CHAPTER

TEN

RECURRENCE RELATIONS AND RECURSIVE ALGORITHMS

10.1 INTRODUCTION

Suppose we ask a friend the age of his oldest daughter. He could tell us directly that she is 19 years old. Or he could tell us that she is 6 years older than his second daughter. If we ask for the age of the second daughter, instead of telling us that she is 13 years old, he might tell us that she is 5 years older than his third daughter. In turn, he could tell us that his third daughter is 2 years older than his only son. When he tells us that his only son is 6 years old, we would have no difficulty in figuring out that his third daughter is 8 years old, his second daughter is 13 years old, and his oldest daughter is 19 years old.

Let us consider another example. Suppose we ask for instructions to get from our house to the railroad station. We are told, "Go on Prospect Avenue, then go east on Green Street. After passing the public library, go onto Springfield Avenue and then go north on Neil Street. At the bus depot, turn right at the traffic light onto University Avenue. At the second traffic light, turn left and you'll see the railroad station." This, of course, is a perfectly clear way to instruct someone to go from our house to the railroad station. However, there is an alternative way to give the instruction, which is simply, "*Go to the bus depot* and turn right at the traffic light onto University Avenue. At the second traffic light, turn left and you'll see the railroad station." We note that such an instruction consists of two parts: one part tells us how to go from the bus depot to the railroad station explicitly, and the other simple and concise part makes use of our knowledge of how to get from our house to the bus depot. Suppose we are not sure how to get from our house to the bus depot. In that case, we can be further instructed to

"*Go to the public library* and onto Springfield Avenue. Then go north on Neil Street to the bus depot." Needless to say, either we know how to get from our house to the public library, or we should ask for further instruction.

The reader has undoubtedly realized what we are trying to say. In the first example, instead of telling us the age of his oldest daughter directly, our friend chose to tell us the age of his oldest daughter in terms of the age of his second-oldest daughter. Then, instead of telling us directly the age of his second-oldest daughter, he chose to tell us the age of his second-oldest daughter in terms of the age of his third-oldest daughter, and so on. In the second example, instead of spelling out explicitly all the details of some instruction, we specified the instruction partly in terms of knowledge we already have. We can make several observations about these two examples. First, using our prior knowledge can be a concise way to give information or instruction; for example, in directing someone to the railroad station, a great deal of information can be compressed into the simple statement, "Go to the bus depot." Secondly, we do need to do some work to make use of the knowledge we already have. In the first example, at a certain point, our friend must tell us directly the age of one of his children so that we can determine the ages of the other children. In the second example, we need to find out how to get to the public library before we can use the given instruction on how to get from there to the railroad station. Third, we might try to refer to some prior knowledge in successive steps (the age of our friend's son, that of his third-oldest daughter, and so on, or how to go to the public library, how to go to the bus depot, and so on.) Such a chain of references can only be terminated when we reach a point where we know explicitly what to do without referring to other prior knowledge.† In this chapter, we shall apply what we have just learned first to the specification of discrete numeric functions, and then to the specification of algorithms.

10.2 RECURRENCE RELATIONS

Consider the numeric function $\mathbf{a} = (3^0, 3^1, 3^2, \ldots, 3^r, \ldots)$. Clearly, the function can be specified by a general expression for a_r, namely,

$$a_r = 3^r \qquad r \geqq 0$$

† An American professor has written a paper in English, his native language, and wishes to have it published in a French journal. Since he knows no French, he asked his French colleague, Professor Bestougeff, to translate the paper into French for him. After the translation was completed, the American professor felt that he should include a footnote in the paper to acknowledge his colleague's contribution. He wrote the footnote, "The author wishes to thank Professor Bestougeff for translating this paper into French for him," in English, and asked Professor Bestougeff to translate it into French for him, which the professor gladly did. It then occurred to the American professor that he should also acknowledge Professor Bestougeff for translating the footnote for him. So, he wrote another footnote, "The author wishes to thank Professor Bestougeff for translating the preceding footnote into French for him," he then asked Professor Bestougeff to translate that into French, and copied the French translation *twice* as two additional footnotes. His problem was completely solved!

