PART I
Automation and Control Technologies

chapter 3

Introduction to Automation

CHAPTER CONTENTS

Automation is the technology by which a process or procedure is accomplished without human assistance. It is implemented using a *program of instructions* combined with a *control system* that executes the instructions. To automate a process, *power* is required, both to drive the process itself and to operate the program and control system. Although automation can be applied in a wide variety of areas, it is most closely associated with the manufacturing industries. It was in the context of manufacturing that the term was originally coined by an engineering manager at Ford Motor Company in 1946 to describe the variety of automatic transfer devices and feed mechanisms that had been installed in Ford's production plants (Historical Note 3.1). It is ironic that nearly all modern applications of automation are controlled by computer technologies that were not available in 1946.

In this part of the book, we examine technologies that have been developed to automate manufacturing operations. The position of automation and control technologies in the larger production system is shown in Figure 3.1. In the present chapter, we provide an

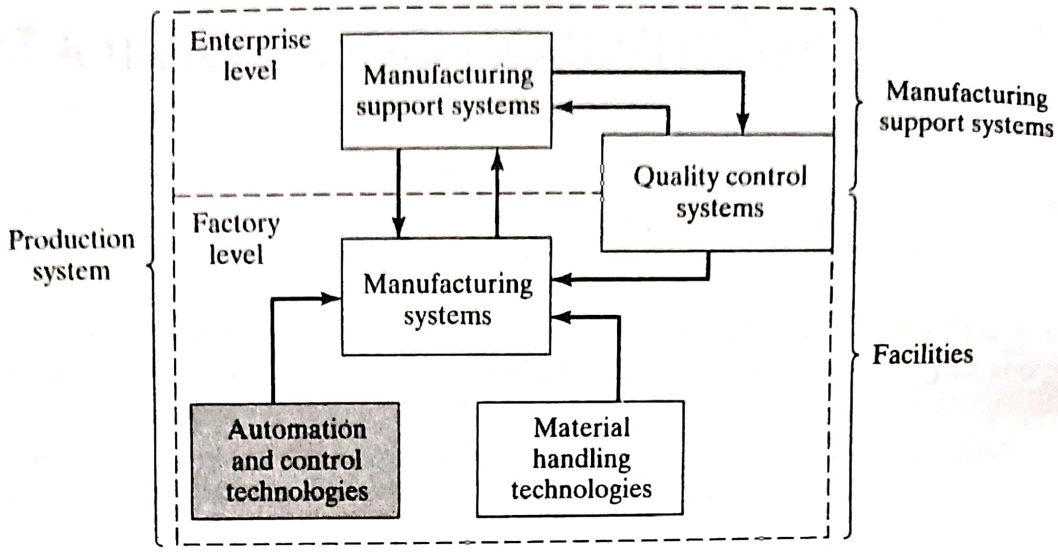

Figure 3.1 Automation and control technologies in the production system.

overview of automation: What are the elements of an automated system? What are some of the advanced features beyond the basic elements? And what are the levels in an enterprise where automation can be applied? In the following two chapters, we discuss industrial control systems and the hardware components of these systems. These two chapters serve as a foundation for the remaining chapters in our coverage of automation and control technologies. These technologies are: (1) numerical control (Chapter 6), (2) industrial robotics (Chapter 7), and (3) programmable logic controllers (Chapter 8).

Historical Note 3.1 History of automation[1]

The history of automation can be traced to the development of basic mechanical devices, such as the wheel (circa 3200 B.C.), lever, winch (circa 600 B.C.), cam (circa A.D. 1000), screw (A.D. 1405), and gear in ancient and medieval times. These basic devices were refined and used to construct the mechanisms in waterwheels, windmills (circa A.D. 650), and steam engines (A.D. 1765). These machines generated the power to operate other machinery of various kinds, such as flour mills (circa 85 B.C.), weaving machines (flying shuttle, 1733), machine tools (boring mill, 1775), steamboats (1787), and railroad locomotives (1803). *Power*, and the capacity to generate it and transmit it to operate a process, is one of the three basic elements of an automated system.

After his first steam engine in 1765, James Watt and his partner, Matthew Boulton, made several improvements in the design. One of the improvements was the flying-ball governor (around 1785), which provided feedback to control the throttle of the engine. The governor consisted of a ball on the end of a hinged lever attached to the rotating shaft. The lever was connected to the throttle valve. As the speed of the rotating shaft increased, the ball was forced to move outward by centrifugal force; this in turn caused the lever to reduce the valve opening and slow the motor speed. As rotational speed decreased, the ball and lever relaxed, thus allowing the valve to open. The flying-ball governor was one of the first examples in engineering of feedback control, an important type of *control system*—the second basic element of an automated system.

The third basic element of an automated system is for the actions of the system or machine to be directed by a *program of instructions*. One of the first examples of machine pro-

[1] Sources of most of the dates in this Historical Note: (1) R. Platt, ***Smithsonian Visual Timeline of Inventions*** (London: Dorling Kindersley Ltd., 1994); and (2) "The Power of Invention," ***Newsweek Special Issue***, Winter 1997–98 (pp. 6-79).

gramming was the Jacquard loom, invented around 1800. This loom was a machine for weaving cloth from yarn. The program of instructions that determined the weaving pattern of the cloth consisted of a metal plate containing holes. The hole pattern in the plate directed the shuttle motions of the loom, which in turn determined the weaving pattern. Different hole patterns yielded different cloth patterns. Thus, the Jacquard loom was a programmable machine, one of the first of its kind.

By the early 1800s, the three basic elements of automated systems—power source, controls, and programmable machines—had been developed, although these elements were primitive by today's standards. It took many years of refinement and many new inventions and developments, both in these basic elements as well as in the enabling infrastructure of the manufacturing industries, before fully automated production systems were to become a common reality. Important examples of these inventions and developments include *interchangeable parts* (circa 1800, Historical Note 2.1); *electrification* (starting in 1881); the *moving assembly line* (1913, Historical Note 17.1); mechanized *transfer lines* for mass production, whose programs were fixed by their hardware configuration (1924, Historical Note 18.1); a mathematical theory of *control systems* (1930s and 1940s); and the MARK I electromechanical *computer* at Harvard University (1944). These inventions and developments had all been realized by the end of World War II.

