16 Scheduling

LEARNING OBJECTIVES

After completing this chapter, you should be able to:

LO16.1 Explain what scheduling involves and the importance of good scheduling.

LO16.2 Compare product and service scheduling hierarchies.

LO16.3 Describe scheduling needs in high-volume systems.

LO16.4 Describe scheduling needs in intermediate-volume systems.

LO16.5 Describe scheduling needs in job shops.

LO16.6 Use and interpret Gantt charts.

LO16.7 Use the assignment method for loading.

LO16.8 Give examples of commonly used priority rules.

LO16.9 Discuss the theory of constraints and that approach to scheduling.

LO16.10 Summarize some of the unique problems encountered in service systems, and describe some of the approaches used for scheduling service systems.

CHAPTER OUTLINE

© U. Baumgarten via Getty

Airline travel can be difficult when flights are delayed or canceled due to weather problems. And even though it may be clear and dry in some areas, flights in those places can still be affected by weather in other areas. Because of all the interdependencies, a problem in one area, especially around major hub airports like Chicago, Atlanta, and New York, has a cascading effect with impacts throughout the nation. This results in massive scheduling problems. Flight arrivals and departures have to be rescheduled, which then means flight crews, terminal gates, connections, and baggage and freight also must be rescheduled. Airline and air traffic control software scheduling systems include and optimize thousands of variables.

Within an organization, **scheduling** pertains to establishing the timing of the use of specific resources of that organization. It relates to the use of equipment, facilities, and human activities. Scheduling occurs in every organization, regardless of the nature of its activities. For example, manufacturers must schedule production, which means developing schedules for workers, equipment, purchases, maintenance, and so on. Hospitals must schedule admissions, surgery, nursing assignments, and support services such as meal preparation, security, maintenance, and cleaning. Educational institutions must schedule classrooms, instruction, and students. And lawyers, doctors, dentists, hairdressers, and auto repair shops must schedule appointments.

LO16.1 Explain what scheduling involves and the importance of good scheduling.

Scheduling Establishing the timing of the use of equipment, facilities, and human activities in an organization.

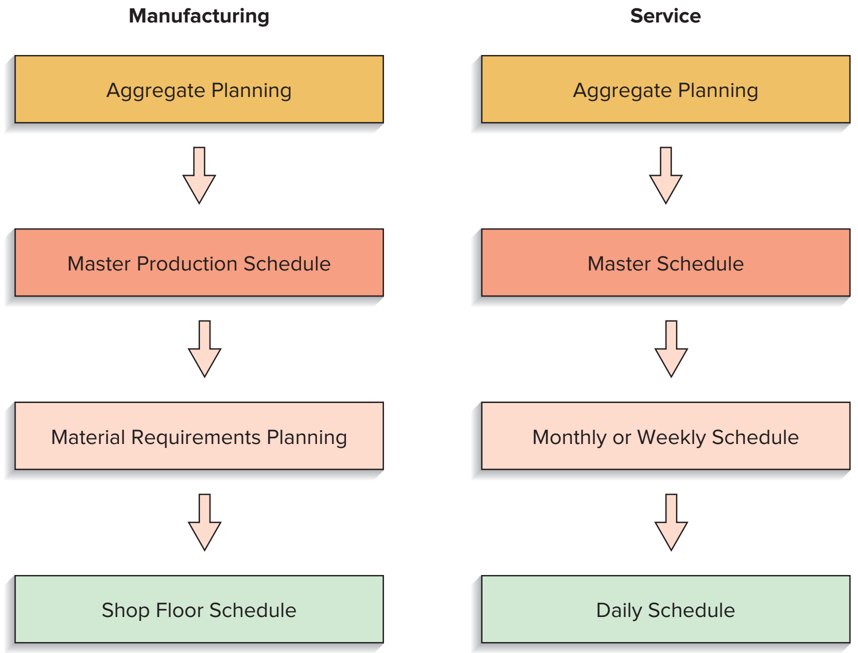
In the decision-making hierarchy, scheduling decisions are the final step in the transformation process before actual output occurs. Many decisions about system design and operation have been made long before scheduling decisions. They include the capacity of the system, product or service design, equipment selection, selection and training of workers, and aggregate planning and master scheduling. Consequently, scheduling decisions must be made within the constraints established by many other decisions, making them fairly narrow in scope and latitude. Figure 16.1 depicts scheduling hierarchies for manufacturing and service scheduling.

Effective scheduling can yield cost savings, increases in productivity, and other benefits. For example, in hospitals, effective scheduling can save lives and improve patient care. In educational institutions, it can reduce the need for expansion of facilities. In competitive environments, effective scheduling can give a company a competitive advantage in terms of customer service (shorter wait time for their orders) if its competitors are less effective with their scheduling.

Generally, the objectives of scheduling are to achieve trade-offs among conflicting goals, which include efficient utilization of staff, equipment, and facilities, and minimization of customer waiting time, inventories, and process times.

This chapter covers scheduling in both manufacturing and service environments. Although the two environments have many similarities, some basic differences are important.

FIGURE 16.1
Scheduling hierarchies



LO16.2 Compare product and service scheduling hierarchies.

16.1 SCHEDULING OPERATIONS

LO16.3 Describe scheduling needs in high-volume systems.

Scheduling tasks are largely a function of the volume of system output. High-volume systems require approaches substantially different from those required by job shops, and project scheduling requires still different approaches. In this chapter, we will consider scheduling for high-volume systems, intermediate-volume systems, and low-volume (job shop) scheduling. Project scheduling is discussed in Chapter 17.

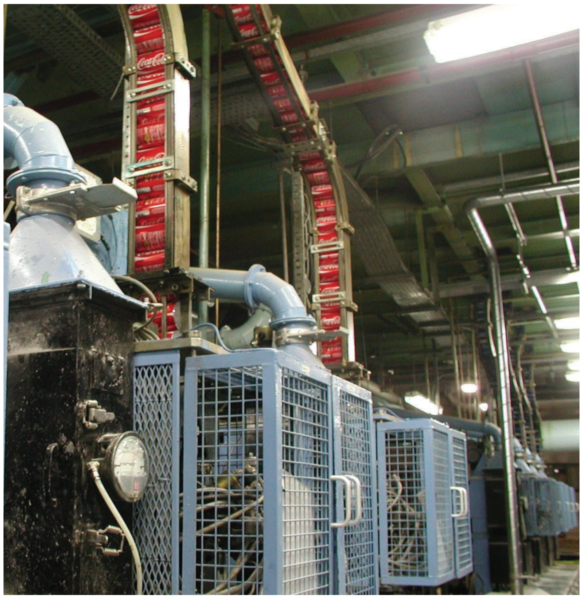
Scheduling in High-Volume Systems

Scheduling encompasses allocating workloads to specific work centers and determining the sequence in which operations are to be performed. High-volume systems are characterized by standardized equipment and activities that provide identical or highly similar operations on customers or products as they pass through the system. The goal is to obtain a smooth rate of flow of goods or customers through the system in order to get a high utilization of labor and equipment. High-volume systems, where jobs follow the same sequence, are often referred to as **flow systems**; scheduling in these systems is referred to as **flow-shop scheduling**, although flow-shop scheduling also can be used in medium-volume systems. Examples of high-volume products include autos, smartphones, radios and televisions, office supplies, toys, and appliances. In process industries, examples include petroleum refining, sugar refining, mining, waste treatment, and the manufacturing of fertilizers. Examples of services include cafeteria lines, news broadcasts, and mass inoculations. Because of the highly repetitive nature of these systems, many of the loading and sequence decisions are determined during the design of the system. The use of highly specialized tools and equipment, arrangement of equipment, use of specialized material-handling equipment, and division of labor are all designed to enhance the flow of work through the system, since all items follow virtually the same sequence of operations.

Flow system High–volume system in which jobs all follow the same sequence.

Flow-shop scheduling Scheduling for flow systems.

A major aspect in the design of flow systems is *line balancing,* which concerns allocating the required tasks to workstations so that they satisfy technical (sequencing) constraints and are balanced with respect to equal work times among stations. Highly balanced systems result in the maximum utilization of equipment and personnel as well as the highest possible rate of output. Line balancing is discussed in Chapter 6.



Courtesy Rexam PLC.

At the Wakefield, U.K. factory, Rexam produces 5,000 cans per minute and delivers them to its main customer, Coca-Cola Bottling. To meet Coke's lean manufacturing requirements, Rexam needed to spray cans the same every single time. They added a spray monitor system to immediately identify spray malfunctions. The early diagnostics of the spray monitor system can save Rexam from coating hundreds of cans improperly. The objectives in upgrading were to improve quality, reduce variation, and reduce costs.

In setting up flow systems, designers must consider the potential discontent of workers in connection with the specialization of job tasks in these systems; high work rates are often achieved by dividing the work into a series of relatively simple tasks assigned to different workers. The resulting jobs tend to be boring and monotonous and may give rise to fatigue, absenteeism, turnover, and other problems, all of which tend to reduce productivity and disrupt the smooth flow of work. These problems and potential solutions are elaborated on in Chapter 7, which deals with the design of work systems.

In spite of the built-in attributes of flow systems related to scheduling, a number of scheduling problems remain. One stems from the fact that few flow systems are *completely* devoted to a single product or service; most must handle a variety of sizes and models. Thus, an automobile manufacturer will assemble many different combinations of cars—two-door and four-door models, some with air-conditioning and some not, some with deluxe trim and others with standard trim, some with CD players, some with tinted glass, and so on. The same can be said for producers of appliances, electronic equipment, and toys. Each change involves slightly different inputs of parts, materials, and processing requirements that must be scheduled into the line. If the line is to operate smoothly, a supervisor must coordinate the flow of materials and the work, which includes the inputs, processing, and outputs, as well as purchases. In addition to achieving a smooth flow, it is important to avoid excessive buildup of inventories. Again, each variation in size or model will tend to have somewhat different inventory requirements, so that additional scheduling efforts will be needed.

One source of scheduling concern is possible disruptions in the system that result in less than the desired output. These can be caused by equipment failures, material shortages, accidents, and absences. In practice, it is usually impossible to increase the rate of output to compensate for these factors, mainly because flow systems are designed to operate at a given rate. Instead, strategies involving subcontracting or overtime are often required, although subcontracting on short notice is not always feasible. Sometimes work that is partly completed can be made up off the line.

The reverse situation can also impose scheduling problems although these are less severe. This happens when the desired output is less than the usual rate. However, instead of slowing the ensuing rate of output, it is usually necessary to operate the system at the usual rate, but for fewer hours. For instance, a production line might operate temporarily for seven hours a day instead of eight.

High-volume systems usually require automated or specialized equipment for processing and handling. Moreover, they perform best with a high, uniform output. Shutdowns and start-ups are generally costly, and especially costly in process industries. Consequently, the following factors often determine the success of such a system:

- **Process and product design.** Here, cost and manufacturability are important, as is achieving a smooth flow through the system.

- **Preventive maintenance.** Keeping equipment in good operating order can minimize breakdowns that would disrupt the flow of work.
- **Rapid repair when breakdowns occur.** This can require specialists as well as stocks of critical spare parts.
- **Optimal product mixes.** Techniques such as linear programming can be used to determine optimal blends of inputs to achieve desired outputs at minimal costs. This is particularly true in the manufacture of fertilizers, animal feeds, and diet foods.
- **Minimization of quality problems.** Quality problems can be extremely disruptive, requiring shutdowns while problems are resolved. Moreover, when output fails to meet quality standards, not only is there the loss of output but also a waste of the labor, material, time, and other resources that went into it.
- **Reliability and timing of supplies.** Shortages of supplies are an obvious source of disruption and must be avoided. On the other hand, if the solution is to stockpile supplies, that can lead to high carrying costs. Shortening supply lead times, developing reliable supply schedules, and carefully projecting needs are all useful.

LO16.4 Describe scheduling needs in intermediate-volume systems.

Scheduling in Intermediate-Volume Systems

Intermediate-volume system outputs fall between the standardized type of output of the high-volume systems and made-to-order output of job shops. Like the high-volume systems, intermediate-volume systems typically produce standard outputs. If manufacturing is involved, the products may be for stock rather than for special order. However, the volume of output in such cases is not large enough to justify continuous production. Instead, it is more economical to process these items *intermittently*. Thus, intermediate-volume work centers periodically shift from one job to another. In contrast to a job shop, the run (batch) sizes are relatively large. Examples of products made in these systems include canned foods, baked goods, paint, and cosmetics.

The three basic issues in these systems are the *run size* of jobs, the *timing* of jobs, and the *sequence* in which jobs should be processed.

Sometimes, the issue of run size can be determined by using a model such as the economic run size model discussed in Chapter 13 on inventory management. The run size that would minimize setup and inventory costs is

$$Q_p = \sqrt{\frac{2DS}{H}}\sqrt{\frac{p}{p-u}}, \text{ where } S = \text{Setup cost} \qquad (16\text{-}1)$$

Setup cost may be an important consideration. Setup costs may depend on the order in which jobs are processed; similar jobs may require less setup change between them. For example, jobs in a print shop may be sequenced by ink color to reduce the number of setups needed. This opens up the possibility of reducing setup cost and time by taking processing sequence into account. It also makes sequencing more complex, and it requires estimating job setup costs for every sequence combination.