As was pointed out in Chap. 9, the function can also be specified by its generating function, namely,

$$A(z) = \frac{1}{1 - 3z}$$

According to our discussion in Sec. 10.1, we note that there is still another way to specify the numeric function. Since the value of a_r is three times the value of a_{r-1}, for all r, once we know the value of a_{r-1} we can compute the value of a_r. The value of a_{r-1} can, in turn, be computed as three times the value of a_{r-2}, which, again, is equal to three times the value of a_{r-3}. Eventually, we need the value of a_0, which is known to be 1. Thus, we note that the relation

$$a_r = 3a_{r-1}$$

together with the information that $a_0 = 1$ also completely specifies the numeric function **a**.

As another example, consider the sequence of numbers† known as the *Fibonacci sequence of numbers.* The sequence starts with the two numbers 1, 1 and contains numbers that are equal to the sum of their two immediate predecessors. A portion of the sequence is

$$1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

It is quite difficult in this case to obtain a general expression for the rth number in the sequence by observation, which, incidentally, is

$$a_r = \frac{1}{\sqrt{5}}\left(\frac{1 + \sqrt{5}}{2}\right)^{r+1} - \frac{1}{\sqrt{5}}\left(\frac{1 - \sqrt{5}}{2}\right)^{r+1}$$

Nor is it obvious what the generating function for the numeric function is. Incidentally, it is

$$A(z) = \frac{1}{1 - z - z^2}$$

On the other hand, the sequence can be described by the relation

$$a_r = a_{r-1} + a_{r-2}$$

together with the information $a_0 = 1$ and $a_1 = 1$.

For a numeric function $(a_0, a_1, a_2, \ldots, a_r, \ldots)$, an equation relating a_r, for any r, to one or more of the a_i's, $i < r$, is called a *recurrence relation.* A recurrence relation is also called a *difference equation,* and those two terms will be used interchangeably. In many discrete computation problems, it is sometimes easier to obtain a specification of a numeric function in terms of a recurrence relation than to obtain a general expression for the value of the numeric function at r or a closed-form expression for its generating function. It is clear that according to the

† Clearly, a numeric function can be viewed simply as a sequence of real numbers, and conversely.

recurrence relation, we can carry out a step-by-step computation to determine a_r from $a_{r-1}, a_{r-2}, \ldots$, to determine a_{r+1} from $a_r, a_{r-1}, \ldots$, and so on, provided that the value of the function at one or more points is given so that the computation can be initiated. These given values of the function are called *boundary conditions.* In the first example above, the boundary condition is $a_0 = 1$, and in the second example above, the boundary conditions are $a_0 = 1$ and $a_1 = 1$. We thus conclude that a numeric function can be described by a recurrence relation together with an appropriate set of boundary conditions. The numeric function is also referred to as the *solution of the recurrence relation.*

One step beyond determining the values of a numeric function in a step-by-step computation according to a given recurrence relation is to obtain from the recurrence relation either a general expression for the solution or a closed-form expression for its generating function. Unfortunately, no general method of solution for handling all recurrence relations is known. In the following, we shall study a class of recurrence relations known as *linear recurrence relations with constant coefficients.*

10.3 LINEAR RECURRENCE RELATIONS WITH CONSTANT COEFFICIENTS

A recurrence relation of the form

$$C_0 a_r + C_1 a_{r-1} + C_2 a_{r-2} + \cdots + C_k a_{r-k} = f(r) \tag{10.1}$$

where C_i's are constants, is called a *linear recurrence relation with constant coefficients.* The recurrence relation in (10.1) is known as a *k*th-*order recurrence relation*, provided that both C_0 and C_k are nonzero. For example,

$$2a_r + 3a_{r-1} = 2^r$$

is a first-order linear recurrence relation with constant coefficients. Also, both

$$3a_r - 5a_{r-1} + 2a_{r-2} = r^2 + 5 \tag{10.2}$$

and

$$a_r + 7a_{r-2} = 0$$

are second-order linear recurrence relations with constant coefficients. In this chapter we shall restrict our discussion to linear recurrence relations with constant coefficients both because we frequently encounter this class of recurrence relations and know how to handle them quite well.