Since 1945, many new inventions and developments have contributed significantly to automation technology. Del Harder coined the word *automation* around 1946 in reference to the many automatic devices that the Ford Motor Company had developed for its production lines. The first electronic digital computer was developed at University of Pennsylvania in 1946. The first *numerical control* machine tool was developed and demonstrated in 1952 at Massachusetts Institute of Technology based on a concept proposed by John Parsons and Frank Stulen (Historical Note 6.1). By the late 1960s and early 1970s, digital computers were being connected to machine tools. In 1954, the first *industrial robot* was designed and patented (issued 1961) by George Devol (Historical Note 7.1). The first commercial robot was installed to unload parts in a die casting operation in 1961. In the late 1960s, the first *flexible manufacturing system* in the United States was installed at Ingersoll Rand Company to perform machining operations on a variety of parts (Historical Note 16.1). Around 1969, the first *programmable logic controller* was introduced (Historical Note 8.1). In 1978, the first commercial *personal computer* (PC) had been introduced by Apple Computer, although a similar product had been introduced in kit form as early as 1975.

Developments in computer technology were made possible by advances in electronics, including the *transistor* (1948), *hard disk* for computer memory (1956), *integrated circuits* (1960), the *microprocessor* (1971), *random accesss memory* (1984), megabyte capacity memory chips (circa 1990), and the *Pentium* microprocessors (1993). Software developments related to automation have been equally important, including the *FORTRAN* computer programming language (1955), the *APT* programming language for numerical control (NC) machine tools (1961), the *UNIX* operating system (1969), the *VAL* language for robot programming (1979), Microsoft *Windows* (1985), and the *JAVA* programming language (1995). Advances and enhancements in these technologies continue.

BASIC ELEMENTS OF AN AUTOMATED SYSTEM

An automated system consists of three basic elements: (1) *power* to accomplish the process and operate the system, (2) a *program of instructions* to direct the process, and (3) a *control system* to actuate the instructions. The relationship amongst these elements is illustrated in Figure 3.2. All systems that qualify as being automated include these three basic elements in one form or another.

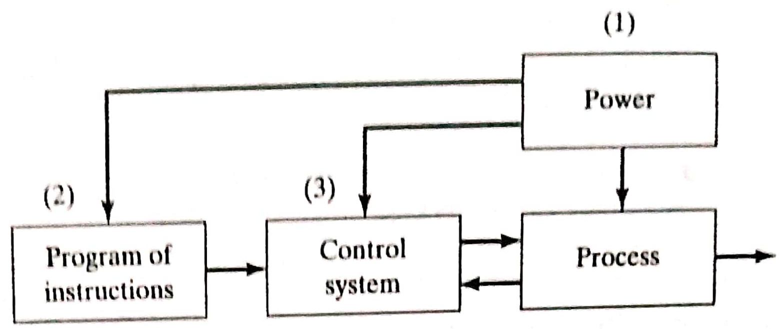

Figure 3.2 Elements of an automated system: (1) power, (2) program of instructions, and (3) control systems.

3.1.1 Power to Accomplish the Automated Process

An automated system is used to operate some process, and power is required to drive the process as well as the controls. The principal source of power in automated systems is electricity. Electric power has many advantages in automated as well as nonautomated processes:

- Electrical power is widely available at moderate cost. It is an important part of our industrial infrastructure.
- Electrical power can be readily converted to alternative energy forms: mechanical, thermal, light, acoustic, hydraulic, and pneumatic.
- Electrical power at low levels can be used to accomplish functions such as signal transmission, information processing, and data storage and communication.
- Electrical energy can be stored in long-life batteries for use in locations where an external source of electrical power is not conveniently available.

Alternative power sources include fossil fuels, solar energy, water, and wind. However, their exclusive use is rare in automated systems. In many cases when alternative power sources are used to drive the process itself, electrical power is used for the controls that automate the operation. For example, in casting or heat treatment, the furnace may be heated by fossil fuels, but the control system to regulate temperature and time cycle is electrical. In other cases, the energy from these alternative sources is converted to electric power to operate both the process and its automation. When solar energy is used as a power source for an automated system, it is generally converted in this way.

Power for the Process. In production, the term *process* refers to the manufacturing operation that is performed on a work unit. In Table 3.1, a list of common manufacturing processes is compiled along with the form of power required and the resulting action on the work unit. Most of the power in manufacturing plants is consumed by these kinds of operations. The "power form" indicated in the middle column of the table refers to the energy that is applied directly to the process. As indicated above, the power source for each operation is usually converted from electricity.

In addition to driving the manufacturing process itself, power is also required for the following material handling functions:

- *Loading and unloading the work unit.* All of the processes listed in Table 3.1 are accomplished on discrete parts. These parts must be moved into the proper position

TABLE 3.1 Common Manufacturing Processes and Their Power Requirements

Process	Power Form	Action Accomplished
Casting	Thermal	Melting the metal before pouring into a mold cavity where solidification occurs.
Electric discharge machining (EDM)	Electrical	Metal removal is accomplished by a series of discrete electrical discharges between electrode (tool) and workpiece. The electric discharges cause very high localized temperatures that melt the metal.
Forging	Mechanical	Metal workpart is deformed by opposing dies. Workparts are often heated in advance of deformation, thus thermal power is also required.
Heat treating	Thermal	Metallic work unit is heated to temperature below melting point to effect microstructural changes.
Injection molding	Thermal and mechanical	Heat is used to raise temperature of polymer to highly plastic consistency, and mechanical force is used to inject the polymer melt into a mold cavity.
Laser beam cutting	Light and thermal	A highly coherent light beam is used to cut material by vaporization and melting.
Machining	Mechanical	Cutting of metal is accomplished by relative motion between tool and workpiece.
Sheet metal punching and blanking	Mechanical	Mechanical power is used to shear metal sheets and plates.
Welding	Thermal (maybe mechanical)	Most welding processes use heat to cause fusion and coalescence of two (or more) metal parts at their contacting surfaces. Some welding processes also apply mechanical pressure to the surfaces.

and orientation for the process to be performed, and power is required for this transport and placement function. At the conclusion of the process, the work unit must similarly be removed. If the process is completely automated, then some form of mechanized power is used. If the process is manually operated or semiautomated, then human power may be used to position and locate the work unit.