In another vein, companies are working to reduce setup times and, hence, experience less downtime for equipment changeover. Tactics include offline setups, snap-on parts, modular setups, and flexible equipment designed to handle a variety of processing requirements.

Another difficulty arises because usage is not always as smooth as assumed in the model. Some products will tend to be used up faster than expected and have to be replenished sooner. Also, because multiple products are to be processed, it is not always possible to schedule production to correspond with optimum run times.

Another approach frequently used is to base production on a master schedule developed from customer orders and forecasts of demand. Companies engaged in assembly operations would then use an MRP approach (described in Chapter 12) to determine the quantity and projected timing of jobs for components. The manager would then compare projected requirements with projected capacity and develop a feasible schedule from that information. Companies engaged in producing processed rather than assembled goods (e.g., food products, such as canned goods and beverages; publishing, such as magazines; paints and cleaning supplies) would use a somewhat different approach; the *time-phasing* information provided by MRP would not be an important factor.

16.2 SCHEDULING IN LOW-VOLUME SYSTEMS

LO16.5 Describe scheduling needs in job shops.

The characteristics of low-volume systems (job shops) are considerably different from those of high- and intermediate-volume systems. Products are made to order, and orders usually differ considerably in terms of processing requirements, materials needed, processing time, and processing sequence and setups. Because of these circumstances, **job-shop scheduling** is usually fairly complex. This is compounded by the impossibility of establishing firm schedules prior to receiving the actual job orders.

Job-shop scheduling Scheduling for low–volume systems with many variations in requirements.

Job-shop processing gives rise to two basic issues for schedulers: how to distribute the workload among work centers and what job processing sequence to use.

Loading

Loading refers to the assignment of jobs to processing (work) centers. Loading decisions involve assigning specific jobs to work centers and to various machines in the work centers. In cases where a job can be processed only by a specific center, loading presents little difficulty. However, problems arise when two or more jobs are to be processed and there are a number of work centers capable of performing the required work. In such cases, the operations manager needs some way of assigning jobs to the centers.

Loading The assignment of jobs to processing centers.

When making assignments, managers often seek an arrangement that will minimize processing and setup costs, minimize idle time among work centers, or minimize job completion time, depending on the situation.

Gantt Charts. Visual aids called **Gantt charts** are used for a variety of purposes related to loading and scheduling. They derive their name from Henry Gantt, who pioneered the use of charts for industrial scheduling in the early 1900s. Gantt charts can be used in a number of different ways, two of which are illustrated in Figure 16.2, which shows scheduling classrooms for a university and scheduling hospital operating rooms for a day.

Gantt charts Chart used as visual aid for loading and scheduling purposes.

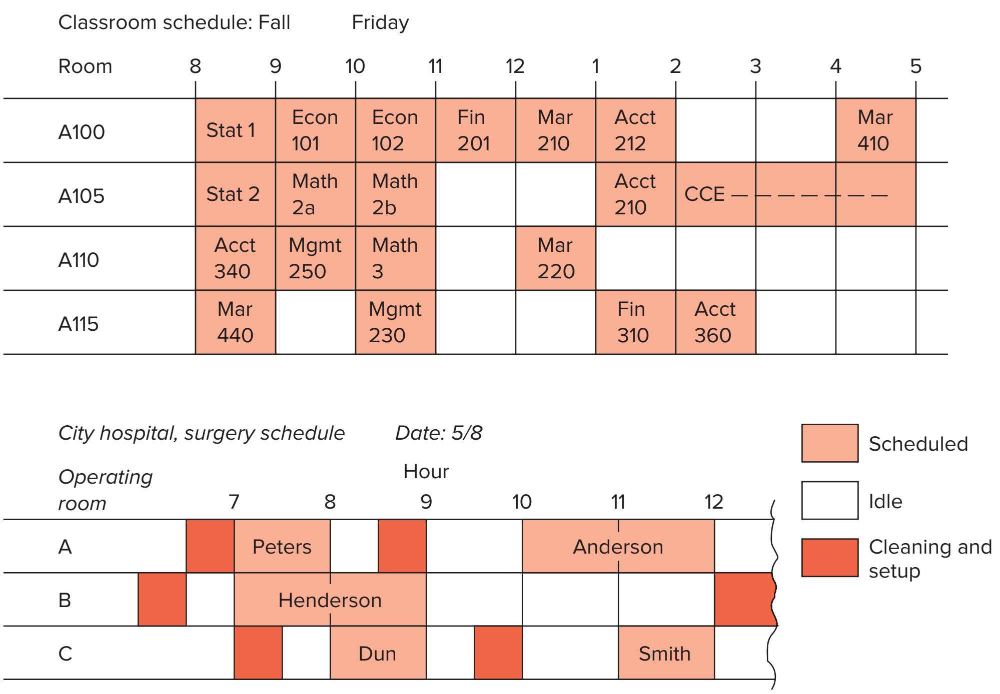


FIGURE 16.2
Examples of charts used for scheduling

LO16.6 Use and interpret Gantt charts..

The purpose of Gantt charts is to organize and visually display the actual or intended use of resources in a *time framework*. In most cases, a time scale is represented horizontally, and resources to be scheduled are listed vertically. The use and idle times of resources are reflected in the chart.

Managers may use the charts for trial-and-error schedule development to get an idea of what different arrangements would involve. Thus, a tentative surgery schedule might reveal insufficient allowance for surgery that takes longer than expected and can be revised accordingly. Use of the chart for classroom scheduling would help avoid assigning two different classes to the same room at the same time.

There are a number of different types of Gantt charts. Two of the most commonly used are the *load chart* and the *schedule chart.*

Load chart A Gantt chart that shows the loading and idle times for a group of machines or list of departments.

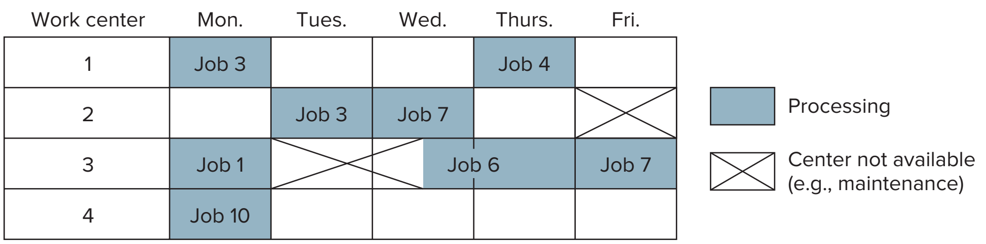
A **load chart** depicts the loading and idle times for a group of machines or a list of departments. Figure 16.3 illustrates a typical load chart. This chart indicates that work center 3 is completely loaded for the entire week, center 4 will be available from Tuesday to Friday, and the other two centers have idle time scattered throughout the week. This information can help a manager rework loading assignments to better utilize the centers. For instance, if all centers perform the same kind of work, the manager might want to free one center for a long job or a rush order. The chart also shows when certain jobs are scheduled to start and finish, and where to expect idle time.

FIGURE 16.3

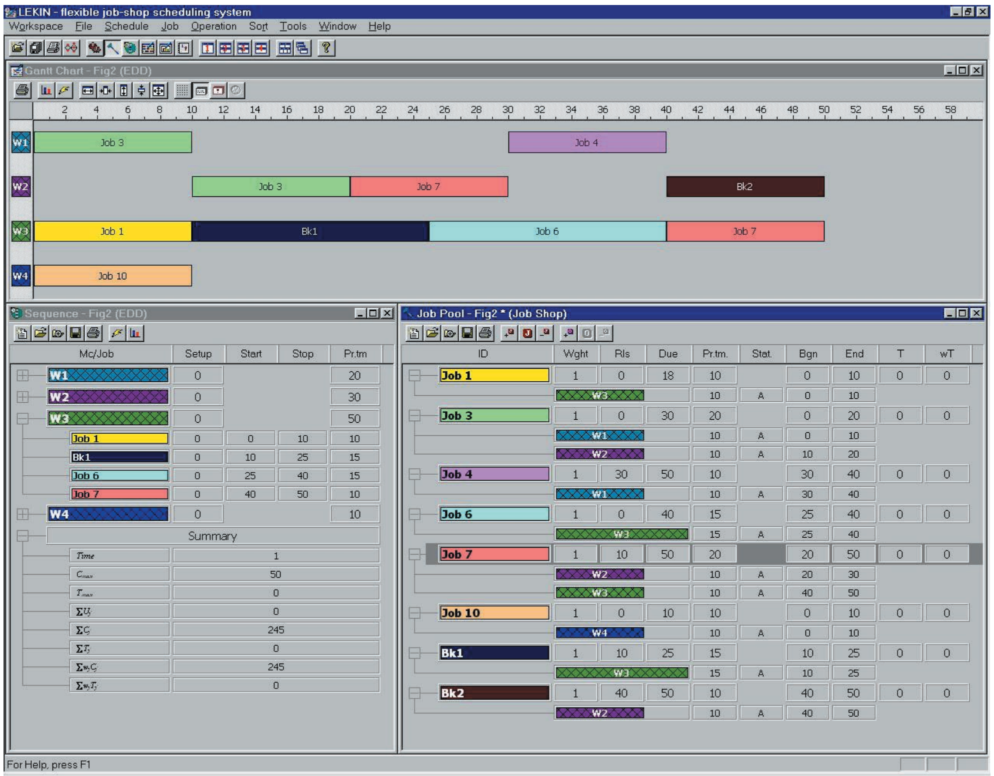
A. Sample Gantt load chart. B. The same Gantt chart using the Lekin software, developed at New York University, includes multiple scheduling routines along with graphics.

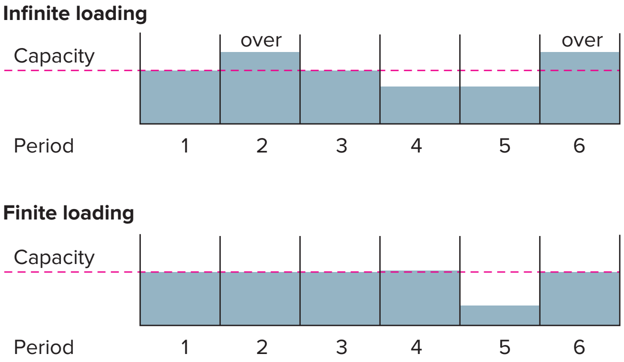
Source: © Pinedo and Feldman. Used with permission.

A.



B.





Two different approaches are used to load work centers: *infinite* loading and *finite* loading. **Infinite loading** assigns jobs to work centers without regard to the capacity of the work center. As you can see in the diagram that follows, this can lead to overloads in some time periods and underloads in others. The priority sequencing rules described in this chapter use infinite loading. One possible result of infinite loading is the formation of queues in some (or all) work centers. That requires a second step to correct the imbalance. **Finite loading** projects actual job starting and stopping times at each work center, taking into account the capacities of each work center and the processing times of jobs, so that capacity is not exceeded. One output of finite loading is a detailed projection of hours each work center will operate. Schedules based on finite loading may have to be updated often, perhaps daily, due to processing delays at work centers and the addition of new jobs or cancellation of current jobs. The following diagram illustrates these two approaches.

Infinite loading Jobs are assigned to work centers without regard to the capacity of the work center.

Finite loading Jobs are assigned to work centers taking into account the work center capacity and job processing times.

With infinite loading, a manager may need to make some response to overloaded work centers. Among the possible responses are shifting work to other periods or other centers, working overtime, or contracting out a portion of the work. Note that the last two options in effect increase capacity to meet the workload.

Finite loading may reflect a fixed upper limit on capacity. For example, a bus line will have only so many buses. Hence, the decision to place into service a particular number of buses fixes capacity. Similarly, a manufacturer might have one specialized machine that it operates around the clock. Thus, it is operated at the upper limit of its capacity, so finite loading would be called for.

There are two general approaches to scheduling—forward scheduling and backward scheduling. **Forward scheduling** means scheduling ahead from a point in time; **backward scheduling** means scheduling backward from a due date. Forward scheduling is used if the issue is "How long will it take to complete this job?" Backward scheduling would be used if the issue is "When is the latest the job can be started and still be completed by the due date?" Forward scheduling enables the scheduler to determine the earliest possible completion time for each job and, thus, the amount of lateness or the amount of slack can be determined. That information can be combined with information from other jobs in setting up a schedule for all current jobs.

Forward scheduling Scheduling ahead from a point in time.

Backward scheduling Scheduling backward from a due date.