Consider the recurrence relation in (10.2). Suppose we are given that $a_3 = 3$ and $a_4 = 6$, we can compute a_5 as

$$a_5 = \frac{-1}{3}[-5 \times 6 + 2 \times 3 - (5^2 + 5)] = 18$$

we can then compute a_6 as

$$a_6 = \frac{-1}{3}[-5 \times 18 + 2 \times 6 - (6^2 + 5)] = \frac{119}{3}$$

and so on. Also, we can compute

$$a_2 = \frac{-1}{2}[3 \times 6 - 5 \times 3 - (4^2 + 5)] = 9$$

$$a_1 = \frac{-1}{2}[3 \times 3 - 5 \times 9 - (3^2 + 5)] = 25$$

$$a_0 = \frac{-1}{2}[3 \times 9 - 5 \times 25 - (2^2 + 5)] = \frac{107}{2}$$

and so on. We conclude that (10.2), together with the values $a_3 = 3$ and $a_4 = 6$, completely specifies the discrete numeric function **a**.

In general, for a kth-order linear recurrence relation with constant coefficients as shown in (10.1), if k consecutive values of the numeric function **a**, a_{m-k}, $a_{m-k+1}, \ldots, a_{m-1}$ are known for some m, the value of a_m can be calculated according to (10.1), namely,

$$a_m = -\frac{1}{C_0}[C_1 a_{m-1} + C_2 a_{m-2} + \cdots + C_k a_{m-k} - f(m)]$$

Furthermore, the value of a_{m+1} can be computed as

$$a_{m+1} = -\frac{1}{C_0}[C_1 a_m + C_2 a_{m-1} + \cdots + C_k a_{m-k+1} - f(m+1)]$$

and the values of $a_{m+2}, a_{m+3}, \ldots$ can be computed in a similar manner. Also, the value of a_{m-k-1} can be computed as

$$a_{m-k-1} = -\frac{1}{C_k}[C_0 a_{m-1} + C_1 a_{m-2} + \cdots + C_{k-1} a_{m-k} - f(m-1)]$$

and the value of a_{m-k-2} can be computed as

$$a_{m-k-2} = -\frac{1}{C_k}[C_0 a_{m-2} + C_1 a_{m-3} + \cdots + C_{k-1} a_{m-k-1} - f(m-2)]$$

The values of $a_{m-k-3}, a_{m-k-4}, \ldots$ can be computed in a similar manner. Indeed, for a kth-order linear recurrence relation, the values of k consecutive a_i's are always sufficient to determine the numeric function **a** uniquely. In other words, the values of k consecutive a_i's constitute an appropriate set of boundary conditions.

On the other hand, for a kth-order linear recurrence relation with constant coefficients, fewer than k values of the numeric function will not be sufficient to determine the numeric function uniquely. For example, let

$$a_r + a_{r-1} + a_{r-2} = 4 \tag{10.3}$$

If we are given that $a_0 = 2$, we can find many numeric functions that will satisfy the recurrence relation as well as the given boundary condition. Thus,

$$2, 0, 2, 2, 0, 2, 2, 0, 2, 2, 0, \ldots$$

$$2, 2, 0, 2, 2, 0, 2, 2, 0, 2, 2, \ldots$$

$$2, 5, -3, 2, 5, -3, 2, 5, -3, 2, \ldots$$

are all possibilities. Yet, more than k values of the numeric function might make it impossible for the existence of a numeric function that satisfies the recurrence relation and the given boundary conditions. For example, for the recurrence relation in (10.3), if we were given that

$$a_0 = 2 \qquad a_1 = 2 \qquad a_2 = 2$$

then obviously a_0, a_1, and a_2 do not satisfy the recurrence relation. Consequently, no **a** can satisfy (10.3) and the boundary conditions.