- *Material transport between operations.* In addition to loading and unloading at a given operation, the work units must be moved between operations. We consider the material handling technologies associated with this transport function in Chapter 10.

Power for Automation. Above and beyond the basic power requirements for the manufacturing operation, additional power is required for automation. The additional power is used for the following functions:

- *Controller unit.* Modern industrial controllers are based on digital computers, which require electrical power to read the program of instructions, make the control calculations, and execute the instructions by transmitting the proper commands to the actuating devices.
- *Power to actuate the control signals.* The commands sent by the controller unit are carried out by means of electromechanical devices, such as switches and motors, called *actuators* (Section 5.2). The commands are generally transmitted by means of low-voltage control signals. To accomplish the commands, the actuators require more power,

and so the control signals must be amplified to provide the proper power level for the actuating device.

- *Data acquisition and information processing.* In most control systems, data must be collected from the process and used as input to the control algorithms. In addition, a requirement of the process may include keeping records of process performance or product quality. These data acquisition and record keeping functions require power, although in modest amounts.

3.1.2 Program of Instructions

The actions performed by an automated process are defined by a program of instructions. Whether the manufacturing operation involves low, medium, or high production (Section 1.1), each part or product style made in the operation requires one or more processing steps that are unique to that style. These processing steps are performed during a work cycle. A new part is completed during each work cycle (in some manufacturing operations, more than one part is produced during the work cycle; e.g., a plastic injection molding operation may produce multiple parts each cycle using a multiple cavity mold). The particular processing steps for the work cycle are specified in a *work cycle program*. Work cycle programs are called *part programs* in numerical control (Chapter 6). Other process control applications use different names for this type of program.

Work Cycle Programs. In the simplest automated processes, the work cycle consists of essentially one step, which is to maintain a single process parameter at a defined level, for example, maintain the temperature of a furnace at a designated value for the duration of a heat treatment cycle. (We assume that loading and unloading of the work units into and from the furnace is performed manually and is therefore not part of the automatic cycle.) In this case, programming simply involves setting the temperature dial on the furnace. To change the program, the operator simply changes the temperature setting. An extension of this simple case is when the single-step process is defined by more than one process parameter, for example, a furnace in which both temperature and atmosphere are controlled.

In more complicated systems, the process involves a work cycle consisting of multiple steps that are repeated with no deviation from one cycle to the next. Most discrete part manufacturing operations are in this category. A typical sequence of steps (simplified) is: (1) load the part into the production machine, (2) perform the process, and (3) unload the part. During each step, there are one or more activities that involve changes in one or more process parameters. *Process parameters* are inputs to the process, such as temperature setting of a furnace, coordinate axis value in a positioning system, valve opened or closed in a fluid flow system, and motor on or off. Process parameters are distinguished from *process variables*, which are outputs from the process; for example, the actual temperature of the furnace, the actual position of the axis, the actual flow rate of the fluid in the pipe, and the rotational speed of the motor. As our list of examples suggests, the changes in process parameter values may be continuous (gradual changes during the processing step; for example, gradually increasing temperature during a heat treatment cycle) or discrete (stepwise changes; for example, on/off). Different process parameters may be involved in each step.

EXAMPLE 3.1 An Automated Turning Operation

Consider an automated turning operation in which a cone-shaped geometry is generated. Assume the system is automated and that a robot is used to load and unload the work unit. The work cycle consists of the following steps: (1) load starting workpiece, (2) position cutting tool prior to turning, (3) turn, (4) reposition tool to a safe location at end of turning, and (5) unload finished workpiece. Identify the activity(ies) and process parameter(s) in each step of the operation.

Solution: In step (1), the activities consist of the robot manipulator reaching for the raw workpart, lifting and positioning the part into the chuck jaws of the lathe, then removing the manipulator to a safe position to await unloading. The process parameters for these activities are the axis values of the robot manipulator (which change continuously), the gripper value (open or closed), and the chuck jaw value (open or closed).

In step (2), the activity involves the movement of the cutting tool to a "ready" position. The process parameters associated with this activity are the x- and z-axis position of the tool.

Step (3) is the turning operation. It requires the simultaneous control of three process parameters: rotational speed of the workpiece (rev/min), feed (mm/rev), and radial distance of the cutting tool from the axis of rotation. To cut the conical shape, radial distance must be changed continuously at a constant rate for each revolution of the workpiece. For a consistent finish on the surface, the rotational speed must be continuously adjusted to maintain a constant surface speed (m/min); and for equal feed marks on the surface, the feed must be set at a constant value. Depending on the angle of the cone, multiple turning passes may be required to gradually generate the desired contour. Each pass represents an additional step in the sequence.

Steps (4) and (5) involve the reverse activities as steps (2) and (1), respectively, and the process parameters are the same.

Many production operations consist of multiple steps, sometimes more complicated than our turning example. Examples of these operations include automatic screw machine cycles, sheet metal stamping operations, plastic injection molding, and die casting. Each of these manufacturing processes has been used for many decades. In earlier versions of these operations, the work cycles were controlled by hardware components, such as limit switches, timers, cams, and electromechanical relays. In effect, the hardware components and their arrangements served as the program of instructions that directed the sequence of steps in the processing cycle. Although these devices were quite adequate in performing their sequencing function, they suffered from the following disadvantages: (1) They often required considerable time to design and fabricate, thus forcing the production equipment to be used for batch production only; (2) making even minor changes in the program was difficult and time consuming; and (3) the program was in a physical form that is not readily compatible with computer data processing and communication.