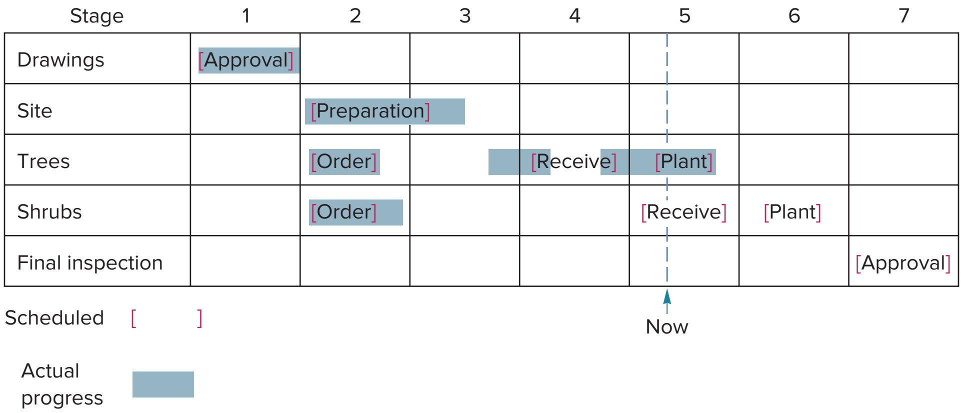
A manager often uses a **schedule chart** to monitor the progress of jobs. The vertical axis on this type of Gantt chart shows the orders or jobs in progress, and the horizontal axis shows time. The chart indicates which jobs are on schedule and which are behind or ahead.

Schedule chart A Gantt chart that shows the orders or jobs in progress and whether they are on schedule.

A typical schedule chart is illustrated in Figure 16.4. It shows the current status of a landscaping job with planned and actual starting and finishing times for the five stages of the job. The chart indicates that approval and the ordering of trees and shrubs was on schedule. The site preparation was a bit behind schedule. The trees were received earlier than expected, and planting is ahead of schedule. However, the shrubs have not yet been received. The chart indicates some slack between scheduled receipt of shrubs and shrub planting, so if the shrubs arrive by the end of the week, it appears the schedule can still be met.

Despite the obvious benefits of Gantt charts and the fact that they are widely used, they possess certain limitations, the chief one being the need to repeatedly update a chart to keep it

FIGURE 16.4
Progress chart for landscaping job



current. In addition, a chart does not directly reveal costs associated with alternative loadings. Finally, a job's processing time may vary depending on the work center; certain stations or work centers may be capable of processing some jobs faster than other stations. Again, that situation would increase the complexity of evaluating alternative schedules.

In addition to Gantt charts, managers often rely on input/output reports to manage work flow.

Input/output (I/O) control Managing work flow and queues at work centers.

Input/Output Control. **Input/output (I/O) control** refers to monitoring the work flow and queue lengths at work centers. The purpose of I/O control is to manage work flow so that queues and waiting times are kept under control. Without I/O control, demand may exceed processing capacity, causing an overload at a center. Conversely, work may arrive slower than the rate a work center can handle, leaving the work center underutilized. Ideally, a balance can be struck between the input and output rates, thereby achieving effective use of work center capacities without experiencing excessive queues at the work centers. A simple example of I/O control is the use of stoplights on some expressway on-ramps. These regulate the flow of entering traffic according to the current volume of expressway traffic.

Figure 16.5 illustrates an input/output report for a work center. A key portion of the report is the backlog of work waiting to be processed. The report also reveals deviations-from-planned for both inputs and outputs, thereby enabling a manager to determine possible sources of problems.

The deviations in each period are determined by subtracting "planned" from "actual." For example, in the first period, subtracting the planned input of 100 hours from the actual input

FIGURE 16.5
A sample input/output report for a work center showing input and output in hours of processing time

Note: Figures represent standard hours of processing time.

*Given, not derived from the data.

			Period					
			1	2	3	4	5	6
Input	Planned		100	100	90	90	90	90
	Actual		120	95	80	88	93	94
	Deviation		+20	–5	–10	–2	+3	+4
	Cum. dev.		+20	+15	+5	+3	+6	+10
Output	Planned		110	110	100	100	100	95
	Actual		110	105	95	101	103	96
	Deviation		0	–5	–5	+1	+3	+1
	Cum. dev.		0	+5	+10	+9	+6	+5
Backlog		40*	50	40	25	12	2	0

of 120 hours produces a deviation of +20 hours. Similarly, in the first period, the planned and actual outputs are equal, producing a deviation of 0 hours.

The backlog for each period is determined by subtracting the "actual output" from the "actual input" and adjusting the backlog from the previous period by that amount. For example, in the second period actual output exceeds actual input by 10 hours. Hence, the previous backlog of 50 hours is reduced by 10 hours to 40 hours.

Another approach that can be used to assign jobs to resources is the *assignment method.*

Assignment Method of Linear Programming. The **assignment model** is a special-purpose linear programming model that is useful in situations that call for assigning tasks or other work requirements to resources. Typical examples include assigning jobs to machines or workers, territories to salespeople, and repair jobs to repair crews. The idea is to obtain an optimum *matching* of tasks and resources. Commonly used criteria include costs, profits, efficiency, and performance.

Assignment model A linear programming model for optimal assignment of tasks and resources.

LO16.7 Use the assignment method for loading.

SCREENCAM TUTORIAL

Table 16.1 illustrates a typical problem, where four jobs are to be assigned to four workers. The problem is arranged in a format that facilitates evaluation of assignments. The numbers in the body of the table represent the value or cost associated with each job-worker combination. In this case, the numbers represent costs. Thus, it would cost $8 for worker A to do job 1, $6 for worker B to do job 1, and so on. If the problem involved minimizing the cost for job 1 alone, it would clearly be assigned to worker C, since that combination has the lowest cost. However, that assignment does not take into account the other jobs and their costs, which is important since the lowest-cost assignment for any one job may not be consistent with a minimum-cost assignment when all jobs are considered.

If there are to be *n* matches, there are *n!* different possibilities. In this case, there are 4! = 24 different matches. One approach is to investigate each match and select the one with the lowest cost. However, if there are 12 jobs, there would be 479 million different matches! A much simpler approach is to use a procedure called the **Hungarian method** to identify the lowest-cost solution.

Hungarian method Method of assigning jobs by a one-for-one matching to identify the lowest–cost solution.

To be able to use the Hungarian method, a one-for-one matching is required. Each job, for example, must be assigned to only one worker. It is also assumed that every worker is capable of handling every job, and that the costs or values associated with each assignment combination are known and fixed (i.e., not subject to variation). The number of rows and columns must be the same. Solved Problem 1 at the end of the chapter shows what to do if they aren't the same.

Once the relevant cost information has been acquired and arranged in tabular form, the basic procedure of the Hungarian method is as follows:

1. Subtract the smallest number in each row from every number in the row. This is called a *row reduction.* Enter the results in a new table.
2. Subtract the smallest number in each column of the new table from every number in the column. This is called a *column reduction.* Enter the results in another table.
3. Test whether an optimum assignment can be made. You do this by determining the *minimum* number of lines (horizontal or vertical) needed to cross out (cover) all zeros. If the number of lines equals the number of rows, an optimum assignment is possible. In that case, go to step 6. Otherwise go on to step 4.
4. If the number of lines is less than the number of rows, modify the table in this way:
   a. Subtract the smallest uncovered number from every uncovered number in the table.
   b. Add the smallest uncovered number to the numbers at *intersections* of cross-out lines.
   c. Numbers crossed out but not at intersections of cross-out lines carry over to the next table.
5. Repeat steps 3 and 4 until an optimal table is obtained.
6. Make the assignments. Begin with rows or columns with only one zero. Match items that have zeros, using only one match for each row and each column. Eliminate both the row and the column after the match.

TABLE 16.1
A typical assignment problem showing job times for each job/ worker combination

		WORKER			
		A	B	C	D
	1	8	6	2	4
Job	2	6	7	11	10
	3	3	5	7	6
	4	5	10	12	9

EXAMPLE 1

mhhe.com/stevenson13e

Using the Assignment Method to Make Job Assignments

Determine the optimum assignment of jobs to workers for the following data (from Table 16.1):

		WORKER				
		A	B	C	D	Row Minimum
	1	8	6	2	4	2
Job	2	6	7	11	10	6
	3	3	5	7	6	3
	4	5	10	12	9	5

SOLUTION

a. Subtract the smallest number in each row from every number in the row, and enter the results in a new table. The result of this row reduction is:

		WORKER			
		A	B	C	D
	1	6	4	0	2
Job	2	0	1	5	4
	3	0	2	4	3
	4	0	5	7	4
Column Minimum		0	1	0	2

b. Subtract the smallest number in each column from every number in the column, and enter the results in a new table. The result of this column reduction is:

		WORKER			
		A	B	C	D
	1	6	3	0	0
Job	2	0	0	5	2
	3	0	1	4	1
	4	0	4	7	2

c. Determine the *minimum* number of lines needed to cross out all zeros. (Try to cross out as many zeros as possible when drawing lines.)

		WORKER			
		A	B	C	D
	1	6	3	0	0
Job	2	0	0	5	2
	3	0	1	4	1
	4	0	4	7	2

d. Since only three lines are needed to cross out all zeros and the table has four rows, this is not the optimum. Note that the smallest uncovered value is 1.

e. Subtract the smallest uncovered value from every uncovered number that hasn't been crossed out, and add it to numbers that are at the intersections of covering lines. The results are as follows:

		WORKER			
		A	B	C	D
	1	7	3	0	0
job	2	1	0	5	2
	3	0	0	3	0
	4	0	3	6	1

f. Determine the minimum number of lines needed to cross out all zeros (four). Since this equals the number of rows, you can make the optimum assignment.

		WORKER			
		A	**B**	**C**	**D**
	1	7	3	0	0
Job	**2**	1	0	5	2
	3	0	0	3	0
	4	0	3	6	1

g. Make assignments: Start with rows and columns with only one zero. Match jobs with machines that have a zero cost.

		WORKER			
		A	**B**	**C**	**D**
	1	7	3	0	0
Job	**2**	1	0	5	2
	3	0	0	3	0
	4	0	3	6	1

Assignment	**Cost**
1-C	$ 2
2-B	7
3-D	6
4-A	5
	20

The assignment problem can also be solved using an Excel template, as seen in Table 16.2. The ones in the solution matrix denote assignments (i.e., assign C to job 1), and the zeros denote no assignment for a worker/machine and job combination.

TABLE 16.2 Excel solution to Example 1

Input Matrix: Worker/Machine

Job	A	B	C	D	E	F	G	H	
1	8	6	2	4					1
2	6	7	11	10					1
3	3	5	7	6					1
4	5	10	12	9					1
5									0
6									0
7									0
8									0
	1	1	1	1	0	0	0	0	

Total Jobs = 4

Do not change or delete unshaded cells. Total Workers/Machines = 4

Solution Matrix: Worker/Machine

Job	A	B	C	D	E	F	G	H	
1	0	0	1	0	0	0	0	0	1
2	0	1	0	0	0	0	0	0	1
3	0	0	0	1	0	0	0	0	1
4	1	0	0	0	0	0	0	0	1
5	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0
	1	1	1	1	0	0	0	0	

Total Cost = 20

As you can see, the process is relatively simple. The simplicity of the Hungarian method belies its usefulness when the assumptions are met. Not only does it provide a rational method for making assignments, it guarantees an optimal solution, often without the use of a computer, which is necessary only for fairly large problems. When profits instead of costs are involved, the profits can be converted to *relative costs* by subtracting every number in the table from the largest number and then proceeding as in a minimization problem.

It is worth knowing that one extension of this technique can be used to prevent undesirable assignments. For example, union rules may prohibit one person's assignment to a particular job, or a manager might wish to avoid assigning an unqualified person to a job. Whatever the reason, specific combinations can be avoided by assigning a relatively high cost to that combination. In the previous example, if we wish to avoid combination 1-A, assigning a cost of $50 to that combination will achieve the desired effect, because $50 is much greater than the other costs.

Sequencing

Although loading decisions determine the machines or work centers that will be used to process specific jobs, they do not indicate the *order* in which the jobs waiting at a given work center are to be processed. **Sequencing** is concerned with determining job processing order. Sequencing decisions determine both the order in which jobs are processed at various work centers and the order in which jobs are processed at individual **workstations** *within* the work centers.

Sequencing Determining the order in which jobs at a work center will be processed.

Workstation An area where one or a few workers and/or machines perform similar work.

If work centers are lightly loaded and if jobs all require the same amount of processing time, sequencing presents no particular difficulties. However, for heavily loaded work centers, especially in situations where relatively lengthy jobs are involved, the order of processing can be very important in terms of costs associated with jobs waiting for processing and in terms of idle time at the work centers. In this section, we will examine some of the ways in which jobs are sequenced.