The values of k nonconsecutive a_i's might or might not constitute an appropriate set of boundary conditions, depending on the specific recurrence relation we have. We shall not study the problem of what constitutes an appropriate set of boundary conditions here, since it is not a significantly important one. See, however, Prob. 10.8.

We should point out that if a kth-order recurrence relation is not a linear recurrence relation with constant coefficients, k consecutive values of the numeric functions might not specify uniquely a solution. For example, consider the recurrence relation

$$a_r^2 + a_{r-1} = 5$$

Given that $a_0 = 1$, we note that

$$1, 2, \sqrt{3}, \ldots$$

$$1, 2, -\sqrt{3}, \ldots$$

$$1, -2, \sqrt{7}, \ldots$$

are all solutions to the recurrence relation that satisfy the boundary condition.

We shall restrict our discussion to the solution of linear recurrence relations with constant coefficients. There is a significant advantage to restrict ourselves to this class of recurrence relations. Since we know that for a given set of boundary conditions the solution to a linear recurrence relation with constant coefficients is unique, as long as we are able to find a numeric function that satisfies the recurrence relation as well as the boundary conditions, it is *the* solution we are looking for. Such an argument should remove some of the "mystery" about the solution procedure we are going to present. We shall determine the solution by "guessing" what it will be. The justification for guessing is simply that it works: when the procedure yields a solution to the recurrence relation, it will be the correct solution.

10.4 HOMOGENEOUS SOLUTIONS

The (total) solution of a linear difference equation with constant coefficients is the sum of two parts, the *homogeneous solution*, which satisfies the difference equation when the right-hand side of the equation is set to 0, and the *particular solution*, which satisfies the difference equation with $f(r)$ on the right-hand side.† In other words, the discrete numeric function that is the solution of the difference equation is the sum of two discrete numeric functions—one is the homogeneous solution and the other is the particular solution. Let $\mathbf{a}^{(h)} = (a_0^{(h)}, a_1^{(h)}, \ldots, a_r^{(h)}, \ldots)$ denote the homogeneous solution and $\mathbf{a}^{(p)} = (a_0^{(p)}, a_1^{(p)}, \ldots, a_r^{(p)}, \ldots)$ denote the particular solution to the difference equation. Since

$$C_0 a_r^{(h)} + C_1 a_{r-1}^{(h)} + \cdots + C_k a_{r-k}^{(h)} = 0 \tag{10.4}$$

and

$$C_0 a_r^{(p)} + C_1 a_{r-1}^{(p)} + \cdots + C_k a_{r-k}^{(p)} = f(r)$$

we have

$$C_0(a_r^{(h)} + a_r^{(p)}) + C_1(a_{r-1}^{(h)} + a_{r-1}^{(p)}) + \cdots + C_k(a_{r-k}^{(h)} + a_{r-k}^{(p)}) = f(r)$$

Clearly, the total solution, $\mathbf{a} = \mathbf{a}^{(h)} + \mathbf{a}^{(p)}$ satisfies the difference equation.

At this point, the reader is probably somewhat puzzled by the fact that we are interested in the homogeneous solution to the difference equation. In the first place, it seems that the particular solution would be the solution we seek. In the second place, the homogeneous solution does not even satisfy the given difference equation. [Note that the right-hand side of (10.4) is set to 0.] However, we recall that many discrete numeric functions satisfy a given difference equation, but only one of them satisfies the given boundary conditions at the same time. As it will be seen later in Sec. 10.6, the particular solution alone will not, in general, satisfy the boundary conditions, while we can adjust the homogeneous solution so that the total solutions will satisfy the difference equation as well as the boundary conditions.