Modern controllers used in automated systems are based on digital computers. Instead of cams, timers, relays, and other hardware devices, the programs for computer-controlled equipment are contained in magnetic tape, diskettes, compact disks (CD-ROMs), computer memory, and other modern storage technologies. Virtually all new equipment that perform the above mass production operations are designed with some type of computer controller to execute their respective processing cycles. The use of digital computers as the process controller allows improvements and upgrades to be made in the control programs, such as the addition of control functions not foreseen during initial equipment design. These kinds of control changes are often difficult to make with the previous hardware devices.

The work cycle may include manual steps, where the operator performs certain activities during the work cycle, and the automated system performs the rest. A common example is the loading and unloading of parts by the operator into and from a numerical control machine between machining cycles, where the machine performs the cutting operation under part program control. Initiation of the cutting operation of each cycle is triggered by the operator activating a "start" button after the part has been loaded.

Decision-Making in the Programmed Work Cycle. In our previous discussion of automated work cycles, the only two features of the work cycle are (1) the number and sequence of processing steps and (2) the process parameter changes in each step. Each work cycle consists of the same steps and associated process parameter changes with no variation from one cycle to the next. The program of instructions is repeated each work cycle without deviation. In fact, many automated manufacturing operations require decisions to be made during the programmed work cycle to cope with variations in the cycle. In many cases, the variations are routine elements of the cycle, and the corresponding instructions for dealing with them are incorporated into the regular part program. These cases include:

- *Operator interaction.* Although the program of instructions is intended to be carried out without human interaction, the controller unit may require input data from a human operator in order to function. For example, in an automated engraving operation, the operator may have to enter the alphanumeric characters that are to be engraved on the work unit (e.g., plaque, trophy, belt buckle). Having entered the characters, the engraving operation is accomplished automatically by the system. (An everyday example of operator interaction with an automated system is a bank customer using an automated teller machine. The customer must enter the codes indicating what transaction is to be accomplished by the teller machine.)
- *Different part or product styles processed by the system.* In this instance, the automated system is programmed to perform different work cycles on different part or product styles. An example is an industrial robot that performs a series of spot welding operations on car bodies in a final assembly plant. These plants are often designed to build different body styles on the same automated assembly line, such as two-door and four-door sedans. As each car body enters a given welding station on the line, sensors identify which style it is, and the robot performs the correct series of welds for that style.
- *Variations in the starting work units.* In many manufacturing operations the starting work units are not consistent. A good example is a sand casting as the starting work unit in a machining operation. The dimensional variations in the raw castings sometimes necessitate an extra machining pass to bring the machined dimension to the specified value. The part program must be coded to allow for the additional pass when necessary.

In all of these examples, the routine variations can be accommodated in the regular work cycle program. The program can be designed to respond to sensor or operator inputs by executing the appropriate subroutine corresponding to the input. In other cases, the variations in the work cycle are not routine at all. They are infrequent and unexpected, such as the failure of an equipment component. In these instances, the program must include contingency procedures or modifications in the sequence to cope with conditions that lie outside the normal routine. We discuss these measures later in the chapter in the context of advanced automation functions (Section 3.2).

TABLE 3.2 Features of Work Cycle Programs Used in Automated Systems

Program Feature	*Examples or Alternatives*
Steps in work cycle	Example: • Typical sequence of steps: (1) load, (2), process, (3) unload
Process parameters (inputs) in each step	Alternatives: • One parameter versus multiple parameters that must be changed during the step • Continuous parameters versus discrete parameters • Parameters that change during the step; for example, a positioning system whose axes values change during the processing step
Manual steps in work cycle	Alternatives: • Manual steps versus no manual steps (completely automated work cycle) Example: • Operator loading and unloading parts to and from machine
Operator interaction	Alternatives: • Operator interaction versus completely automated work cycle Example: • Operator entering processing information for current workpart
Different part or product styles	Alternatives: • Identical part or product style each cycle (mass or batch production) versus different part or product styles each cycle (flexible automation)
Variations in starting work units	Example: • Variations in starting dimensions or part features

A variety of production situations and work cycle programs has been discussed here. The features of work cycle programs (part programs) used to direct the operations of an automated system are summarized as in Table 3.2.

3.1.3 Control System

The control element of the automated system executes the program of instructions. The control system causes the process to accomplish its defined function, which for our purpose is to carry out some manufacturing operation. Let us provide a brief introduction to control systems here. The following chapter describes this important industrial technology in more detail.

The controls in an automated system can be either closed loop or open loop. A *closed loop control system*, also known as a *feedback control system*, is one in which the output variable is compared with an input parameter, and any difference between the two is used to drive the output into agreement with the input. As shown in Figure 3.3, a closed loop control system consists of six basic elements: (1) input parameter, (2) process, (3) output variable, (4) feedback sensor, (5) controller, and (6) actuator. The *input parameter*, often referred to as the *set point*, represents the desired value of the output. In a home temperature control system, the set point is the desired thermostat setting. The *process* is the operation or function being controlled. In particular, it is the *output variable* that is being controlled in the loop. In the present discussion, the process of interest is usually a manufacturing operation, and the output variable is some process variable, perhaps a critical performance

