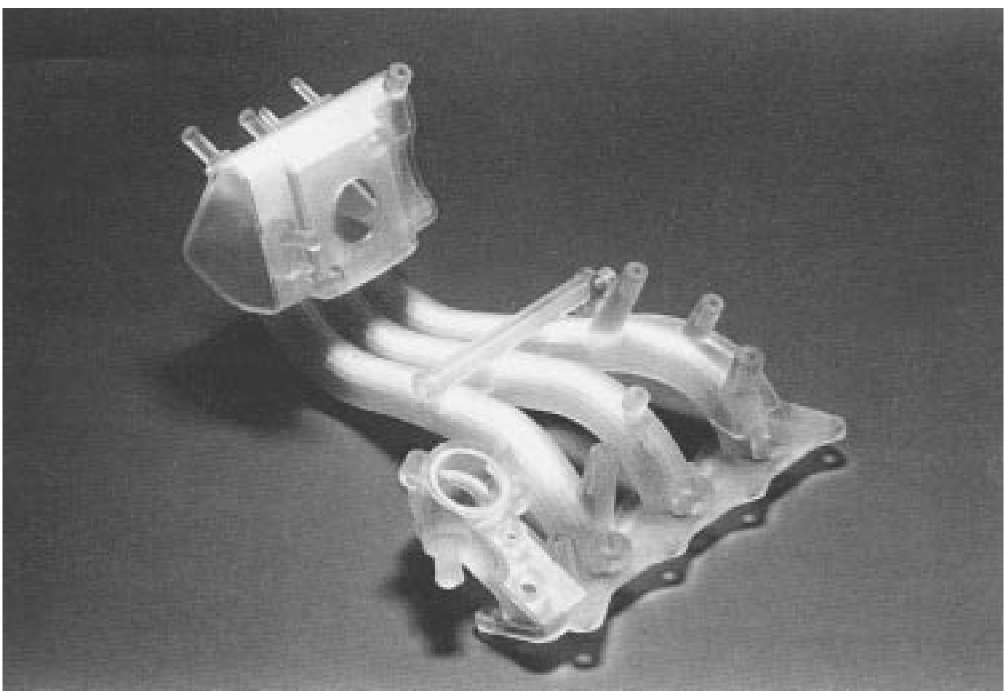

FIGURE 13.5.14 Sample for fluid dynamic analysis.

Although the powder sintering process can apply some metals, there is no suitable rapid-prototyping technique which can directly form products from metal materials at the moment. Research activities are underway to study the feasibility of producing molds from metal materials directly using a three-dimentional printing technique. In a general application, the lost wax models are first made by rapid prototyping and then used for creating metal prototypes by investment casting.

Rapid prototyping is still a relatively new technology, and therefore there are considerable opportunities for technical improvements.

General Rapid Prototyping in Production

For rapid prototyping to be carried out, three-dimentional CAD data must be available. Creating the CAD data takes far more time than creating the three-dimentional models based on the CAD data. By realizing efficient concurrent engineering, rapid prototyping will no doubt become a very important tool. In general, many other types of production systems can be included in the list of systems currently termed *general rapid prototyping* (casting and machining, etc.).

Use of Three-Dimentional CAD Data

Casting methods which are able to produce green sand molds satisfy the conditions of rapid prototyping. Expendable pattern casting is also suitable for rapid manufacturing. In this case, a three-dimentional polystyrene foam model is made by machining or binding. Most of the industrial products around us are produced with dies and molds. Because they are expensive to manufacture, dies and molds are unsuitable for making prototypes and for small-lot production. This may be a reason why rapid prototyping was developed; however, some prototype production methods do involve the use of dies and molds. Flexible prototype production has been carried out in sheet metal forming with the use of the turret punch press, laser beam cutting machine, and NC press brake. Producing dies and molds rapidly and manufacturing using such dies and molds also fit into the category of general rapid prototyping in the broad sense. Examples include what is known as the low-cost blanking dies using steel rule, deep drawing die made of zinc alloy and bismuth alloy.