A homogeneous solution of a linear difference equation with constant coefficients is of the form $A\alpha_1^r$, where α_1 is called a *characteristic root* and A is a constant determined by the boundary conditions. Substituting $A\alpha^r$ for a_r in the difference equation with the right-hand side of the equation set to 0, we obtain

$$C_0 A\alpha^r + C_1 A\alpha^{r-1} + C_2 A\alpha^{r-2} + \cdots + C_k A\alpha^{r-k} = 0$$

This equation can be simplified to

$$C_0 \alpha^k + C_1 \alpha^{k-1} + C_2 \alpha^{k-2} + \cdots + C_k = 0$$

† For a reader with previous exposure to the topic of differential equations, the analogy between the solution of linear difference equations with constant coefficients and that of linear difference equations with constant coefficients should be quite obvious. Indeed, that is the reason we begin to use the phrase *linear difference equations with constant coefficients* here instead of adhering to the phrase *linear recurrence relations with constant coefficients* all the time.

which is called the *characteristic equation* of the difference equation. Therefore, if α_1 is one of the roots of the characteristic equation (it is for this reason that α_1 is called a characteristic root), $A\alpha_1^r$ is a homogeneous solution to the difference equation.

A characteristic equation of kth degree has k characteristic roots. Suppose the roots of the characteristic equation are distinct. In this case it is easy to verify that

$$a_r^{(h)} = A_1\alpha_1^r + A_2\alpha_2^r + \cdots + A_k\alpha_k^r$$

is also a homogeneous solution to the difference equation, where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are the distinct characteristic roots and $A_1, A_2, \ldots, A_k$ are constants which are to be determined by the boundary conditions.†

Example 10.1 Let us revisit the example of the Fibonacci sequence of numbers discussed in Sec. 10.2. The recurrence relation for the Fibonacci sequence of numbers is

$$a_r = a_{r-1} + a_{r-2}$$

The corresponding characteristic equation is

$$\alpha^2 - \alpha - 1 = 0$$

which has the two distinct roots

$$\alpha_1 = \frac{1+\sqrt{5}}{2} \qquad \alpha_2 = \frac{1-\sqrt{5}}{2}$$

It follows that

$$a_r^{(h)} = A_1\left(\frac{1+\sqrt{5}}{2}\right)^r + A_2\left(\frac{1-\sqrt{5}}{2}\right)^r$$

is a homogeneous solution where the two constants A_1 and A_2 are to be determined from the boundary conditions $a_0 = 1$ and $a_1 = 1$. □

Now suppose that some of the roots of the characteristic equation are multiple roots. Let α_1 be a root of multiplicity m. We shall show that the corresponding homogeneous solution is

$$(A_1 r^{m-1} + A_2 r^{m-2} + \cdots + A_{m-2} r^2 + A_{m-1} r + A_m)\alpha_1^r$$

where the A_i's are constants to be determined by the boundary conditions. It is clear that $A_m\alpha_1^r$ is a homogeneous solution of the difference equation in (10.1). To show that $A_{m-1}r\alpha_1^r$ is also a homogeneous solution, we recall that α_1 not only is a root of the equation

$$C_0\alpha^r + C_1\alpha^{r-1} + C_2\alpha^{r-2} + \cdots + C_k\alpha^{r-k} = 0 \tag{10.5}$$

† The question of the uniqueness of the solution will be discussed later.

but also is a root of the derivative equation of (10.5),

$$C_0 r\alpha^{r-1} + C_1(r-1)\alpha^{r-2} + C_2(r-2)\alpha^{r-3} + \cdots + C_k(r-k)\alpha^{r-k-1} = 0 \quad (10.6)$$

because α_1 is a multiple root of (10.5). Multiplying (10.6) by $A_{m-1}\alpha$ and replacing α by α_1, we obtain

$$C_0 A_{m-1} r\alpha_1^r + C_1 A_{m-1}(r-1)\alpha_1^{r-1} + C_2 A_{m-1}(r-2)\alpha_1^{r-2} + \cdots + C_k A_{m-1}(r-k)\alpha_1^{r-k} = 0$$

which shows that $A_{m-1} r\alpha_1^r$ is indeed a homogeneous solution.