Typically, a number of jobs will be waiting for processing. **Priority rules** are simple heuristics used to select the order in which the jobs will be processed. Some of the most common are listed in Table 16.3. The rules generally rest on the assumption that job setup cost and time are *independent* of processing sequence. In using these rules, job processing times and due dates are important pieces of information. **Job time** usually includes setup and processing times. Jobs that require similar setups can lead to reduced setup times if the sequencing rule takes this into account (the rules described here do not). Due dates may be the result of delivery times promised to customers, material requirements planning (MRP) processing, or managerial decisions. They are subject to revision and must be kept current to give meaning to sequencing choices. Also, it should be noted that due dates associated with all rules except slack per operation (S/O) and critical ratio (CR) are for the operation about to be performed; due dates for S/O and CR are typically final due dates for orders rather than intermediate, departmental deadlines.

Priority rules Simple heuristics used to select the order in which jobs will be processed.

Job time Time needed for setup and processing of a job.

The priority rules can be classified as either *local* or *global*. **Local priority rules** take into account information pertaining only to a single workstation; **global priority rules** take into account information pertaining to multiple workstations. First come, first served (FCFS), shortest processing time (SPT), and earliest due date (EDD) are local rules; CR and S/O are global rules. Rush can be either local or global. As you might imagine, global rules require more effort than local rules. A major complication in global sequencing is that not all jobs

Local priority rules Focus on information pertaining to a single workstation when establishing a job sequence.

Global priority rules Incorporate information from multiple workstations when establishing a job sequence.

TABLE 16.3
Possible priority rules

First come, first served (FCFS): Jobs are processed in the order in which they arrive at a machine or work center.
Shortest processing time (SPT): Jobs are processed according to processing time at a machine or work center, shortest job first.
Earliest due date (EDD): Jobs are processed according to due date, earliest due date first.
Critical ratio (CR): Jobs are processed according to smallest ratio of time remaining until due date to processing time remaining.
Slack per operation (S/O): Jobs are processed according to average slack time (time until due date minus remaining time to process). Compute by dividing slack time by number of remaining operations, including the current one.
Rush: Emergency or preferred customers first.

LO16.8 Give examples of commonly used priority rules.

TABLE 16.4
Assumptions of priority rules

The set of jobs is known; no new jobs arrive after processing begins; and no jobs are canceled.
Setup time is independent of processing sequence.
Setup time is deterministic.
Processing times are deterministic rather than variable.
There will be no interruptions in processing such as machine breakdowns, accidents, or worker illness.

require the same processing or even the same order of processing. As a result, the set of jobs is different for different workstations. Local rules are particularly useful for bottleneck operations, but they are not limited to those situations.

A number of assumptions apply when using the priority rules; Table 16.4 lists them. In effect, the priority rules pertain to *static* sequencing: For simplicity, it is assumed that there is no variability in either setup or processing times, or in the set of jobs. The assumptions make the scheduling problem manageable. In practice, jobs may be delayed or canceled, and new jobs may arrive, requiring schedule revisions.

The effectiveness of any given sequence is frequently judged in terms of one or more *performance measures.* The most frequently used performance measures follow:

- **Job flow time** is the amount of time it takes from when a job arrives until it is complete. It includes not only actual processing time but also any time waiting to be processed, transportation time between operations, and any waiting time related to equipment breakdowns, unavailable parts, quality problems, and so on. The average flow time for a group of jobs is equal to the total flow time for the jobs divided by the number of jobs.

Job flow time The amount of time from when a job arrives until it is finished.

- **Job lateness** is the amount of time the job completion date is expected to exceed the date the job was due or promised to a customer. It is the difference between the actual completion time and the due date. If only differences for jobs with completion times that exceed due dates are recorded, and zeros are assigned to jobs that are early, the term used is job *tardiness.*

Job lateness The difference between the actual completion date and the due date.

- **Makespan** is the total time needed to complete a *group* of jobs. It is the length of time between the start of the first job in the group and the completion of the last job in the group. If processing involves only one work center, makespan will be the same regardless of the priority rule being used.

Makespan Total time needed to complete a group of jobs from the beginning of the first job to the completion of the last job.

- **Average number of jobs.** Jobs that are in a shop are considered to be work-in-process inventory. The average work-in-process for a group of jobs can be computed using the following formula:

  $$\text{Average number of jobs} = \text{Total flow time} \div \text{Makespan}$$

  If the jobs represent equal amounts of inventory, the average number of jobs will also reflect the average work-in-process inventory.

Of the priority rules, rush scheduling is quite simple and needs no explanation. The other rules and performance measures are illustrated in the following two examples.

EXAMPLE 2

excel
mhhe.com/stevenson13e

Determining Job Sequences Using Various Rules

Processing times (including setup times) and due dates for six jobs waiting to be processed at a work center are given in the following table. Determine the sequence of jobs, the average flow time, average tardiness, and average number of jobs at the work center, for each of these rules:

a. FCFS

b. SPT

c. EDD

d. CR

Job	Processing Time (days)	Due Date (days from present time)
A	2	7
B	8	16
C	4	4
D	10	17
E	5	15
F	12	18

Assume jobs arrived in the order shown.

SOLUTION

a. The FCFS sequence is simply A-B-C-D-E-F. The measures of effectiveness are as follows (see table):

(1) *Average flow time:* 120/6 = 20 days.

(2) *Average tardiness:* 54/6 = 9 days.

(3) The *makespan* is 41 days. *Average number of jobs at the work center:* 120/41 = 2.93.

Job Sequence	(1) Processing Time	(2) Flow Time	(3) Due Date	(2) – (3) Days Tardy [0 if negative]
A	2	2	7	0
B	8	10	16	0
C	4	14	4	10
D	10	24	17	7
E	5	29	15	14
F	12	41	18	23
	41	120		54

The flow time column indicates *cumulative* processing time, so summing these times and dividing by the total number of jobs processed indicates the average time each job spends at the work center. Similarly, find the average number of jobs at the center by summing the flow times and dividing by the total processing time.

The Excel solution is shown in Table 16.5.

b. Using the SPT rule, the job sequence is A-C-E-B-D-F (see the following table). The resulting values for the three measures of effectiveness are:

(1) *Average flow time:* 108/6 = 18 days.

(2) *Average tardiness:* 40/6 = 6.67 days.

(3) *Average number of jobs at the work center:* 108/41 = 2.63.

Job Sequence	(1) Processing Time	(2) Flow Time	(3) Due Date	(2) – (3) Days Tardy [0 if negative]
A	2	2	7	0
C	4	6	4	2
E	5	11	15	0
B	8	19	16	3
D	10	29	17	12
F	12	41	18	23
	41	108		40

TABLE 16.5 Excel solution for Example 2a

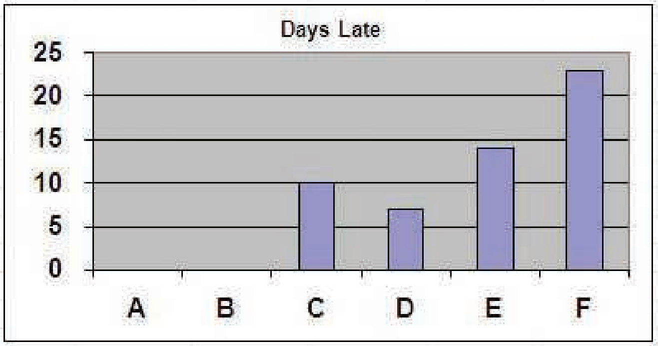
Job Sequencing
<Back Notes

Clear | FCFS | SPT | DD | CR | S/O

Current Date: 23

Job	Process. Time	Due Date	Remain Oper.	CR Sequence	Critical Ratio	Slack	S/O	Flow Time	Days Late
A	2	7		3		5		2	0
B	8	16		5		8		10	0
C	4	4		1		0		14	10
D	10	17		6		7		24	7
E	5	15		4		10		29	14
F	12	18		2		6		41	23
Totals	41		0			36		120	54

Method	FCFS
Average Flow Time	20.00
Average Tardiness	9.00
Average Number of Jobs	2.93



See notes below.

Notes:

1. Enter Job Name, Processing Time, and Due Date for each job.
2. For the FCFS, SPT, and DD rules, simply press the appropriate button.
3. For the CR rule, perform the following BEFORE pressing the CR button:
   a. Select job with lowest Critical Ratio
   b. Schedule that job next by entering next sequence number (start with 1) in the CR Sequence column.
   c. Add the processing time for that job to the current date.
   d. Repeat steps a, b, and c until all jobs have been scheduled (i.e. the CR Sequence column is filled in).
   e. Then press the CR button.
4. Fill in the Remaining Operations column and then press the S/O button.

c. Using earliest due date as the selection criterion, the job sequence is C-A-E-B-D-F. The measures of effectiveness are as follows (see table):

(1) *Average flow time:* 110/6 = 18.33 days.

(2) *Average tardiness:* 38/6 = 6.33 days.

(3) *Average number of jobs at the work center:* 110/41 = 2.68.

Job Sequence	(1) Processing Time	(2) Flow Time	(3) Due Date	(2) – (3) Days Tardy [0 if negative]
C	4	4	4	0
A	2	6	7	0
E	5	11	15	0
B	8	19	16	3
D	10	29	17	12
F	12	41	18	23
	41	110		38

d. Using the critical ratio we find:

Job Sequence	Processing Time	Due Date	Critical Ratio Calculation
A	2	7	(7 − 0)/2 = 3.5
B	8	16	(16 − 0)/8 = 2.0
C	4	4	(4 − 0)/4 = 1.0 (lowest)
D	10	17	(17 − 0)/10 = 1.7
E	5	15	(15 − 0)/5 = 3.0
F	12	18	(18 − 0)/12 = 1.5

At day 4 [C completed], the critical ratios are

Job Sequence	Processing Time	Due Date	Critical Ratio Calculation
A	2	7	(7 − 4)/2 = 1.5
B	8	16	(16 − 4)/8 = 1.5
C	—	—	—
D	10	17	(17 − 4)/10 = 1.3
E	5	15	(15 − 4)/5 = 2.2
F	12	18	(18 − 4)/12 = 1.17 (lowest)

At day 16 [C and F completed], the critical ratios are

Job Sequence	Processing Time	Due Date	Critical Ratio Calculation
A	2	7	(7 − 16)/2 = −4.5 (lowest)
B	8	16	(16 − 18)/8 = 0.0
C	—	—	—
D	10	17	(17 − 16)/10 = 0.1
E	5	15	(15 − 16)/5 = −0.2
F	—	—	—

At day 18 [C, F, and A completed], the critical ratios are:

Job Sequence	Processing Time	Due Date	Critical Ratio Calculation
A	—	—	—
B	8	16	(16 − 18)/8 = −0.25
C	—	—	—
D	10	17	(17 − 18)/10 = −0.10
E	5	15	(15 − 18)/5 = −0.60 (lowest)
F	—	—	—

At day 23 [C, F, A, and E completed], the critical ratios are

Job Sequence	Processing Time	Due Date	Critical Ratio Calculation
A	—	—	—
B	8	16	(16 − 23)/8 = −0.875 (lowest)
C	—	—	—
D	10	17	(17 − 23)/10 = −0.60
E	—	—	—
F	—	—	—

The job sequence is C-F-A-E-B-D, and the resulting values for the measures of effectiveness are as follows:

(1) *Average flow time:* 133/6 = 22.17 days.

(2) *Average tardiness:* 58/6 = 9.67 days.

(3) *Average number of jobs at the work center:* 133/41 = 3.24.

Sequence	(1) Processing Time	(2) Flow Time	(3) Due Date	(2) − (3) Days Tardy
C	4	4	4	0
F	12	16	18	0
A	2	18	7	11
E	5	23	15	8
B	8	31	16	15
D	10	41	17	24
	41	133		58

The results of these four rules are summarized in Table 16.6.

In Example 2, the SPT rule was the best according to two of the measures of effectiveness and a little worse than the EDD rule on average tardiness. The CR rule was the worst in every case. For a different set of numbers, the EDD rule (or perhaps another rule not mentioned here) might prove superior to SPT in terms of average job tardiness or some other measure of effectiveness. However, SPT is always superior in terms of minimizing flow time and, hence, in terms of minimizing the average number of jobs at the work center and completion time. This results in faster job completion, which has the potential to generate more revenue.

Generally speaking, the FCFS rule and the CR rule turn out to be the least effective of the rules.

The primary limitation of the FCFS rule is that long jobs will tend to delay other jobs. If a process consists of work on a number of machines, machine idle time for downstream workstations will increase. However, for service systems in which customers are directly involved, the FCFS rule is by far the dominant priority rule, mainly because of the inherent fairness but also because of the inability to obtain realistic estimates of processing time for individual jobs. The FCFS rule also has the advantage of simplicity. If other measures are important when there is high customer contact, companies may adopt the strategy of moving processing to the "backroom" so they don't necessarily have to follow FCFS.

Because the SPT rule always results in the lowest (i.e., optimal) average completion (flow) time, it can result in lower in-process inventories. And because it often provides the lowest (optimal) average tardiness, it can result in better customer service levels. Finally, since it always involves a lower average number of jobs at the work center, there tends to be less congestion in the work area. SPT also minimizes downstream idle time. However, due dates are often uppermost in managers' minds, so they may not use SPT because it doesn't incorporate due dates.