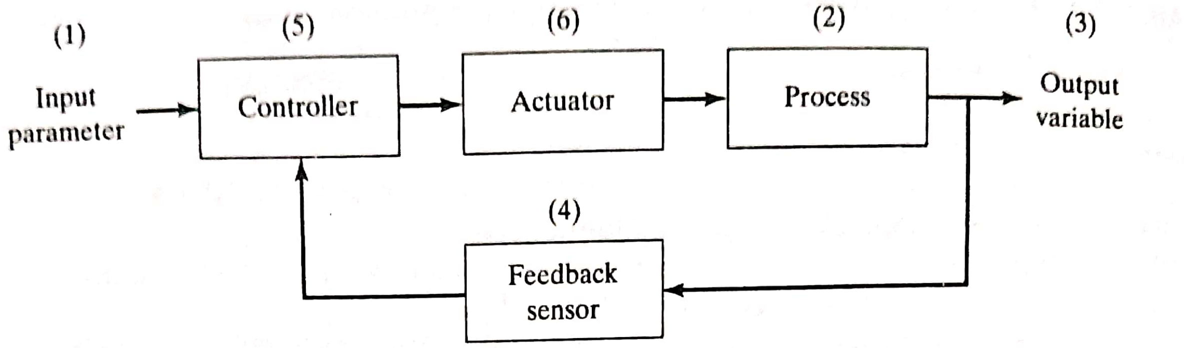

Figure 3.3 A feedback control system.

measure in the process, such as temperature or force or flow rate. A *sensor* is used to measure the output variable and close the loop between input and output. Sensors perform the feedback function in a closed loop control system. The *controller* compares the output with the input and makes the required adjustment in the process to reduce the difference between them. The adjustment is accomplished using one or more *actuators*, which are the hardware devices that physically carry out the control actions, such as an electric motor or a flow valve. It should be mentioned that our model in Figure 3.3 shows only one loop. Most industrial processes require multiple loops, one for each process variable that must be controlled.

In contrast to the closed loop control system, an *open loop control system* operates without the feedback loop, as in Figure 3.4. In this case, the controls operate without measuring the output variable, so no comparison is made between the actual value of the output and the desired input parameter. The controller relies on an accurate model of the effect of its actuator on the process variable. With an open loop system, there is always the risk that the actuator will not have the intended effect on the process, and that is the disadvantage of an open loop system. Its advantage is that it is generally simpler and less expensive than a closed loop system. Open loop systems are usually appropriate when the following conditions apply: (1) The actions performed by the control system are simple, (2) the actuating function is very reliable, and (3) any reaction forces opposing the actuation are small enough to have no effect on the actuation. If these characteristics are not applicable, then a closed loop control system may be more appropriate.

Consider the difference between a closed loop and open loop system for the case of a positioning system. Positioning systems are common in manufacturing to locate a workpart relative to a tool or workhead. Figure 3.5 illustrates the case of a closed loop posi-

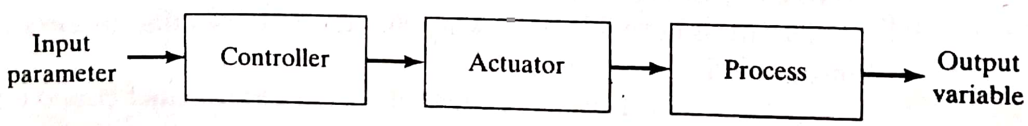

Figure 3.4 An open loop control system.

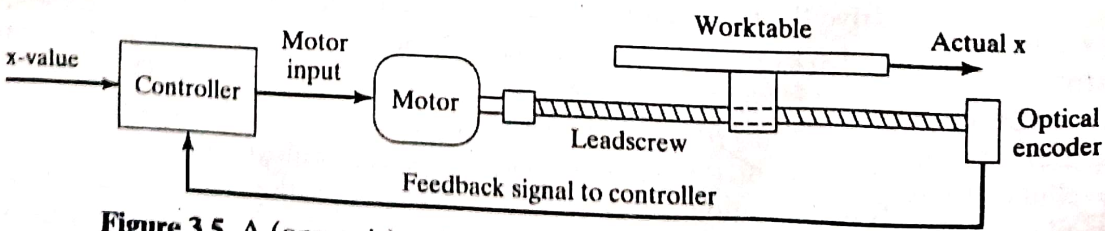

Figure 3.5 A (one-axis) positioning system consisting of a leadscrew driven by a dc servomotor.

tioning system. In operation, the system is directed to move the worktable to a specified location as defined by a coordinate value in a Cartesian (or other) coordinate system. Most positioning systems have at least two axes (e.g., an $x-y$ positioning table) with a control system for each axis, but our diagram only illustrates one of these axes. A dc servomotor connected to a leadscrew is a common actuator for each axis. A signal indicating the coordinate value (e.g., x-value) is sent from the controller to the motor that drives the leadscrew, whose rotation is converted into linear motion of the positioning table. As the table moves closer to the desired x-coordinate value, the difference between the actual x-position and the input x-value is reduced. The actual x-position is measured by a feedback sensor (e.g., an optical encoder). The controller continues to drive the motor until the actual table position corresponds to the input position value.

For the open loop case, the diagram for the positioning system would be similar to the preceding, except that no feedback loop is present and a stepper motor is used in place of the dc servomotor. A stepper motor is designed to rotate a precise fraction of a turn for each pulse received from the controller. Since the motor shaft is connected to the leadscrew, and the leadscrew drives the worktable, each pulse converts into a small constant linear movement of the table. To move the table a desired distance, the number of pulses corresponding to that distance is sent to the motor. Given the proper application, whose characteristics match the preceding list of operating conditions, an open loop positioning system works with high reliability.

We consider the engineering analysis of closed loop and open loop positioning systems in the context of numerical control in a subsequent chapter (Section 6.6).

ADVANCED AUTOMATION FUNCTIONS

In addition to executing work cycle programs, an automated system may be capable of executing advanced functions that are not specific to a particular work unit. In general, the functions are concerned with enhancing the performance and safety of the equipment. Advanced automation functions include the following: (1) safety monitoring, (2) maintenance and repair diagnostics, and (3) error detection and recovery.

Advanced automation functions are made possible by special subroutines included in the program of instructions. In some cases, the functions provide information only and do not involve any physical actions by the control system. An example of this case includes reporting a list of preventive maintenance tasks that should be accomplished. Any actions taken on the basis of this report are decided by the human operators and managers of the system and not by the system itself. In other cases, the program of instructions must be physically executed by means of the control system using available actuators. A simple example of this case is a safety monitoring system that sounds an alarm when a human worker gets dangerously close to the automated system.