The fact that α_1 satisfies the second, third, ..., $(m-1)$st derivative equations of (10.5) enables us to prove that $A_{m-2} r^2\alpha_1^r$, $A_{m-3} r^3\alpha_1^r$, ..., $A_1 r^{m-1}\alpha_1^r$ are also homogeneous solutions in a similar manner.

Example 10.2 Consider the difference equation:

$$a_r + 6a_{r-1} + 12a_{r-2} + 8a_{r-3} = 0$$

The characteristic equation is

$$\alpha^3 + 6\alpha^2 + 12\alpha + 8 = 0$$

Thus,

$$a_r^{(h)} = (A_1 r^2 + A_2 r + A_3)(-2)^r$$

is a homogeneous solution since -2 is a triple characteristic root. □

Example 10.3 Consider the difference equation

$$4a_r - 20a_{r-1} + 17a_{r-2} - 4a_{r-3} = 0$$

The characteristic equation is

$$4\alpha^3 - 20\alpha^2 + 17\alpha - 4 = 0$$

and the characteristic roots are $\frac{1}{2}$, $\frac{1}{2}$, 4. Consequently, the homogeneous solution is

$$a_r^{(h)} = (A_1 r + A_2)(\tfrac{1}{2})^r + A_3(4)^r$$ □

10.5 PARTICULAR SOLUTIONS

There is no general procedure for determining the particular solution of a difference equation. However, in simple cases, this solution can be obtained by the method of inspection. As will be demonstrated in the examples in this section, we first set up the general form of the particular solution according to the form of $f(r)$, and then determine the exact solution according to the given difference equation. Consider the difference equation

$$a_r + 5a_{r-1} + 6a_{r-2} = 3r^2 \quad (10.7)$$

We assume that the general form of the particular solution is

$$P_1 r^2 + P_2 r + P_3 \tag{10.8}$$

where P_1, P_2, and P_3 are constants to be determined. Substituting the expression in (10.8) into the left-hand side of (10.7), we obtain

$$P_1 r^2 + P_2 r + P_3 + 5P_1(r-1)^2 + 5P_2(r-1) + 5P_3 + 6P_1(r-2)^2 + 6P_2(r-2) + 6P_3$$

which simplifies to

$$12P_1 r^2 - (34P_1 - 12P_2)r + (29P_1 - 17P_2 + 12P_3) \tag{10.9}$$

Comparing (10.9) with the right-hand side of (10.7), we obtain the equations

$$12P_1 = 3$$
$$34P_1 - 12P_2 = 0$$
$$29P_1 - 17P_2 + 12P_3 = 0$$

which yield

$$P_1 = \tfrac{1}{4} \qquad P_2 = \tfrac{17}{24} \qquad P_3 = \tfrac{115}{288}$$

Therefore, the particular solution is

$$a_r^{(p)} = \tfrac{1}{4}r^2 + \tfrac{17}{24}r + \tfrac{115}{288}$$

In general, when $f(r)$ is of the form of a polynomial of degree t in r

$$F_1 r^t + F_2 r^{t-1} + \cdots + F_t r + F_{t+1}$$

the corresponding particular solution will be of the form

$$P_1 r^t + P_2 r^{t-1} + \cdots + P_t r + P_{t+1}$$

Example 10.4 Consider the difference equation

$$a_r + 5a_{r-1} + 6a_{r-2} = 3r^2 - 2r + 1 \tag{10.10}$$

The particular solution is of the form

$$P_1 r^2 + P_2 r + P_3 \tag{10.11}$$

Substituting (10.11) into (10.10), we obtain

$$P_1 r^2 + P_2 r + P_3 + 5P_1(r-1)^2 + 5P_2(r-1) + 5P_3 + 6P_1(r-2)^2 + 6P_2(r-2) + 6P_3 = 3r^2 - 2r + 1$$

which simplifies to

$$12P_1 r^2 - (34P_1 - 12P_2)r + (29_1 - 17P_2 + 12P_3) = 3r^2 - 2r + 1 \tag{10.12}$$