The major disadvantage of the SPT rule is that it tends to make long jobs wait, perhaps for rather long times (especially if new, shorter jobs are continually added to the system). That can be troubling if long jobs are from the company's best customers. Various modifications may be used in an effort to avoid this. For example, after waiting for a given time period, any remaining jobs are automatically moved to the head of the line. This is known as the *truncated* SPT rule.

TABLE 16.6
Comparison of the four rules for Example 2

Rule	Average Flow Time (days)	Average Tardiness (days)	Average Number of Jobs at the Work Center
FCFS	20.00	9.00	2.93
SPT	18.00	6.67	2.63
EDD	18.33	6.33	2.68
CR	22.17	9.67	3.24

The EDD rule directly addresses due dates and minimizes lateness. Although it has intuitive appeal, its main limitation is that it does not take processing time into account. One possible consequence is that it can result in some jobs waiting a long time, which adds to both in-process inventories and shop congestion.

The CR rule is easy to use and has intuitive appeal. Although it had the poorest showing in Example 2 for all three measures, it usually does quite well in terms of minimizing job tardiness. Therefore, if job tardiness is important, the CR rule might be the best choice among the rules.

Let's take a look now at the S/O (slack per operation) rule.

EXAMPLE 3

eXcel
mhhe.com/stevenson13e

Scheduling Jobs Using the S/O Rule

Use the S/O rule to schedule the following jobs. Note that processing time includes the time remaining for the current and subsequent operations. In addition, you will need to know the number of operations remaining, including the current one.

Job	Remaining Processing Time	Due Date	Remaining Number of Operations
A	4	14	3
B	16	32	6
C	8	8	5
D	20	34	2
E	10	30	4
F	18	30	2

SOLUTION

Determine the difference between the due date and the processing time for each operation. Divide the difference by the number of remaining operations, and rank them from low to high. This yields the sequence of jobs:

Job	(1) Remaining Processing Time	(2) Due Date	(3) (2) − (1) Slack	(4) Remaining Number of Operations	(5) (3) ÷ (4) Ratio	(6) Rank
A	4	14	10	3	3.33	3
B	16	32	16	6	2.67	2
C	8	8	0	5	0	1
D	20	34	14	2	7.00	6
E	10	30	20	4	5.00	4
F	18	30	12	2	6.00	5

The indicated sequence (see column 6) is C-B-A-E-F-D.

Using the S/O rule, the designated job sequence may change after any given operation, so if that happened, it would be necessary to reevaluate the sequence after each operation. Note that any of the previously mentioned priority rules could be used on a station-by-station basis for this situation; the only difference is that the S/O approach incorporates downstream information in arriving at a job sequence.

In reality, many priority rules are available to sequence jobs, and some other rule might provide superior results for a given set of circumstances. The purpose in examining these few rules is to provide insight into the nature of sequencing rules. Each shop or organization should consider carefully its own circumstances and the measures of effectiveness it feels are important, when selecting a rule to use.

The following section describes a special-purpose algorithm that can be used to sequence a set of jobs that must all be processed at the same two machines or work centers.

Sequencing Jobs through Two Work Centers[1]

Johnson's rule is a technique that managers can use to minimize the makespan for a group of jobs to be processed on two machines or at two successive work centers (sometimes referred to as a two-machine flow shop).[2] It also minimizes the total idle time at the work centers. For the technique to work, several conditions must be satisfied:

Johnson's rule Technique for minimizing makespan for a group of jobs to be processed on two machines or at two work centers.

- Job time (including setup and processing) must be known and constant for each job at each work center.
- Job times must be independent of the job sequence.
- All jobs must follow the same two-step work sequence.
- A job must be completed at the first work center before the job moves on to the second work center.

Application of Johnson's rule begins with a listing of all jobs to be scheduled, and how much time will be required by each job at each workstation. The sequence is determined by following these steps:

1. Select the job with the shortest time. If the shortest time is at the first work center, schedule that job first; if the time is at the second work center, schedule the job last. Break ties arbitrarily.
2. Eliminate the job and its time from further consideration.
3. Repeat steps 1 and 2, working toward the center of the sequence, until all jobs have been scheduled.

Successful application of these steps identifies the sequence with the minimum makespan, or all work is completed as soon as possible. However, precisely *when* a certain job will be completed (its flow time) or when idle time will occur is not apparent by inspecting the sequence. To determine such detailed performance information, it is generally easiest to create a Gantt chart illustrating the finished sequence, as demonstrated in Example 4.

When significant idle time at the second work center occurs, job splitting at the first center just prior to the occurrence of idle time may alleviate some of it and also shorten throughput time. In Example 4, this is not a concern. The last solved problem at the end of this chapter illustrates the use of job splitting.

Using Johnson's Rule to Sequence Jobs

EXAMPLE 4

excel
mhhe.com/stevenson13e

A group of six jobs is to be processed through a two-machine flow shop. The first operation involves cleaning and the second involves painting. Determine a sequence that will minimize the total completion time for this group of jobs. Processing times are as follows:

	PROCESSING TIME (hours)	
Job	Work Center 1	Work Center 2
A	5	5
B	4	3
C	8	9
D	2	7
E	6	8
F	12	15

To employ Johnson's rule, create a "blank" sequence first, such as:

1st	2nd	3rd	4th	5th	6th

[1] For a description of a heuristic that can be used for the case where a set of jobs is to be processed through more than two work centers, see Thomas Vollmann et al., *Manufacturing Planning and Control Systems,* 5th ed. (New York: Irwin/McGraw-Hill, 2004).

[2] S. M. Johnson, "Optimal Two- and Three-Stage Production with Setup Times Included," *Naval Research Quarterly* 1 (March 1954), pp. 61–68.

SOLUTION

a. Select the job with the shortest processing time. It is job D, with a time of two hours.

b. Since the time is at the first center, schedule job D first. Eliminate job D from further consideration.

1st	2nd	3rd	4th	5th	6th
D					

c. Job B has the next shortest time. Since it is at the second work center, schedule it last and eliminate job B from further consideration. We now have

1st	2nd	3rd	4th	5th	6th
D					B

d. The remaining jobs and their times are

Job	**1**	**2**
A	5	5
C	8	9
E	6	8
F	12	15

Note that there is a tie for the shortest remaining time; job A has the same time at each work center. It makes no difference, then, whether we place it toward the beginning or the end of the sequence. Suppose it is placed arbitrarily toward the end. We now have

1st	2nd	3rd	4th	5th	6th
D				A	B

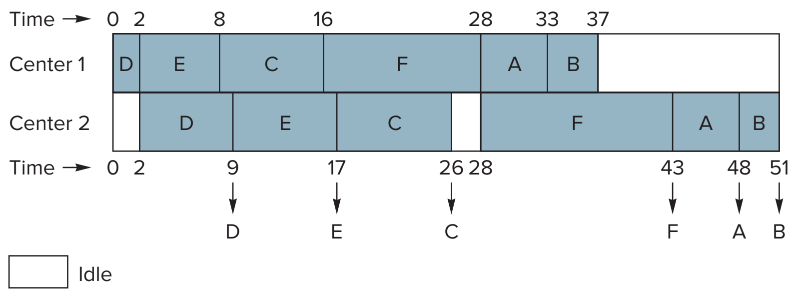
e. The shortest remaining time is six hours for job E at work center 1. Thus, schedule that job toward the beginning of the sequence (after job D):

1st	2nd	3rd	4th	5th	6th
D	E			A	B

f. Job C has the shorter time of the remaining two jobs. Since it is for the first work center, place it third in the sequence. Finally, assign the remaining job (F) to the fourth position and the result is

1st	2nd	3rd	4th	5th	6th
D	E	C	F	A	B

g. Construct a Gantt chart to reveal flow time and idle time information. Be very careful not to schedule the beginning of work at center 2 *before* work at center 1 has been completed for any given job. Traditionally, it is assumed that center 1 must finish and pass the job to center 2, which can cause idle time in center 2's schedule, such as in the case of job F as follows:



Thus, the group of jobs will take 51 hours to complete. The second work center will wait two hours for its first job and also wait two hours after finishing job C. Center 1 will be finished in 37 hours. Of course, idle periods at the beginning or end of the sequence could be used to do other jobs or for maintenance or setup/teardown activities.

Sequencing Jobs When Setup Times Are Sequence-Dependent

The preceding discussion and examples assumed that machine setup times are independent of processing order, but in many instances that assumption is not true. Consequently, a manager may want to schedule jobs at a workstation taking those dependencies into account. The goal is to minimize total setup time.

Consider the following table, which shows workstation machine setup times based on job processing order. For example, if job A is followed by job B, the setup time for B will be six hours. Furthermore, if job A is completed first, followed by job B, job C will then follow job B and have a setup time of four hours. If a job is done first, its setup time will be the amount shown in the setup time column to the right of the job. Thus, if job A is done first, its setup time will be three hours.

			Resulting following job setup time (hrs.) is		
		Setup time (hrs.)	A	B	C
If the preceding job is	**A**	**3**	—	6	2
	B	**2**	1	—	4
	C	**2**	5	3	—

The simplest way to determine which sequence will result in the lowest total setup time is to list each possible sequence and determine its total setup time. In general, the number of different alternatives is equal to $n!$, where n is the number of jobs. Here, n is 3, so $n! = 3 \times 2 \times 1 = 6$. The six alternatives and their total setup times are as follows:

Sequence	**Setup Times**	**Total**
A-B-C	3 + 6 + 4 = 13	
A-C-B	3 + 2 + 3 = 8	
B-A-C	2 + 1 + 2 = 5 (best)	
B-C-A	2 + 4 + 5 = 11	
C-A-B	2 + 5 + 6 = 13	
C-B-A	2 + 3 + 1 = 6	

Hence, to minimize total setup time, the manager would select sequence B-A-C.

This procedure is relatively simple to do manually when the number of jobs is two or three. However, as the number of jobs increases, the list of alternatives quickly becomes larger. For example, six jobs would have 720 alternatives. In such instances, a manager would employ a computer to generate the list and identify the best alternative(s). (Note that more than one alternative may be tied for the lowest setup time.)

Why Scheduling Can Be Difficult

Scheduling can be difficult for a number of reasons. One is that, in reality, an operation must deal with variability in setup times, processing times, interruptions, and changes in the set of

© Alex Mcaclean/Getty

jobs. Another major reason is that, except for small job sets, there is no method for identifying the optimal schedule, and it would be virtually impossible to sort through the vast number of possible alternatives to obtain the best schedule. As a result, scheduling is far from an exact science and, in many instances, is an ongoing task for a manager.

Computer technology reduces the burden of scheduling and makes real-time scheduling possible.

Minimizing Scheduling Difficulties

There are a number of actions that managers can consider to minimize scheduling problems:

- Setting realistic due dates.
- Focusing on bottleneck operations: First, try to increase the capacity of the operations. If that is not possible or feasible, schedule the bottleneck operations first, and then schedule the nonbottleneck operations around the bottleneck operations.
- Considering lot splitting for large jobs. This usually works best when there are relatively large differences in job times. Note that this doesn't apply to single-unit jobs.

The Theory of Constraints

LO16.9 Discuss the theory of constraints and that approach to scheduling.

Another approach to scheduling was developed and promoted by Eli Goldratt.[3] He first described it in his book *The Goal.* Goldratt avoided much of the complexity often associated with scheduling problems by simply focusing on *bottleneck* operations (i.e., those for which there was insufficient capacity—in effect, a work center with zero idle time). He reasoned the output of the system was limited by the output of the bottleneck operation(s); therefore, it was essential to schedule the nonbottleneck operations in a way that minimized the idle time of the bottleneck operation(s). Thus, idle time of nonbottleneck operations was not a factor in overall productivity of the system, as long as the bottleneck operations were used effectively. These observations have been refined into a series of scheduling principles which include:

- An hour lost at a bottleneck operation is an hour lost by the system. The bottleneck operation determines the overall capacity of the system.
- Saving time through improvements of a nonbottleneck will not increase the ultimate output of the system.
- Activation of a resource is not the same as utilization of a resource. Because a nonbottleneck operation is active does not necessarily mean it is being useful.

These principles are also the foundation of a specific scheduling technique for intermittent production systems, one that many firms have found simpler and less time-consuming to use than traditional analytical techniques. This technique uses a *drum-buffer-rope* conceptualization to manage the system. The "drum" is the schedule; it sets the pace of production. The goal is to schedule to make maximum use of bottleneck resources. The "buffer" refers to potentially constraining resources outside of the bottleneck. The role of the buffer is to keep a small amount of inventory ahead of the bottleneck operation to minimize the risk of having it be idle. The "rope" represents the synchronizing of the sequence of operations to ensure effective use of the bottleneck operations. The goal is to avoid costly and time-consuming multiple setups, particularly of capacity-constrained resources, so they do not become bottlenecks too.

[3]Eli Goldratt, *The General Theory of Constraints* (New Haven, CT: Avraham Y. Institute, 1989).