Among the many general rapid-prototyping systems that exist, those newly developed rapid-prototyping methods discussed are gradually becoming methods for creating complicated products accurately with the use of three-dimentional CAD data. In terms of the total cost, applications of these new methods are still limited, but the spread of three-dimentional CAD data and technological progress of rapid prototyping should make them one of the common manufacturing techniques in the near future.

References

Ashley, S. 1992. Rapid prototyping systems, *Mech. Eng.,* April, 34–43.

Jacobs, P.F. 1992. *Rapid Prototyping and Manufacturing,* Society of Manufacturing Engineers, Dearborn, MI.

Rapid Prototyping in Europe and Japan, Japan and World Technology Evaluation Centers CJTEC/WTECS Report, September 1996, Loyola College, Baltimore, MD.

Sachs, E. et al. 1990. Three dimensional printing: rapid tooling and prototypes directly from a CAD model, *CIRP Ann.,* 39(1), 201–204.

Solid Freeform Fabrication Symposium, University of Texas, Austin, TX 1991–1993.

13.6 Underlying Paradigms in Manufacturing Systems and Enterprise Management for the 21st Century

Quality Systems

H. E. Cook

Introduction

Quality engineering has been described as the process of minimizing the sum of the total costs and the functional losses of manufactured products. Total costs include variable costs, investment, maintenance/repair costs, environmental losses, and costs of disposal or recycling. Functional losses arise from deviations from ideal performance. A subset of total quality management, quality engineering, focuses on parameter and tolerance design after the target specifications for the product have been developed as part of system design.

In contrast to quality engineering, total quality management embraces the entire product realization process. Its objective should be to maximize the net value of the product to society which includes buyer, seller, and the rest of society. Product value is determined solely by the customer and can be set equal to the maximum amount the customer would be willing to pay for the product. For a product to be purchased, its price must be less than its perceived value to the customer at the time of purchase. Consumer surplus is the difference between value and price.

The true value of a product is formed by the customer after assessing the product's performance over the complete time period that he or she used it. Functional quality loss is also known as the cost of inferior quality which is equal to the loss of value incurred by a product as a result of its attributes being off their ideal specification points (Figure 13.6.1). When manufacturing costs are added to value, the resulting sum (equal to total quality less environmental losses) is maximized when the attribute is off its ideal specification because of the impossibly high costs required to make a product perfect.

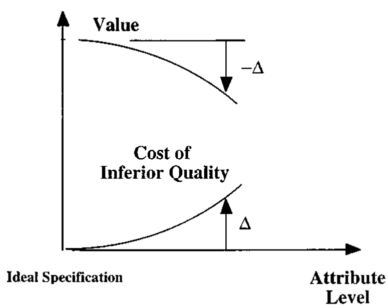

FIGURE 13.6.1 The relation of product value to the cost of inferior quality.

Requirements Flow

The systems viewpoint, as expressed by the flow of requirements shown in Figure 13.6.2, is helpful in considering the full ramifications of total quality management. Every system can be divided into subsystems and every system is but a subsystem of a larger system. Each task receives input requirements from its customer (either internal or external) and sends output requirements to its suppliers.

Task Objectives

A major objective of the system task is to assess customer needs, to translate those needs into a complete set of system-level specifications for the product, and to send a key (but partial) set of subsystem requirements to those responsible for the subsystem tasks. The system specifications and subsystem requirements developed by the system task should be such that (1) customers will want to purchase the product in a competitive marketplace and, with use, find that the product meets or exceeds their

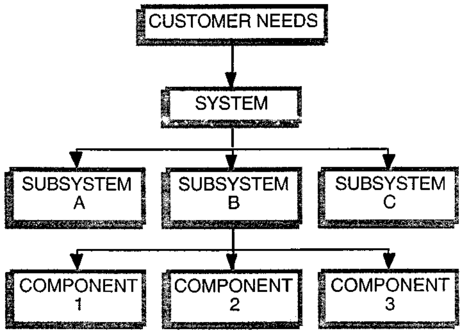

FIGURE 13.6.2 Flow of requirements from customer through enterprise.

expectations, (2) the product will meet the profitability objectives of the enterprise, and (3) all environmental rules and regulations are met. The system task also has the responsibility of resolving conflicts which arise between subsystem tasks.