3.2.1 Safety Monitoring

One of the significant reasons for automating a manufacturing operation is to remove worker(s) from a hazardous working environment. An automated system is often installed to perform a potentially dangerous operation that would otherwise be accomplished manually by human workers. However, even in automated systems, workers are still needed to service the system, at periodic time intervals if not full-time. Accordingly, it is important that the automated system be designed to operate safely when workers are in attendance. In addition, it is essential that the automated system carry out its process in a way that is not self-destructive. Thus, there are two reasons for providing an automated system with a safety monitoring capability: (1) to protect human workers in the vicinity of the system and (2) to protect the equipment associated with the system.

Safety monitoring means more than the conventional safety measures taken in a manufacturing operation, such as protective shields around the operation or the kinds of manual devices that might be utilized by human workers, such as emergency stop buttons. *Safety monitoring* in an automated system involves the use of sensors to track the system's operation and identify conditions and events that are unsafe or potentially unsafe. The safety monitoring system is programmed to respond to unsafe conditions in some appropriate way. Possible responses to various hazards might include one or more of the following:

- complete stoppage of the automated system
- sounding an alarm
- reducing the operating speed of the process
- taking corrective actions to recover from the safety violation

This last response is the most sophisticated and is suggestive of an intelligent machine performing some advanced strategy. This kind of response is applicable to a variety of possible mishaps, not necessarily confined to safety issues, and is called error detection and recovery (Section 3.2.3).

Sensors for safety monitoring range from very simple devices to highly sophisticated systems. The topic of sensor technology is discussed in Chapter 5 (Section 5.1). The following list suggests some of the possible sensors and their applications for safety monitoring:

- Limit switches to detect proper positioning of a part in a workholding device so that the processing cycle can begin.
- Photoelectric sensors triggered by the interruption of a light beam; this could be used to indicate that a part is in the proper position or to detect the presence of a human intruder into the work cell.
- Temperature sensors to indicate that a metal workpart is hot enough to proceed with a hot forging operation. If the workpart is not sufficiently heated, then the metal's ductility may be too low, and the forging dies might be damaged during the operation.
- Heat or smoke detectors to sense fire hazards.
- Pressure-sensitive floor pads to detect human intruders into the work cell.
- Machine vision systems to supervise the automated system and its surroundings.

It should be mentioned that a given safety monitoring system is limited in its ability to respond to hazardous conditions by the possible irregularities that have been foreseen by the system designer. If the designer has not anticipated a particular hazard, and consequently has not provided the system with the sensing capability to detect that hazard, then the safety monitoring system cannot recognize the event if and when it occurs.

3.2.2 Maintenance and Repair Diagnostics

Modern automated production systems are becoming increasingly complex and sophisticated, thus complicating the problem of maintaining and repairing them. *Maintenance and repair diagnostics* refers to the capabilities of an automated system to assist in the identification of the source of potential or actual malfunctions and failures of the system. Three modes of operation are typical of a modern maintenance and repair diagnostics subsystem:

1. *Status monitoring*. In the status monitoring mode, the diagnostic subsystem monitors and records the status of key sensors and parameters of the system during normal operation. On request, the diagnostics subsystem can display any of these values and provide an interpretation of current system status, perhaps warning of an imminent failure.

2. *Failure diagnostics*. The failure diagnostics mode is invoked when a malfunction or failure occurs. Its purpose is to interpret the current values of the monitored variables and to analyze the recorded values preceding the failure so that the cause of the failure can be identified.

3. *Recommendation of repair procedure*. In the third mode of operation, the subsystem provides a recommended procedure to the repair crew as to the steps that should be taken to effect repairs. Methods for developing the recommendations are sometimes based on the use of expert systems in which the collective judgments of many repair experts are pooled and incorporated into a computer program that uses artificial intelligence techniques.

Status monitoring serves two important functions in machine diagnostics: (1) providing information for diagnosing a current failure and (2) providing data to predict a future malfunction or failure. First, when a failure of the equipment has occurred, it is usually difficult for the repair crew to determine the reason for the failure and what steps should be taken to make repairs. It is often helpful to reconstruct the events leading up to the failure. The computer is programmed to monitor and record the variables and to draw logical inferences from their values about the reason for the malfunction. This diagnosis helps the repair personnel make the necessary repairs and replace the appropriate components. This is especially helpful in electronic repairs where it is often difficult to determine on the basis of visual inspection which components have failed.

The second function of status monitoring is to identify signs of an impending failure, so that the affected components can be replaced before failure actually causes the system to go down. These part replacements can be made during the night shift or other time when the process is not operating, with the result that the system experiences no loss of regular operation.

3.2.3 Error Detection and Recovery

In the operation of any automated system, there are hardware malfunctions and unexpected events that occur during operation. These events can result in costly delays and loss of production until the problem has been corrected and regular operation is restored. Traditionally, equipment malfunctions are corrected by human workers, perhaps with the aid of a maintenance and repair diagnostics subroutine. With the increased use of computer control for manufacturing processes, there is a trend toward using the control computer not only to diagnose the malfunctions but also to automatically take the necessary corrective action to restore the system to normal operation. The term *error detection and recovery* is used when the computer performs these functions.

Error Detection. As indicated by the term, error detection and recovery consists of two steps: (1) error detection and (2) error recovery. The *error detection* step uses the automated system's available sensor systems to determine when a deviation or malfunction has occurred, correctly interpret the sensor signal(s), and classify the error. Design of the error detection subsystem must begin with a classification of the possible errors that can occur during system operation. The errors in a manufacturing process tend to be very application specific. They must be anticipated in advance in order to select sensors that will enable their detection.