Comparing the two sides of (10.12), we obtain the equations

$$12P_1 = 3$$
$$34P_1 - 12P_2 = 2$$
$$29P_1 - 17P_2 + 12P_3 = 1$$

which yield

$$P_1 = \tfrac{1}{4} \qquad P_2 = \tfrac{13}{24} \qquad P_3 = \tfrac{71}{288}$$

Therefore, the particular solution is

$$a_r^{(p)} = \tfrac{1}{4}r^2 + \tfrac{13}{24}r + \tfrac{71}{288} \qquad \square$$

Example 10.5 Consider the difference equation

$$a_r - 5a_{r-1} + 6a_{r-2} = 1 \tag{10.13}$$

Since $f(r)$ is a constant, the particular solution will also be a constant P. Substituting P into (10.13), we obtain

$$P - 5P + 6P = 1$$

That is,

$$2P = 1$$

or

$$a_r^{(p)} = \tfrac{1}{2} \qquad \square$$

As another example, consider the difference equation

$$a_r + 5a_{r-1} + 6a_{r-2} = 42 \cdot 4^r \tag{10.14}$$

We assume that the general form of the particular solution is

$$P4^r \tag{10.15}$$

Substituting the expression in (10.15) into the left-hand side of (10.14), we obtain

$$P4^r + 5P4^{r-1} + 6P4^{r-2}$$

which simplifies to

$$\tfrac{21}{8}P4^r \tag{10.16}$$

Comparing (10.16) with the right-hand side of (10.14), we obtain

$$P = 16$$

Therefore, the particular solution is

$$a_r^{(p)} = 16 \cdot 4^r$$

In general, when $f(r)$ is of the form β^r, the corresponding particular solution is of the form $P\beta^r$, if β is not a characteristic root of the difference equation. Furthermore, when $f(r)$ is of the form

$$(F_1 r^t + F_2 r^{t-1} + \cdots + F_t r + F_{t+1})\beta^r$$

the corresponding particular solution is of the form

$$(P_1 r^t + P_2 r^{t-1} + \cdots + P_t r + P_{t+1})\beta^r$$

if β is not a characteristic root of the difference equation. Consider the following example.

Example 10.6 Consider the difference equation

$$a_r + a_{r-1} = 3r2^r \tag{10.17}$$

The general form of the particular solution is

$$(P_1 r + P_2)2^r \tag{10.18}$$

Substituting (10.18) into (10.17), we obtain

$$(P_1 r + P_2)2^r + [P_1(r-1) + P_2]2^{r-1} = 3r2^r$$

which simplifies to

$$\tfrac{3}{2}P_1 r2^r + (-\tfrac{1}{2}P_1 + \tfrac{3}{2}P_2)2^r = 3r2^r \tag{10.19}$$

Comparing the two sides of (10.19), we obtain the equations

$$\tfrac{3}{2}P_1 = 3$$

$$-\tfrac{1}{2}P_1 + \tfrac{3}{2}P_2 = 0$$

Thus,

$$P_1 = 2 \qquad P_2 = \tfrac{2}{3}$$

and the particular solution is

$$a_r^{(p)} = (2r + \tfrac{2}{3})2^r$$

□

For the case that β is a characteristic root of multiplicity $m-1$, when $f(r)$ is of the form

$$(F_1 r^t + F_2 r^{t-1} + \cdots + F_t r + F_{t+1})\beta^r$$

the corresponding particular solution is of the form

$$r^{m-1}(P_1 r^t + P_2 r^{t-1} + \cdots + P_t r + P_{t+1})\beta^r$$

Let us examine the following examples.