The subsystem task receives the key requirements from its internal customer and translates these into a complete set of subsystem requirements and sends a key (but partial) set of component requirements to those responsible for the component tasks. In turn, those responsible for each component task translate the requirements received into a complete set of component requirements and a partial (but key) set of raw material requirements. Requirements set at each level include controls on variable costs, investment, performance, reliability, durability, service, disposal, environmental quality, package, assembly, and timing for both production and prototype parts. Synchronization is very important to total quality management as parts should be received exactly when needed with minimal inventory.

Parts Flow

The response to the requirements flow is a parts flow in the opposite direction that begins with the conversion of raw materials into components. This is followed by the assembly of components into subsystems which are shipped to the system task for final assembly. The process is completed by shipping the finished product to the customer. Thus, each task shown in Figure 13.6.1 has both a planning or design function as well as other functions including manufacturing, assembly, purchasing, marketing, service, accounting, and finance. The actions taken to meet customer needs should be traceable as the requirements flow through the enterprise. With the systems viewpoint, all parameters are measured or computed at the full system level including value, costs, and investment.

Task Management

Within each task shown in Figure 13.6.2 are subtasks. The combined flow of requirements and parts between several subtasks is shown in Figure 13.6.3 using a modified IDEF representation. Requirements are shown as controls which flow from left to right, and parts, in response, are shown as flowing from right to left.

Each task is accomplished by exercising its authority, responsibility, and capability. Authority to set requirements on parts should rest fully and undiluted with the task receiving the parts. The task which ships the parts should possess the full authority, responsibility, and capability to manufacture the parts for its customer. Before sourcing of parts, demonstration of capability by the manufacturer is a vital element of sound quality engineering. Capability is ultimately determined by the set of tools which the task has at its disposal and includes the skills and experience of the people as well as the hardware and software used by them.

Because a broad range of skills is needed, the required expertise is generated by forming a team to carry out the task. Quality tools used by the teams include structured methodologies such as Taguchi methods (design of experiments), quality function deployment, failure mode effects and criticality

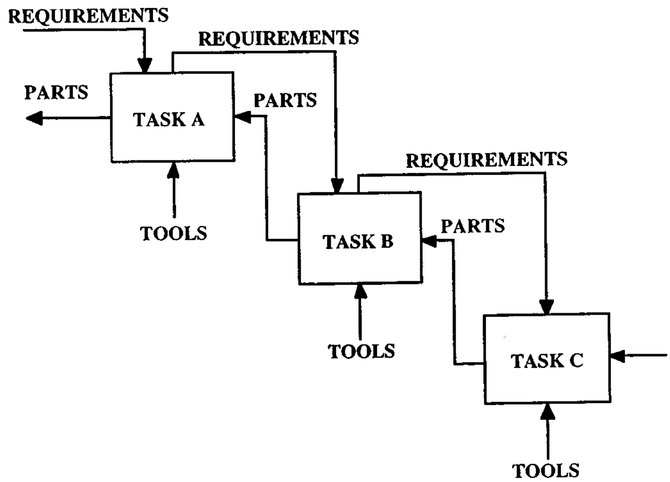

FIGURE 13.6.3 The flow of requirements and parts in a modified IDEF0 (function model) representation.

analysis, statistical process control, cost analysis, and value analysis. It is highly recommended that final authority and responsibility for each task rest with one person.

Fundamental and Bottom-Line Metrics

A variety of parameters can be used to measure the progress of quality improvements. These include things such as the degree of customer satisfaction expressed for the product, the frequency of repair, and the variance found in product dimensions and performance levels that directly impact value to the customer. Repair and operating costs borne by the customer subtract from value if they are greater than what was anticipated by the customer. Likewise, resale price subtracts form value if it is below what the customer expected. Performance degradation of the product over its lifetime of use subtracts from its value. Noise and atmospheric pollution caused by the manufacture and use of the product subtract from the net value of the product to society. The costs to manufacture and develop the product subtract from the net value received by the manufacturer. Moreover, products which are not improved in value and reduced in costs at the pace of competing products will likely be eliminated from the market in time.