The drum-buffer-rope approach provides a basis for developing a schedule that achieves maximum output and shorter lead times while avoiding carrying excess inventory. Use of the drum-buffer-rope approach generally results in operations capable of consistent on-time delivery, reduced inventory, and shorter lead times, as well as a reduction in disruptions that require expediting.

Goldratt also developed a system of varying batch sizes to achieve the greatest output of bottleneck operations. He used the term **process batch** to denote the basic lot size for a job and the term **transfer batch** to denote a portion of the basic lot that could be used during production to facilitate utilization of bottleneck operations. In effect, a lot could be split into two or more parts. Splitting a large lot at one or more operations preceding a bottleneck operation would reduce the waiting time of the bottleneck operation.

Process batch The economical quantity to produce upon the activation of a given operation.

Transfer batch The quantity to be transported from one operation to another, assumed to be smaller than the first operation's process batch.

Traditional management has emphasized maximizing output of every operation. In contrast to that approach, the **theory of constraints** has as its goal maximizing flow through the entire system, which it does by emphasizing balancing the flow through the various operations. It begins with identifying the bottleneck operation. Next, there is a five-step procedure to improve the performance of the bottleneck operation:

Theory of constraints Production planning approach that emphasizes balancing flow throughout a system, and pursues a perpetual five-step improvement process centered around the system's currently most restrictive constraint.

1. Determine what is constraining the operation.
2. Exploit the constraint (i.e., make sure the constraining resource is used to its maximum).
3. Subordinate everything to the constraint (i.e., focus on the constraint).
4. Determine how to overcome (eliminate) the constraint.
5. Repeat the process for the next highest constraint.

(Note the similarity to the plan-do-study-act [PDSA] approach discussed in Chapter 9.)

The goal, of course, is to make *money*. The theory of constraints uses three metrics to assess the effectiveness of improvements:

- **Throughput:** The rate at which the system generates *money* through sales (i.e., the contribution margin, or sales revenue less variable costs; labor costs are considered to be part of operating expense)
- **Inventory:** Inventory represents *money* tied up in goods and materials used in a process
- **Operating expense:** All the *money* the system spends to convert inventory into throughput; this includes utilities, scrap, depreciation, and so on

Goldratt's ideas are applicable to both manufacturing and service environments.

16.3 SCHEDULING SERVICES

LO16.10 Summarize some of the unique problems encountered in service systems, and describe some of the approaches used for scheduling service systems.

Scheduling service systems presents certain problems not generally encountered in manufacturing systems. This is due primarily to (1) the inability to store or inventory services and (2) the random nature of customer requests for service. In some situations, the second difficulty can be moderated by using appointment or reservation systems, but the inability to store services in most cases is a fact of life that managers must contend with.

The approach used to schedule services generally depends on whether customer contact is involved. In back-office operations, where there is little or no customer contact—such as processing mail-order requests, loan approvals, and tax preparation—the same priority rules described in the preceding pages are used. The goal is to maximize worker efficiency, and work is often processed in batches. A key factor can be the due date, say for rush orders, orders where the customer has paid a premium for faster than normal delivery. That is similar to the situation that occurs in front-office operations, where there is a high degree of customer contact, and efficiency may become secondary to keeping customer waiting times to reasonable levels, so scheduling the workforce to meet demand becomes a priority. Having too few workers causes waiting lines to form, but having more workers than needed increases labor costs, which can have a substantial impact on profits, particularly in service systems where labor is the major cost involved.

An ideal situation is one that has a smooth flow of customers through the system. This would occur if each new customer arrives at the precise instant that the preceding customer's service is completed, as in a physician's office, or in air travel if the demand just equals the number of available seats. In each of these situations customer waiting time would be minimized, and the service system staff and equipment would be fully utilized. Unfortunately, the random nature of customer requests for service that generally prevails in service systems makes it nearly impossible to provide service capability that matches demand. Moreover, if service times are subject to variability—say, because of differing processing requirements—the inefficiency of the system is compounded. The inefficiencies can be reduced if arrivals can be scheduled (e.g., appointments), as in the case of doctors and dentists. However, in many situations appointments are not possible (supermarkets, gas stations, theaters, hospital emergency rooms, repair of equipment breakdowns). Chapter 18, on waiting lines, focuses on those kinds of situations. There, the emphasis is on intermediate-term decisions related to service capacity. In this section, we will concern ourselves with short-term *scheduling,* in which much of the capacity of a system is essentially fixed, and the goal is to achieve a certain degree of customer service by efficient utilization of that capacity.

Scheduling in service systems may involve scheduling (1) customers, (2) the workforce, and (3) equipment. Scheduling customers often takes the form of appointment systems or reservation systems.

Appointment Systems

Appointment systems are intended to control the timing of customer arrivals in order to minimize customer waiting while achieving a high degree of capacity utilization. A doctor can use an appointment system to schedule patients' office visits during the afternoon, leaving the mornings free for hospital duties. Similarly, an attorney can schedule client meetings around court appearances. Even with appointments, however, problems can still arise due to lack of punctuality on the part of patients or clients, no-shows, and the inability to completely control the length of contact time (e.g., a dentist might run into complications in filling a tooth and have to spend additional time with a patient, thus backing up later appointments). Some of this can be avoided by trying to match the time reserved for a patient or client with the specific needs of that case rather than setting appointments at regular intervals. Even with the problems of late arrivals and no-shows, the appointment system is a tremendous improvement over random arrivals.

Reservation Systems

Reservation systems are designed to enable service systems to formulate a fairly accurate estimate of the demand on the system for a given time period and to minimize customer disappointment generated by excessive waiting or inability to obtain service. Reservation systems are widely used by resorts, hotels and motels, restaurants, and some modes of transportation (e.g., airlines, car rentals). In the case of restaurants, reservations enable management to spread out or group customers so that demand matches service capabilities. Late arrivals and no-shows can disrupt the system. One approach to the no-show problem is to use decision theory (described in the supplement to Chapter 5). The problem also can be viewed as a single-period inventory problem, as described in Chapter 13.

Yield Management

Many companies, especially in the travel and tourist industries, operate with fixed capacities. Examples include hotels and motels, which operate with a fixed number of rooms to rent each night; airlines, which operate with a fixed number of seats to sell on any given flight; and cruise lines, which operate with a fixed number of berths to sell for any given cruise. The number of rooms, seats, or berths can be thought of as perishable inventory. For example, unsold seats on a flight cannot be carried over to the next flight; they are lost. The same is true for hotel rooms and cruise berths. Of course that unsold inventory does not generate income, so companies with fixed capacities must develop strategies to deal with sales.

© Mark Wilson/Getty

The basic yield management concept is applicable to railroads as well as airlines. Yield management is multidisciplinary because it blends elements of marketing, operations, and financial management.

Yield management is the application of pricing strategies to allocate capacity among various categories of demand with the goal of maximizing the revenue generated by the fixed capacity. Demand for fixed capacity usually consists of customers who make advance reservations and walk-ins. Customers who make advance reservations are typically price-sensitive, while walk-ins are often price-insensitive. Companies must decide on the percentage of their limited inventory to allocate to reservations, trading off lower revenue per unit for increased certainty of sales, and how much to allocate to walk-ins, where demand is less certain but revenue per unit is higher.

Yield management The application of pricing strategies to allocate capacity among various categories of demand.

The ability to predict demand is critical to the success of yield management, so forecasting plays a key role in the process. Seasonal variations are generally important, so forecasts must incorporate seasonality and plans must also be somewhat flexible to allow for ever-present random variations.

Scheduling the Workforce

Scheduling customers is demand management. Scheduling the workforce is capacity management. This approach works best when demand can be predicted with reasonable accuracy. This is often true for restaurants, theaters, rush-hour traffic, and similar instances that have repeating patterns of intensity of customer arrivals. Scheduling hospital personnel, police, and telephone operators for catalog sales, credit card companies, and mutual fund companies also comes under this heading. An additional consideration is the extent to which variations in customer demands can be met with workforce flexibility. Thus, capacity can be adjusted by having cross-trained workers who can be temporarily assigned to help out on bottleneck operations during periods of peak demand.

Various constraints can affect workforce scheduling flexibility, including legal, behavioral, technical—such as workers' qualifications to perform certain operations—and budget constraints. Union or federal work rules and vacations can make scheduling more complicated.

Cyclical Scheduling

In many services (e.g., hospitals, police departments, fire departments, restaurants, and supermarkets) the scheduling requirements are fairly similar: Employees must be assigned to work shifts or time slots, and have days off, on a repeating or cyclical basis. Here is a method for determining both a schedule and the minimum number of workers needed.

Generally a basic work pattern is set (e.g., work five consecutive days, have two consecutive days off), and a list of staffing needs for the schedule cycle (usually one week) is given. For example:

Day	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Staff needed	2	4	3	4	6	5	5

A fairly simple but effective approach for determining the minimum number of workers needed is the following: Begin by repeating the staff needs for worker 1. Then,

1. Make the first worker's assignment such that the two days with the lowest need (i.e., lowest sum) are designated as days off. Here Mon–Tues has the two lowest consecutive requirements. Circle those days. (Note, in some instances, Sun–Mon might yield the two lowest days.) In case of a tie, pick the pair with the lowest adjacent requirement day to the left and day to the right. If there is still a tie, pick arbitrarily.

Day	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Staff needed	**2**	**4**	**3**	**4**	**6**	**5**	**5**
Worker 1	2	4	3	4	6	5	5

2. Subtract one from each day's requirement, except for the circled days. Assign the next employee, again using the two lowest consecutive days as days off. Circle those days.

Day	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Staff needed	**2**	**4**	**3**	**4**	**6**	**5**	**5**
Worker 1	2	4	3	4	6	5	5
Worker 2	2	4	2	3	5	4	4

3. Repeat the preceding step for each additional worker until all staffing requirements have been met. However, don't subtract from a value of zero. Note the tie for worker 3: Mon–Tue and Sun–Mon have the lowest consecutive requirements, 4. The Mon–Tue two adjacents are Sun = 3 and Wed = 2, for a total of 5, which is less than the two adjacents for Sun–Mon (Sat = 3 and Tue = 3 for a total of 6). So circle Mon–Tue requirements for worker 3. Worker 4 also has a tie, and adjacents Sat and Tue total 5, whereas Tue–Fri adjacents total 6, so circle Sun–Mon requirements.

Day	Mon	Tue	Wed	Thu	Fri	Sat	Sun	
Staff needed	**2**	**4**	**3**	**4**	**6**	**5**	**5**	
Worker 1	2	4	3	4	6	5	5	
Worker 2	2	4	2	3	5	4	4	
Worker 3	1	3	2	3	4	3	3	
Worker 4	1	3	1	2	3	2	2	(tie)
Worker 5	1	2	0	1	2	1	2	
Worker 6	0	1	0	1	1	0	1	(multiple ties)
Worker 7	0	1	0	0	0	0	0	(multiple ties)
No. working:	**2**	**4**	**3**	**4**	**6**	**5**	**5**	

For Worker 7, circle Wed and Thu 0 0.

To identify the days each worker is working, go across each worker's row to find the nonzero values that are not circled, signifying that the worker is assigned for those days. Similarly, to find the workers who are assigned to work for any particular day, go down that day's column to find the nonzero values that are not circled. *Note:* Worker 6 will only work three days, and worker 7 will only work one day.

Scheduling Multiple Resources

In some situations, it is necessary to coordinate the use of more than one resource. For example, hospitals must schedule surgeons, operating room staffs, recovery room staffs, admissions, special equipment, nursing staffs, and so on. Educational institutions must schedule faculty, classrooms, audiovisual equipment, and students. As you might guess, the greater the number of resources to be scheduled simultaneously, the greater the complexity and the less likely that an optimum schedule can be achieved. The problem is further complicated by the variable nature of such systems. For example, educational institutions frequently change their course offerings, student enrollments change, and students exhibit different course-selection patterns.

Some schools and hospitals are using computer software to assist them in devising acceptable schedules, although many appear to be using intuitive approaches with varying degrees of success.

Airlines are another example of service systems that require the scheduling of multiple resources. Flight crews, aircraft, baggage handling equipment, ticket counters, gate personnel, boarding ramps, food service, cleaning, and maintenance personnel all have to be coordinated. Furthermore, government regulations on the number of hours a pilot can spend flying place an additional restriction on the system. Another interesting variable is that, unlike most systems, the flight crews and the equipment do not remain in one location. Moreover, the crew and the equipment are not usually scheduled as a single unit. Flight crews are often scheduled so that they return to their base city every two days or more, and rest breaks must be considered. On the other hand, the aircraft may be in almost continuous use except for periodic maintenance and repairs. Consequently, flight crews commonly follow different trip patterns than that of the aircraft.

Service systems are prone to slowdowns when variability in demand for services causes bottlenecks. Part of the difficulty lies in predicting which operations will become bottlenecks. Moreover, bottlenecks may shift with the passage of time, so that different operations become bottleneck operations—further complicating the problem.