In analyzing a given production operation, the possible errors can be classified into one of three general categories: (1) random errors, (2) systematic errors, and (3) aberrations. *Random errors* occur as a result of the normal stochastic nature of the process. These errors occur when the process is in statistical control (Section 21.1). Large variations in part dimensions, even when the production process is in statistical control, can cause problems in downstream operations. By detecting these deviations on a part-by-part basis, corrective action can be taken in subsequent operations. *Systematic errors* are those that result from some assignable cause such as a change in raw material properties or a drift in an equipment setting. These errors usually cause the product to deviate from specifications so as to be unacceptable in quality terms. Finally, the third type of error, *aberrations*, results from either an equipment failure or a human mistake. Examples of equipment failures include fracture of a mechanical shear pin, bursts in a hydraulic line, rupture of a pressure vessel, and sudden failure of a cutting tool. Examples of human mistakes include errors in the control program, improper fixture setups, and substitution of the wrong raw materials.

The two main design problems in error detection are: (1) to anticipate all of the possible errors that can occur in a given process and (2) to specify the appropriate sensor systems and associated interpretive software so that the system is capable of recognizing each error. Solving the first problem requires a systematic evaluation of the possibilities under each of the three error classifications. If the error has not been anticipated, then the error detection subsystem cannot correctly detect and identify it.

EXAMPLE 3.2 Error Detection in an Automated Machining Cell

Consider an automated cell consisting of a CNC machine tool, a parts storage unit, and a robot for loading and unloading the parts between the machine and the storage unit. Possible errors that might affect this system can be divided into the following categories: (1) machine and process, (2) cutting tools, (3) workholding fixture, (4) part storage unit, and (5) load/unload robot. Develop a list of possible errors (deviations and malfunctions) that might be included in each of these five categories.

Solution: A list of possible errors in the machining cell is presented in Table 3.3.

Error Recovery. *Error recovery* is concerned with applying the necessary corrective action to overcome the error and bring the system back to normal operation. The problem of designing an error recovery system focuses on devising appropriate strategies and procedures that will either correct or compensate for the variety of errors that can occur in the process. Generally, a specific recovery strategy and procedure must be designed for each different error. The types of strategies can be classified as follows:

1. *Make adjustments at the end of the current work cycle.* When the current work cycle is completed, the part program branches to a corrective action subroutine specifically

TABLE 3.3 Error Detection Step in an Automated Machining Cell: Error Categories and Possible Malfunctions Within Each Category

Error Categories	*Possible Malfunctions*
1. Machine and process	Loss of power, power overload, thermal deflection, cutting temperature too high, vibration, no coolant, chip fouling, wrong part program, defective part
2. Cutting tools	Tool breakage, tool wear-out, vibration, tool not present, wrong tool
3. Workholding fixture	Part not in fixture, clamps not actuated, part dislodged during machining, part deflection during machining, part breakage, chips causing location problems
4. Part storage unit	Workpart not present, wrong workpart, oversized or undersized workpart
5. Load/unload robot	Improper grasping of workpart, robot drops workpart, no part present at pickup

designed for the error detected, executes the subroutine, and then returns to the work cycle program. This action reflects a low level of urgency and is most commonly associated with random errors in the process.

2. *Make adjustments during the current cycle.* This generally indicates a higher level of urgency than the preceding type. In this case, the action to correct or compensate for the detected error is initiated as soon as the error is detected. However, it must be possible to accomplish the designated corrective action while the work cycle is still being executed.
3. *Stop the process to invoke corrective action.* In this case, the deviation or malfunction requires that the execution of the work cycle be suspended during corrective action. It is assumed that the system is capable of automatically recovering from the error without human assistance. At the end of the corrective action, the regular work cycle is continued.
4. *Stop the process and call for help.* In this case, the error requiring stoppage of the process cannot be resolved through automated recovery procedures. This situation arises because: (1) the automated cell is not enabled to correct the problem or (2) the error cannot be classified into the predefined list of errors. In either case, human assistance is required to correct the problem and restore the system to fully automated operation.

Error detection and recovery requires an interrupt system (Section 4.3.2). When an error in the process is sensed and identified, an interrupt in the current program execution is invoked to branch to the appropriate recovery subroutine. This is done either at the end of the current cycle (type 1 above) or immediately (types 2, 3, and 4). At the completion of the recovery procedure, program execution reverts back to normal operation.

EXAMPLE 3.3 Error Recovery in an Automated Machining Cell

For the automated cell of Example 3.2, develop a list of possible corrective actions that might be taken by the system to address certain of the errors.

Solution: A list of possible corrective actions is presented in Table 3.4.

TABLE 3.4 Error Recovery in an Automated Machining Cell: Possible Corrective Actions That Might Be Taken in Response to Errors Detected During the Operation

Errors Detected	Possible Corrective Actions to Recover
Part dimensions deviating due to thermal deflection of machine tool	Adjust coordinates in part program to compensate (category 1 corrective action)
Part dropped by robot during pickup	Reach for another part (category 2 corrective action)
Part is dimensionally oversized	Adjust part program to take a preliminary machining pass across the work surface (category 2 corrective action)
Chatter (tool vibration)	Increase or decrease cutting speed to change harmonic frequency (category 2 corrective action)
Cutting temperature too high	Reduce cutting speed (category 2 corrective action)
Failure of cutting tool	Replace cutting tool with another sharp tool (category 3 corrective action).
No more parts in parts storage unit	Call operator to resupply starting workparts (category 4 corrective action)
Chips fouling machining operation	Call operator to clear chips from work area (category 4 corrective action)

3.3 LEVELS OF AUTOMATION

The concept of automated systems can be applied to various levels of factory operations. One normally associates automation with the individual production machines. However, the production machine itself is made up of subsystems that may themselves be automated. For example, one of the important automation technologies we discuss in this part of the book is numerical control (Chapter 6). A modern numerical control (NC) machine tool is an automated system. However, the NC machine itself is composed of multiple control systems. Any NC machine has at least two axes of motion, and some machines have up to five axes. Each of these axes operates as a positioning system, as described in Section 3.1.3, and is, in effect, itself an automated system. Similarly, a NC machine is often part of a larger manufacturing system, and the larger system may itself be automated. For example, two or three machine tools may be connected by an automated part handling system operating under computer control. The machine tools also receive instructions (e.g., part programs) from the computer. Thus we have three levels of automation and control included here (the positioning system level, the machine tool level, and the manufacturing system level). For our purposes in this text, we can identify five possible levels of automation in a production plant. They are defined next, and their hierarchy is depicted in Figure 3.6.