These metrics can be grouped into one of three categories — value, cost, and the pace of innovation. They represent the fundamental metrics for the product because they determine what the bottom-line metrics of profitability and market share will be. Management of the fundamental metrics is the management of total quality and, likewise, the management of the total enterprise. The level of sustained profitability is the best measure of how well total quality is being managed in competitive markets.

Collaborative Manufacturing

James J. Solberg

Introduction

The world of manufacturing is undergoing rapid change in almost every aspect. It has become something of a cliché to speak of paradigm shifts, but there is no doubt that many assumptions, beliefs, and practices of the past are being seriously questioned. Meanwhile, the daily struggle of manufacturing enterprises to cope with the ordinary problems of producing products and satisfying customers goes on. In ecological terms, manufacturing companies are faced with a competitive struggle for survival (as always), augmented with the additional challenge of a changing climate. It is not surprising that many people are confused about what is happening and what to do about it.

What Is Collaborative Manufacturing (CM)?

CM is a very broad arena which incorporates many other topical themes of the day, including team design, computer-supported collaborative work, agile manufacturing, enterprise integration, virtual enterprises, high-performance distributed computing, concurrent engineering, computer-integrated manufacturing, virtual reality, global sourcing, and business process reengineering. In general, CM is defined by the following attributes:

- Integrated product and process development, including customers and suppliers;
- Flexible manufacturing distributed over networks of cooperating facilities;
- Teamwork among geographically and organizationally distributed units;
- High-technology support for the collaboration, including high-speed information networks and integration methodology;
- Multidisciplines and multiple objectives.

What is the payoff for CM? Companies that can engage effectively in CM will have the potential for

- Better market opportunities;
- A wider range of design and processing options over which to optimize;
- Fewer and looser constraints restricting their capabilities;
- Lower investment costs;
- Better utilization of resources;
- Faster response to changes.

Obviously, we are still a long way from being able to do what was described in our scenario. However, many of the pieces are already available, and many of the enabling technologies are rapidly coming into commercial use. The high-speed networks (well beyond current Internet speeds) are being developed and are certain to become both cost-effective and ubiquitous within a few years. The needed software developments, such as agent programming languages and interoperability standards, are progressing nicely. Nevertheless, an enormous agenda for needed research can be derived from unmet needs, particularly in areas that directly relate traditional manufacturing to information technology. Bridging these two research communities will not be easy, but the effort will offer great rewards to those who succeed.

Who Is Doing CM?

A great deal of the current work in CM and related themes is best accessed through the Internet. Understandably, the most active researchers in these fields are "Internet aware" and use it for both gathering and distributing their knowledge. Consequently, the best way to survey recent work is to browse the net, following links to associated sites. Unfortunately, the medium is so dynamic that material can appear or disappear at any time. A few of the better home pages are given here as starting points. Search engines can pick up more current connections.

- ACORN — A project involving Carnegie Mellon University, MIT, the University of Michigan and Enterprise Integration Technologies, Inc. (EIT):

  (http://www.edrc.cmu.edu:8888/acorn/acorn_front.html).
- Agile Manufacturing projects and organizations — A wide range of research an development projects funded through DARPA and NSF:

  (http://absu.amef.lehigh.edu/NIST-COPIES/pimain.html).
- SHARE — A DARPA-sponsored project to create a methodology and environment for collaborative product development, conducted by EIT and Stanford:

  (http://gummo. stanford.edu/html/SHARE/share.html).