16.4 OPERATIONS STRATEGY

Scheduling can either help or hinder operations strategy. If scheduling is done well, goods or services can be made or delivered in a timely manner. Resources can be used to best advantage and customers will be satisfied. Scheduling not performed well will result in inefficient use of resources and possibly dissatisfied customers.

The implication is clear: Management should not overlook the important role that scheduling plays in the success of an organization and the supply chain, giving a competitive advantage if done well or disadvantage if done poorly. Time-based competition depends on good scheduling. Coordination of materials, equipment use, and employee time is an important function of operations management. It is not enough to have good design, superior quality, and the other elements of a well-run organization if scheduling is done poorly—just as it is not enough to own a well-designed and well-made car, with all the latest features for comfort and safety, if the owner doesn't know how to drive it!

SUMMARY

Scheduling involves the timing and coordination of operations. Such activities are fundamental to virtually every organization. Scheduling problems differ according to whether a system is designed for high volume, intermediate volume, or low volume. Scheduling problems are particularly complex for job shops (low volume) because of the variety of jobs these systems are required to process.

The two major problems in job-shop scheduling are assigning jobs to machines or work centers, and designating the sequence of job processing at a given machine or work center. Gantt load charts are frequently employed to help managers visualize workloads, and they are useful for describing and analyzing sequencing alternatives. In addition, both heuristic and optimizing methods are used to develop loading and sequencing plans. For the most part, the optimization techniques can be used only if certain assumptions can be made.

Customer requirements in service systems generally present very different circumstances than those encountered in manufacturing systems. Some services can use appointments and reservations for scheduling purposes, although not all systems are amenable to this. When multiple resources are involved, the task of balancing the system can be fairly complex.

KEY POINTS

1. Scheduling occurs in every business organization.
2. Scheduling decisions are made within constraints established by decisions on capacity, product or service design, process selection and layout, aggregate planning, and master scheduling.
3. Scheduling decisions occur just prior to the conversion of inputs into outputs.
4. Effective scheduling can reduce costs and increase productivity.

KEY TERMS

assignment model, 699
backward scheduling, 697
finite loading, 697
flow-shop scheduling, 692
flow system, 692
forward scheduling, 697
Gantt chart, 695
global priority rules, 702
Hungarian method, 699
infinite loading, 697
input/output (I/O) control, 698
job flow time, 703
job lateness, 703
job-shop scheduling, 695
job time, 702
Johnson's rule, 709
load chart, 696
loading, 695
local priority rules, 702
makespan, 703
priority rules, 702
process batch, 713
schedule chart, 697
scheduling, 691
sequencing, 702
theory of constraints, 713
transfer batch, 713
workstation, 702
yield management, 715

SOLVED PROBLEMS

Problem 1

The assignment method. The following table contains information on the cost to run three jobs on four available machines. Determine an assignment plan that will minimize costs.

		MACHINE			
		A	**B**	**C**	**D**
	1	12	16	14	10
Job	**2**	9	8	13	7
	3	15	12	9	11

Solution

In order for us to be able to use the assignment method, the numbers of jobs and machines must be equal. To remedy this situation, add a *dummy* job with costs of 0, and then solve as usual.

		MACHINE			
		A	**B**	**C**	**D**
	1	12	16	14	10
Job	**2**	9	8	13	7
	3	15	12	9	11
(dummy)	**4**	0	0	0	0

a. Subtract the smallest number from each row. The results are

		MACHINE			
		A	**B**	**C**	**D**
	1	2	6	4	0
Job	**2**	2	1	6	0
	3	6	3	0	2
	4	0	0	0	0

b. Subtract the smallest number in each column. (Because of the dummy zeros in each column, the resulting table will be unchanged.)

c. Determine the minimum number of lines needed to cross out the zeros. One possible way is as follows:

		MACHINE			
		A	**B**	**C**	**D**
	1	2	6	4	0
Job	**2**	2	1	6	0
	3	6	3	0	2
	4	0	0	0	0

d. Because the number of lines is less than the number of rows, modify the numbers.

(1) Subtract the smallest uncovered number (1) from each uncovered number.

(2) Add the smallest uncovered number to numbers at line intersections. The result is:

		MACHINE			
		A	**B**	**C**	**D**
	1	1	5	4	0
Job	**2**	1	0	6	0
	3	5	2	0	2
	4	0	0	1	1

e. Test for optimality:

		MACHINE			
		A	**B**	**C**	**D**
	1	1	5	4	0
Job	**2**	1	0	6	0
	3	5	2	0	2
	4	0	0	1	1

Because the minimum number of lines equals the number of rows, an optimum assignment can be made.

f. Assign jobs to machines. Start with rows 1 and 3, since they each have one zero, and columns A and C, also with one zero each. After each assignment, cross out all the numbers in that row *and* column. The result is:

		MACHINE			
		A	**B**	**C**	**D**
Job	**1**	1	5	4	0
	2	1	0	6	0
	3	5	2	0	2
	4	0	0	1	1

Notice that there is only one assignment in each row, and only one assignment in each column.

g. Compute total costs, referring to the original table.

1-D	$10
2-B	8
3-C	9
4-A	0
	$27

h. The implication of assignment 4-A is that machine A will not be assigned a job. It may remain idle or be used for another job.

Problem 2

Priority rules. Job times (including processing and setup) are shown in the following table for five jobs waiting to be processed at a work center.

Job	Job Time (hours)	Due Date (hours)
a	12	15
b	6	24
c	14	20
d	3	8
e	7	6

Determine the processing sequence that would result from each of these priority rules:

a. SPT b. EDD

Solution

Assume job times are independent of processing sequence.

	a. SPT		b. EDD	
Job	**Job Time**	**Processing Order**	**Hour Due**	**Processing Order**
a	12	4	15	3
b	6	2	24	5
c	14	5	20	4
d	3	1	8	2
e	7	3	6	1

Problem 3

Priority rules. Using the job times and due dates from Solved Problem 2, determine each of the following performance measures for first-come, first-served processing order: Assume jobs listed in order of arrival.

a. Makespan

b. Average flow time

c. Average tardiness

d. Average number of jobs at the workstation

Solution

Job	Job Time	Flow Time	Hour Due	Hours Tardy
a	12	12	15	0
b	6	18	24	0
c	14	32	20	12
d	3	35	8	27
e	7	42	6	36
Total		139		75

a. Makespan $= 42$ hours

b. Average flow time $= \dfrac{\text{Total flow time}}{\text{Number of jobs}} = \dfrac{139}{5} = 27.80$ hours

c. Average tardiness $= \dfrac{\text{Total hours tardy}}{\text{Number of jobs}} = \dfrac{75}{5} = 15$ hours

d. Average number of jobs at workstation $= \dfrac{\text{Total flow time}}{\text{Makespan}} = \dfrac{139}{42} = 3.31$

Problem 4

S/O rule. Using the following information, determine an order processing sequence using the S/O priority rule.

Order	Processing Time Remaining (days)	Due Date (days)	Number of Operations Remaining
A	20	30	2
B	11	18	5
C	10	6	2
D	16	23	4

Solution

Assume times are independent of processing sequence.

Order	(1) Remaining Processing Time	(2) Due Date	(3) (2) − (1) Slack	(4) Number of Operations	(5) Ratio	(6) Rank (sequence)
A	20	30	10	2	5.00	4
B	11	18	7	5	1.40	2
C	10	6	−4	2	−2.00	1
D	16	23	7	4	1.75	3

(Note that one ratio is negative. When negatives occur, assign the *lowest* rank to the *most negative* number.)

Problem 5

Sequencing jobs through two work centers. Use Johnson's rule to obtain the optimum sequence for processing the jobs shown through work centers A and B.

	JOB TIMES (hours)	
Job	Work Center A	Work Center B
a	2.50	4.20
b	3.80	1.50
c	2.20	3.00
d	5.80	4.00
e	4.50	2.00

Solution

a. Identify the smallest time: job b (1.50 hours at work center B). Because the time is for B, schedule this job last.

b. The next smallest time is job e (2.00 hours at B). Schedule job e next to last.

c. Identify the smallest remaining job time: job c (2.20 hours at center A). Since the time is in the A column, schedule job c first. At this point, we have:
c, ________, ________, e, b

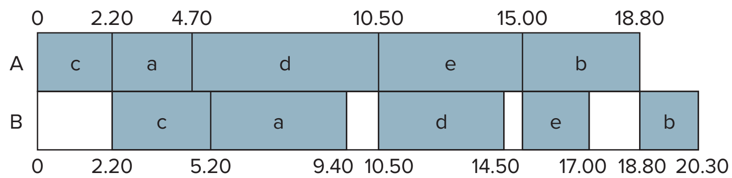
d. The smallest time for the remaining jobs is 2.50 hours for job a at center A. Schedule this job after job c. The one remaining job (job d) fills the remaining slot. Thus, we have c-a-d-e-b.

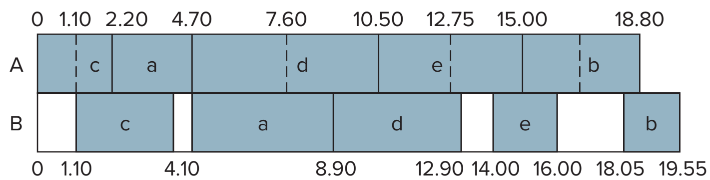
Problem 6

For Solved Problem 5, determine what effect splitting jobs c, d, e, and b in work center A would have on the idle time of work center B and on the throughput time. Assume that each job can be split into two equal parts.

Solution

We assume that the processing sequence remains unchanged and proceed on that basis. The solution from the previous problem is shown in the following chart. The next chart shows reduced idle time at center B when splitting is used.





An inspection of these two figures reveals that throughput time has decreased from 20.30 hours to 19.55 hours. In addition, the original idle time was 5.6 hours. After splitting certain jobs, it was reduced to 4.85 hours, so some improvement was achieved. Note that processing times at B are generally less than at A for jobs toward the end of the sequence. As a result, jobs such as e and b at B were scheduled so that they were *centered* around the finishing times of e and b, respectively, at A, to avoid having to break the jobs due to waiting for the remainder of the split job from A. Thus, the greatest advantage from job splitting generally comes from splitting earlier jobs when Johnson's rule is used for sequencing.

DISCUSSION AND REVIEW QUESTIONS

1. Why is scheduling fairly simple for repetitive systems but fairly complex for job shops?
2. What are the main decision areas of job-shop scheduling?
3. What are Gantt charts? How are they used in scheduling? What are the advantages of using Gantt charts?
4. What are the basic assumptions of the assignment method of linear programming?
5. Briefly describe each of these priority rules:
   a. FCFS
   b. SPT
   c. EDD
   d. S/O
   e. Rush
6. Why are priority rules needed?
7. What problems not generally found in manufacturing systems do service systems present in terms of scheduling the use of resources?
8. Explain forward and backward schedulings and each one's advantage.
9. How are scheduling and productivity related?
10. What factors would you take into account in deciding whether to split a job?
11. Explain the term *makespan.*

TAKING STOCK

1. What general trade-offs are involved in sequencing decisions? In scheduling decisions?
2. Who needs to be involved in setting schedules?
3. How has technology had an impact on scheduling?

CRITICAL THINKING EXERCISES

1. One approach that can be effective in reducing the impact of production bottlenecks in a job shop or batch operations setting is to use smaller lot sizes.
   a. What is the impact of a production bottleneck?
   b. Explain how small lot sizes can reduce the impact of bottleneck operations.
   c. What are the key trade-offs in using small lot sizes for the purpose of reducing the bottleneck effect?
   d. In some cases, the location of a bottleneck will shift (i.e., sometimes it is at workstation 3, another time it is at workstation 12). Furthermore, there can be more than one bottleneck operation at the same time. How would these situations impact scheduling using small lot sizes?
2. Doctors' and dentists' offices frequently schedule patient visits at regularly spaced intervals. What problems can this create? Can you suggest an alternative approach to reduce these problems? Under what circumstances would regularly spaced appointments constitute a reasonable approach to patient scheduling?
3. Name three examples of unethical behavior involving scheduling and state the ethical principle each violates.

PROBLEMS

1. Use the assignment method to determine the best way to assign workers to jobs, given the following cost information. Compute the total cost for your assignment plan.

		JOB		
		A	**B**	**C**
	1	5	8	6
Worker	**2**	6	7	9
	3	4	5	3

2. Rework Problem 1, treating the numbers in the table as profits instead of costs. Compute the total profit.
3. Assign trucks to delivery routes so that total costs are minimized, given the cost data shown. What is the total cost?

		ROUTE				
		A	**B**	**C**	**D**	**E**
	1	4	5	9	8	7
	2	6	4	8	3	5
Truck	**3**	7	3	10	4	6
	4	5	2	5	5	8
	5	6	5	3	4	9

4. Develop an assignment plan that will minimize processing costs, given the information shown, and interpret your answer.