1. *Device level.* This is the lowest level in our automation hierarchy. It includes the actuators, sensors, and other hardware components that comprise the machine level. The devices are combined into the individual control loops of the machine; for example, the feedback control loop for one axis of a CNC machine or one joint of an industrial robot.
2. *Machine level.* Hardware at the device level is assembled into individual machines. Examples include CNC machine tools and similar production equipment, industrial robots, powered conveyors, and automated guided vehicles. Control functions at this

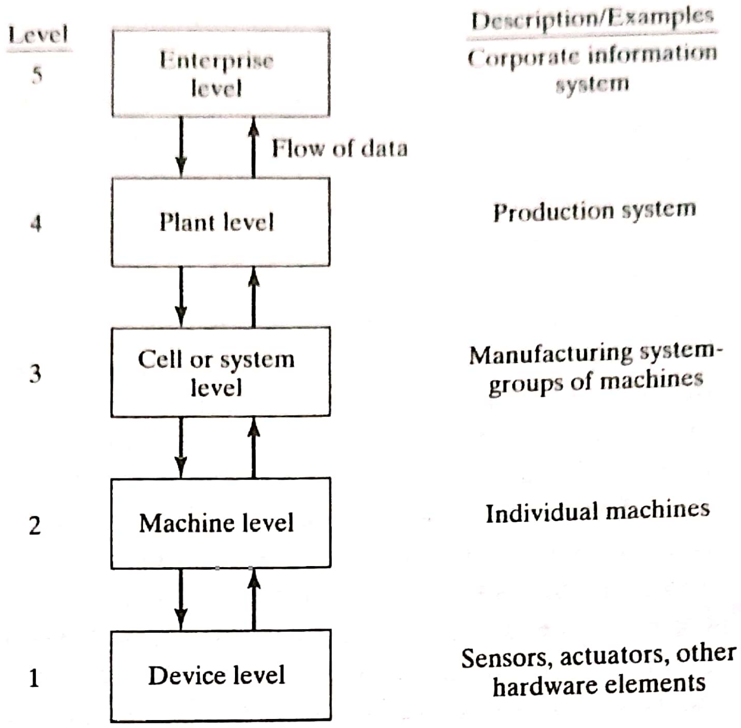

Figure 3.6 Five levels of automation and control in manufacturing.

level include performing the sequence of steps in the program of instructions in the correct order and making sure that each step is properly executed.

3. *Cell or system level.* This is the manufacturing cell or system level, which operates under instructions from the plant level. A manufacturing cell or system is a group of machines or workstations connected and supported by a material handling system, computer, and other equipment appropriate to the manufacturing process. Production lines are included in this level. Functions include part dispatching and machine loading, coordination among machines and material handling system, and collecting and evaluating inspection data.
4. *Plant level.* This is the factory or production systems level. It receives instructions from the corporate information system and translates them into operational plans for production. Likely functions include: order processing, process planning, inventory control, purchasing, material requirements planning, shop floor control, and quality control.
5. *Enterprise level.* This is the highest level, consisting of the corporate information system. It is concerned with all of the functions necessary to manage the company: marketing and sales, accounting, design, research, aggregate planning, and master production scheduling.

Most of the technologies discussed in this part of the book are at level 2 (the machine level), although we discuss level 1 automation technologies (the devices that make up a control system) in Chapter 5. The level 2 technologies include the individual controllers (e.g., programmable logic controllers and digital computer controllers), numerical control machines, and industrial robots. The material handling equipment discussed in Part II also represent technologies at level 2, although some of the handling equipment are themselves sophisticated automated systems. The automation and control issues at level 2 are concerned with the basic operation of the equipment and the physical processes they perform.

Controllers, machines, and material handling equipment are combined into manufacturing cells, or production lines, or similar systems, which make up level 3, considered in Part III. A *manufacturing system* is defined in this book as a collection of integrated equipment designed for some special mission, such as machining a defined part family or assembly of a certain product. Manufacturing systems also include people. Certain highly automated manufacturing systems can operate for extended periods of time without humans present to attend to their needs. But most manufacturing systems include workers as important elements of the system; for example, assembly workers on a conveyorized production line or part loaders/unloaders in a machining cell. Thus, manufacturing systems are designed with varying degrees of automation; some are highly automated, others are completely manual, and there is a wide range between.

The manufacturing systems in a factory are components of a larger system, which we refer to as a production system. We define a *production system* as the people, equipment, and procedures that are organized for the combination of materials and processes that comprise a company's manufacturing operations. Production systems are at level 4, the plant level, while manufacturing systems are at level 3 in our automation hierarchy. Production systems include not only the groups of machines and workstations in the factory but also the support procedures that make them work. These procedures include production control, inventory control, material requirements planning, shop floor control, and quality control. These systems are discussed in Parts IV and V. They are often implemented not only at the plant level but also at the corporate level (level 5).

REFERENCES

[1] BOUCHER, T. O., *Computer Automation in Manufacturing*, Chapman & Hall, London, 1996.

[2] GROOVER, M. P., "Automation," *Encyclopaedia Britannica, Macropaedia*, 15th Edition, Chicago, 1992. Vol. 14, pp. 548–557.

[3] GROOVER, M. P., "Automation," *Handbook of Design, Manufacturing, and Automation*, R. C. Dorf and A. Kusiak, Editors, John Wiley & Sons, Inc., New York, 1994, pp. 3–21.

[4] GROOVER, M. P., "Industrial Control Systems," *Maynard's Industrial Engineering Handbook*, Fifth Edition (forthcoming), Kjell Zandin, Editor, McGraw-Hill Book Company, New York, (forthcoming, 2000).