- SHADE — Another DARPA-sponsored project, SHAred Dependency Engineering, information sharing aspect of concurrent engineering. It is led by the Lockheed AI Center, with help from Stanford and EIT: (http://hitchhiker.space.lockheed.com/aic/shade/papers/shadeoverview.html)
- Succeed — An NSF education coalition, which is investigating collaboration technologies to support research and education: (http://fiddle.ee.vt.edu/succeed/collaboration.html).
- Purdue Center for Collaborative Manufacturing — An NSF-sponsored engineering research center focused on the entire range of CM issues: (http://erc.www.ecn.purdue. edu/erc/).
- Groupware yellow pages: (http://www.consensus.com:8300/GWYP_TOC.html).
- Human computer interaction: (ftp://cheops.cis.ohio-state.edu/pub/hcibib/README.html).
- Computer Supported Collaborative Work yellow pages: (http://www.tft.tele.no/cscw/).
- Cross platform page: (http://www.mps.org/~ebennett/).

References

National Research Council. 1994. *Realizing the Information Future: The Internet and Beyond*, National Academy Press, Washington, D.C.

National Research Council. 1994. *Research Recommendations to Facilitate Distributed Work*, National Academy Press, Washington, D.C.

National Research Council. 1995. *Unit Manufacturing Processes: Issues and Opportunities in Research*, National Academy Press, Washington, D.C.

National Research Council. 1995. *Information Technology for Manufacturing: A Research Agenda*, National Academy Press, Washington, D.C.

Electronic Data Interchange

Chris Wang

Introduction

Electronic data interchange (EDI) is a method to exchange business information between computer systems. In a traditional purchasing environment, buyers, when placing computer-generated orders, will mail them to suppliers, and it could take days before the suppliers receive them and then rekey the orders into their computer system. Using EDI, the buyer's computer system can generate an EDI standard order transaction and transmit it directly to the supplier's inventory system for material pickup. It happens instantly. The benefit of EDI is quite obvious in this case as it reduces material lead time dramatically. Consequently, the objectives of EDI implementation should not be limited to just reducing paperwork and clerical work; instead, it should be used as a methodology to streamline company processes and become competitive in the marketplace.

EDI Elements

EDI consists of the following elements:

- Trading partners — The parties, such as a manufacturer and a supplier, who agree to exchange information.
- Standards — The industry-supplied national, or international formats to which information is converted, allowing disparate computer systems and applications to interchange it. This will be discussed in more detail later.
- Applications — The programs that process business information. For example, an orders application can communicate with an orders entry application of the trading partner.
- Translation — The process of converting business information, usually from a format used by an application, to a standard format, and vice versa.

- Electronic transmission — The means by which the information is delivered, such as a public network. Some companies may choose to build their own transmission facilities. For others, VAN (value-added network) seems to be a good choice as companies do not have to invest heavily in communication equipment and personnel to support it. The VAN provider can handle disparate communication hardware and software and provide wide-area network access at a reasonable cost.

EDI in Manufacturing

In the present-day business environment, many companies are turning to just-in-time (JIT) and other techniques to compete as effectively as possible. EDI can make an important contribution to the success of JIT by ensuring that information exchanged between business partners is also just in time.

In traditional manufacturing, material is stored in quantities much larger than required because of faulty components and possible waste in the production process. To solve problems of carrying safety stock and still producing high-quality product, JIT seems to be an effective technology.

JIT systems are designed to pull raw materials and subassemblies through the manufacturing process only when they are needed and exactly when they are needed. Also, with rapidly changing production needs, orders are getting smaller and are issued more frequently. The traditional paperwork environment simply cannot effectively cope with this change. This is why EDI comes in to play a key role to provide fast, accurate information to achieve these JIT goals. In other words, EDI can provide JIT information in manufacturing processes.