		WORKER		
		A	**B**	**C**
	1	12	8	11
Job	**2**	13	10	8
	3	14	9	14
	4	10	7	12

5. Use the assignment method to obtain a plan that will minimize the processing costs in the following table under these conditions:
   a. The combination 2-D is undesirable
   b. The combinations 1-A and 2-D are undesirable

		WORKER				
		A	**B**	**C**	**D**	**E**
	1	14	18	20	17	18
	2	14	15	19	16	17
Job	**3**	12	16	15	14	17
	4	11	13	14	12	14
	5	10	16	15	14	13

6. The following table contains information concerning four jobs that are awaiting processing at a work center.

Job	Job Time (days)	Due Date (days)
A	14	20
B	10	16
C	7	15
D	6	17

a. Sequence the jobs using (1) FCFS, (2) SPT, (3) EDD, and (4) CR. Assume the list is by order of arrival.
b. For each of the methods in part *a,* determine (1) the average job flow time, (2) the average tardiness, and (3) the average number of jobs at the work center.
c. Is one method superior to the others? Explain.

7. Using the information presented in the following table, identify the processing sequence that would result using (1) FCFS, (2) SPT, (3) EDD, and (4) CR. For each method, determine (1) average job flow time, (2) average job tardiness, and (3) average number of jobs in the system. Jobs are listed in order of arrival. (*Hint:* First determine the total job time for each job by computing the total processing time for the job and then adding in the setup time. All times and due dates are in hours.)

Job	Processing Time per Unit	Units per Job	Setup Time	Due Date
a	.14	45	0.7	4
b	.25	14	0.5	10
c	.10	18	0.2	12
d	.25	40	1.0	20
e	.10	75	0.5	15

8. The following table shows orders to be processed at a machine shop as of 8:00 a.m. Monday. The jobs have different operations they must go through. Processing times are in days. Jobs are listed in order of arrival.

a. Determine the processing sequence at the first work center using each of these rules: (1) FCFS, (2) S/O.
b. Compute the effectiveness of each rule using each of these measures: (1) average flow time, (2) average number of jobs at the work center.

Job	Processing Time (days)	Due Date (days)	Remaining Number of Operations
A	8	20	2
B	10	18	4
C	5	25	5
D	11	17	3
E	9	35	4

9. A wholesale grocery distribution center uses a two-step process to fill orders. Tomorrow's work will consist of filling the seven orders shown. Determine a job sequence that will minimize the time required to fill the orders.

	TIME (hours)	
Order	Step 1	Step 2
A	1.20	1.40
B	0.90	1.30
C	2.00	0.80
D	1.70	1.50
E	1.60	1.80
F	2.20	1.75
G	1.30	1.40

10. The times required to complete each of eight jobs in a two-machine flow shop are shown in the table that follows. Each job must follow the same sequence, beginning with machine A and moving to machine B.

a. Determine a sequence that will minimize makespan time.
b. Construct a chart of the resulting sequence, and find machine B's idle time.

c. For the sequence determined in part *a,* how much would machine B's idle time be reduced by splitting the last two jobs in half?

	TIME (hours)	
Job	**Machine A**	**Machine B**
a	16	5
b	3	13
c	9	6
d	8	7
e	2	14
f	12	4
g	18	14
h	20	11

11. Given the operation times provided:
    a. Develop a job sequence that minimizes idle time at the two work centers.
    b. Construct a chart of the activities at the two centers, and determine each one's idle time, assuming no other activities are involved.

	JOB TIMES (minutes)					
	A	**B**	**C**	**D**	**E**	**F**
Center 1	20	16	43	60	35	42
Center 2	27	30	51	12	28	24

12. A shoe repair operation uses a two-step sequence that all jobs in a certain category follow. All jobs can be split in half at both stations. For the group of jobs listed:
    a. Find the sequence that will minimize total completion time.
    b. Determine the amount of idle time for workstation B.
    c. What jobs are candidates for splitting? Why? If they were split, how much would idle time and makespan time be reduced?

	JOB TIMES (minutes)				
	A	**B**	**C**	**D**	**E**
Workstation A	27	18	70	26	15
Workstation B	45	33	30	24	10

13. The following schedule was prepared by the production manager of Marymount Metal Shop: Determine a schedule that will result in earliest completion of all jobs on this list.

	CUTTING		POLISHING	
Job	**Start**	**Finish**	**Start**	**Finish**
A	0	2	2	5
B	2	6	6	9
C	6	11	11	13
D	11	15	15	20
E	15	17	20	23
F	17	20	23	24
G	20	21	24	28

14. The production manager must determine the processing sequence for seven jobs through the grinding and deburring departments. The same sequence will be followed in both departments. The manager's goal is to move the jobs through the two departments as quickly as possible. The foreman of the deburring department wants the SPT rule to be used to minimize the work-in-process inventory in his department.

	PROCESSING TIME (hours)	
Job	Grinding	Deburring
A	3	6
B	2	4
C	1	5
D	4	3
E	9	4
F	8	7
G	6	2

a. Prepare a schedule using SPT for the grinding department.
b. What is the flow time in the grinding department for the SPT sequence? What is the total time needed to process the seven jobs in both the grinding and deburring departments?
c. Determine a sequence that will minimize the total time needed to process the jobs in both departments. What flow time will result for the grinding department?
d. Discuss the trade-offs between the two alternative sequencing arrangements. At what point would the production manager be indifferent concerning the choice of sequences?

15. A foreman has determined processing times at a work center for a set of jobs and now wants to sequence them. Given the information shown, do the following:
    a. Determine the processing sequence using (1) FCFS, (2) SPT, (3) EDD, and (4) CR. For each sequence, compute the average job tardiness, the average flow time, and the average number of jobs at the work center. The list is in FCFS order.
    b. Using the results of your calculations in part *a,* show that the ratio of average flow time and the average number of jobs measures are equivalent for all four sequencing rules.
    c. Determine the processing sequence that would result using the S/O rule.

Job	Job Time (days)	Due Date	Operations Remaining
a	4.5	10	3
b	6.0	17	4
c	5.2	12	3
d	1.6	27	5
e	2.8	18	3
f	3.3	19	1

16. Given the information in the following table, determine the processing sequence that would result using the S/O rule.

Job	Remaining Processing Time (days)	Due Date	Remaining Number of Operations
a	5	8	2
b	6	5	4
c	9	10	4
d	7	12	3
e	8	10	2

17. Given the following information on job times and due dates, determine the optimal processing sequence using (1) FCFS, (2) SPT, (3) EDD, and (4) CR. For each method, find the average job flow time and the average job tardiness. Jobs are listed in order of arrival.

Job	Job Time (hours)	Due Date (hours)
a	3.5	7
b	2.0	6
c	4.5	18
d	5.0	22
e	2.5	4
f	6.0	20

18. The Budd Gear Co. specializes in heat-treating gears for automobile companies. At 8:00 a.m., when Budd's shop opened today, five orders (listed in order of arrival) were waiting to be processed.

Order	Order Size (units)	Per Unit Time in Heat Treatment (minutes/unit)	Due Date (min. from now)
A	16	4	160
B	6	12	200
C	10	3	180
D	8	10	190
E	4	1	220

a. If the earliest due date rule is used, what sequence should be used?
b. What will be the average job tardiness?
c. What will be the average number of jobs in the system?
d. Would the SPT rule produce better results in terms of job tardiness?

19. The following table contains order-dependent setup times for three jobs. Which processing sequence will minimize the total setup time?

			Following Job's Setup Time (hrs.)		
		Setup Time (hrs.)	A	B	C
Preceding Job	A	2	—	3	5
	B	3	8	—	2
	C	2	4	3	—

20. The following table contains order-dependent setup times for three jobs. Which processing sequence will minimize the total setup time?

			Following Job's Setup Time (hrs.)		
		Setup Time (hrs.)	A	B	C
Preceding Job	A	2.4	—	1.8	2.2
	B	3.2	0.8	—	1.4
	C	2.0	2.6	1.3	—

21. The following table contains order-dependent setup times for four jobs. For safety reasons, job C cannot follow job A, nor can job A follow job C. Determine the processing sequence that will minimize the total setup time. (*Hint:* There are 12 alternatives.)

		Setup Time (hrs.)	Following Job's Setup Time (hrs.) A	B	C	D
Preceding job	A	2	—	5	×	4
	B	1	7	—	3	2
	C	3	x	2	—	2
	D	2	4	3	6	—

22. Given this information on planned and actual inputs and outputs for a service center, determine the work backlog for each period. The beginning backlog is 12 hours of work. The figures shown are standard hours of work.

		PERIOD 1	2	3	4	5
Input	Planned	24	24	24	24	20
	Actual	25	27	20	22	24
Output	Planned	24	24	24	24	23
	Actual	24	22	23	24	24

23. Given the following data on inputs and outputs at a work center, determine the cumulative deviation and the backlog for each time period. The beginning backlog is 7.

		PERIOD 1	2	3	4	5	6
Input	Planned	200	200	180	190	190	200
	Actual	210	200	179	195	193	194

		PERIOD 1	2	3	4	5	6
Output	Planned	200	200	180	190	190	200
	Actual	205	194	177	195	193	200

24. Determine the minimum number of workers needed, and a schedule for the following staffing requirements, giving workers two consecutive days off per cycle (not including Sunday).

Day	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**	**Sat**
Staff needed	**2**	**3**	**1**	**2**	**4**	**3**

25. Determine the minimum number of workers needed, and a schedule for the following staffing requirements, giving workers two consecutive days off per cycle (not including Sunday).

Day	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**	**Sat**
Staff needed	**3**	**4**	**2**	**3**	**4**	**5**

26. Determine the minimum number of workers needed, and a schedule for the following staffing requirements, giving workers two consecutive days off per cycle (not including Sunday).

Day	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**	**Sat**
Staff needed	**4**	**4**	**5**	**6**	**7**	**8**

CASE

HI–HO, YO–YO, INC.

It was a little past 9:00 on a Monday morning when Jeff Baker walked into your office with a box of donuts.

"I've been talking with Anne about a problem we have with short-term capacity in our pad printing operation. You know, that's where we print the logo on the custom lines of yo-yos. We have received more orders than usual for July, and I want to release the orders to pad printing in a way that will enable us to meet our due date commitments in the best way possible. Would you have time to look at the order list (attached) and see what kind of schedule we should follow to do that? By the way, you have established quite a reputation in your short stay here. You have a talent for really explaining why your recommendations are the best approach in a way that all of us 'over-the-hill' managers can understand. Please be sure to do that for me, too. I want to understand why your recommendation is the best schedule and what the trade-offs are for other possible schedules—and none of that philosophical college mumbo-jumbo. Remember, I came up through the ranks. I don't have one of those sheepskins on my wall," he says with a laugh.

Since your schedule was back to normal after that MRP report you did for Anne, you agreed to look at the information. After that compliment, how could you say no? "Try to get back to me within a couple of days," Jeff said as he left your office.

After a few minutes with your old operations management text, you call the production control office to confirm the pad printing schedule. They confirm that pad printing runs one eight-hour shift per day. They tell you that due to a make-up day for flooding in June, pad printing will be running 23 days in July, beginning Friday, July 1 (they will work three Saturdays on July 9, 16, and 23, and take a one-day holiday for July 4).

You thank them for the information and then you begin to develop your plan.

Even though Jeff lacks a college degree, from what you have seen, he is very sharp. And obviously he knows good work when he sees it since he liked, and apparently understood, your past work. You resolve to cover all the bases but in a way that is as clear as possible.

PAD PRINTING ORDER LIST

Job	Date Order Received	Setup Time	Production Time	Due Date[1]
A	6/4	2 hrs.	6 days	11 July
B	6/7	4 hrs.	2 days	8 July
C	6/12	2 hrs.	8 days	25 July
D	6/14	4 hrs.	3 days	19 July
E	6/15	4 hrs.	9 days	29 July

[1]Jobs are due at the beginning of their respective due dates.

Note: Setup time is to set up the pad printer at the start of the job. Setup includes thoroughly cleaning the printing heads and ink reservoirs, installing the new pad(s) and ink supply, and carefully aligning the machine. Setup at the beginning of a new day with the same job is insignificant.

Examine the following rules and write a report to Jeff Baker summarizing your findings and advise him on which rule to use. Rules: FCFS, SPT, EDD, and CR.

SELECTED BIBLIOGRAPHY AND FURTHER READINGS

Goldratt, Eli, and Jeff Cox. *The Goal: A Process of Ongoing Improvement,* 3rd ed. Great Barrington, MA: North River Press, 2012.

Hopp, Wallace J., and Mark L. Spearman. *Factory Physics,* 3rd ed. New York: Irwin/McGraw-Hill, 2007.

Jacobs, F. Robert, William L. Berry, D. Clay Whybark, and Thomas E. Vollmann. *Manufacturing Planning and Control Systems,* 6th ed. New York: Irwin/McGraw-Hill, 2011.

Pinedo, Michael. *Planning and Scheduling in Manufacturing and Services.* New York: Springer, 2005.