- EDI cuts order delivery and lead time — The more control points you have in a process, the greater the number of potential problems. EDI eliminates “control points” for the order process. It eliminates the need to mail orders and rekey order information at the receiving end. It reduces the material lead time for production use.
- Connect applications and processes — With EDI capability, information, such as scheduling, orders, advance delivery notice, statistical process control data, and material safety data sheets can pass quickly and accurately from the supplier’s computer application to the customer’s computer application, so that arriving material can be put to production use with confidence. This meets one of the important goals of JIT, that is, to turn the supplier’s entire production line into a vast stockroom so a company does not have to maintain a huge warehouse and the working capital tied to excess inventory.
- Improve relationship with customers and suppliers — In the supply chain environment, the quicker the chain moves, the better the customers’ needs can be met. With quicker orders, acknowledgments, order changes, and invoices, EDI can satisfy customers’ needs more quickly. Also, the time spent on order tracking and error recovery can now be used in a more productive way and can improve customer/supplier relationships. Companies deeply involved in EDI may see the number of suppliers reduced. This is because through the EDI process, a company can weed out many suppliers who are not efficient and reliable.

EDI Standards

When two organizations exchange business forms electronically, information is encoded and decoded by the computer software of both parties. Therefore, the information must be unambiguous, in order to avoid different interpretations. This relates to the meaning of the terms used, the representation of data used, the codes to be used for data, and the sequence in which data are to be transmitted. All these parameters must be arranged between the two parties on a detailed level.

There are many standard types — it can be based on bilateral agreement, imposed by a dominating party in a certain marketplace, or jointly developed by an industrial group. Some standards have been ratified by international organizations.

The pioneer of EDI standard development was the transportation industry. TDCC (Transportation Data Coordinating Committee) developed sets of standards for transportation mode — air, ocean, motor, and rail. Later, the U.S. grocery industry developed a set of standards, UCS (Uniform Communication Standard), based on TDCC structure.

The TDCC and UCS are more geared toward the business forms exchanged by shipper/carrier, for example, bill of lading. Not until the American National Standards Institute got involved did a general use standard for all industries start to develop, leading to the birth of ANSI X12 standards.

The ANSI X12 is popular in the U.S. Although ANSI X12 is intended for all industries, different user groups still come up with their own conventions to address their specific needs but remain under the X12 umbrella. To name a few, there are AIAG (Automotive Industry Action Group) for the auto industry, CIDX for the chemical industry, and EIDC for the electronic industry.

In Europe, at approximately the same time period, under the leadership of the United Kingdom, the TDI (Trade Data Interchange) was developed. The TDI syntax and structure are quite different from the ANSI X12. To resolve the incompatibility, the U.N. organization UNJEDI was formed to develop an EDI international standard containing features from both TDI and ANSI X12. The result was EDIFACT (EDI for Administrative, Commerce, and Transport). This is the standard to which the world is trying to convert.

EDI Implementation

Before implementation of EDI, planning is critical to success. First, get all the right people involved in planning and implementation of EDI. Ensure that every employee gets EDI education on how to use EDI as a business tool to manage his or her job. Prepare a strategic plan to get approval and support from top management. Top management should be aware of the significant benefits of EDI and has to appreciate the potential of EDI as a business methodology to improve the bottom line.

As part of a strategic plan, it is crucial to perform an operational evaluation. This evaluation details how the internal departments of the company function. For each paper document under evaluation, information flow is tracked, processing procedures are scrutinized, time is measured, and costs are calculated. This will provide top management with valuable information as important as industry trends and competition information. The operational evaluation provides the company with detailed documentation about how it does business in a paper-based environment. This information then serves as a benchmark against which to measure projected costs and benefits of the EDI model.

Once the strategic plan is in place, available resources must be allocated to the departments that will generate the most benefits for the company. Once EDI is implemented in the company, the next step is to sell it to trading partners to maximize the EDI investment.

Summary

EDI is on a fast growing path. EDI software and communication services are available and not expensive. It will not be too long before EDI becomes mandatory as a business practice. If a company is determined to implement EDI, it should look beyond just connecting two computer systems. To achieve the best return on the EDI investment, one should try to use EDI to improve the existing processes within the organization and the relationship with customers and suppliers.

References

Gerf, V.G. 1991. Prospects for electronic data interchange, *Telecommunications*, Jan., 57–60.

Mandell, M. 1991. EDI speeds Caterpillar's global march, *Computerworld*, 25(32), 58.