Example 10.7 Consider the difference equation

$$a_r - 2a_{r-1} = 3 \cdot 2^r \tag{10.20}$$

Because 2 is a characteristic root (of multiplicity 1), the general form of the particular solution is

$$Pr2^r \tag{10.21}$$

Substituting (10.21) into (10.20), we obtain

$$Pr2^r - 2P(r-1)2^{r-1} = 3 \cdot 2^r$$

that is,

$$P2^r = 3 \cdot 2^r$$

or

$$P = 3$$

Thus, the particular solution is

$$a_r^{(p)} = 3r2^r \qquad \square$$

Example 10.8 For the difference equation

$$a_r - 4a_{r-1} + 4a_{r-2} = (r+1)2^r \tag{10.22}$$

since 2 is a double characteristic root, the general form of the particular solution is

$$r^2(P_1 r + P_2)2^r \tag{10.23}$$

Substituting (10.23) into (10.22), we obtain, after simplification:

$$6P_1 r2^r = r2^r$$

$$(-6P_1 + 2P_2)2^r = 2^r$$

which yield

$$P_1 = \tfrac{1}{6} \qquad P_2 = 1$$

Thus, the particular solution is

$$a_r^{(p)} = r^2\left(\frac{r}{6} + 1\right)2^r \qquad \square$$

Example 10.9 Consider the difference equation

$$a_r = a_{r-1} + 7 \tag{10.24}$$

Since 1 is a characteristic root of the difference equation and 7 can be written as $7 \cdot 1^r$, the general form of the particular solution is Pr. (The reader should find out what happens if we assume the general form of the particular solution to be P instead.) Substituting $a_r^{(p)} = Pr$ into (10.24), we obtain

$$Pr = P(r-1) + 7$$

that is,

$$P = 7 \qquad \square$$

Example 10.10 For the difference equation

$$a_r - 2a_{r-1} + a_{r-2} = 7$$

we let $a_r^{(p)} = Pr^2$. We ask the reader to carry out the substitution to confirm that $P = \frac{7}{2}$. $\square$

Example 10.11 Consider the difference equation

$$a_r - 5a_{r-1} + 6a_{r-2} = 2^r + r$$

The general form of the particular solution is

$$P_1 r 2^r + P_2 r + P_3$$

(Note that 2 is a characteristic root of the difference equation.) Substitution and comparison will yield

$$P_1 = -2 \qquad P_2 = \tfrac{1}{2} \qquad P_3 = \tfrac{7}{4}$$

and

$$a_r^{(p)} = -r2^{r+1} + \tfrac{1}{2}r + \tfrac{7}{4} \qquad \square$$

10.6 TOTAL SOLUTIONS

Finally, we must combine the homogeneous solution and the particular solution and determine the undetermined coefficients in the homogeneous solution. For a kth-order difference equation, the k undetermined coefficients $A_1, A_2, \ldots, A_k$ in the homogeneous solution can be determined by the boundary conditions, a_{r_0}, $a_{r_0+1}, \ldots, a_{r_0+k-1}$, for any r_0. Suppose the characteristic roots of the difference equation are all distinct. The total solution is of the form

$$a_r = A_1\alpha_1^r + A_2\alpha_2^r + \cdots + A_k\alpha_k^r + p(r)$$

where $p(r)$ is the particular solution. Thus, for $r = r_0, r_0 + 1, \ldots, r_0 + k - 1$, we have the system of linear equations:

$$\begin{aligned} a_{r_0} &= A_1\alpha_1^{r_0} + A_2\alpha_2^{r_0} + \cdots + A_k\alpha_k^{r_0} + p(r_0) \\ a_{r_0+1} &= A_1\alpha_1^{r_0+1} + A_2\alpha_2^{r_0+1} + \cdots + A_k\alpha_k^{r_0+1} + p(r_0+1) \qquad (10.25) \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r_0+k-1} &= A_1\alpha_1^{r_0+k-1} + A_2\alpha_2^{r_0+k-1} + \cdots + A_k\alpha_k^{r_0+k-1} + p(r_0+k-1) \end{aligned}$$

These k linear equations can be solved for $A_1, A_2, \ldots, A_k$. For example, for the difference equation in (10.14), the total solution is

$$a_r = A_1(-2)^r + A_2(-3)^r + 16 \cdot 4^r$$

Suppose we are given the boundary conditions $a_2 = 278$ and $a_3 = 962$. Solving the equations

$$278 = 4A_1 + 9A_2 + 256$$

$$962 = -8A_1 - 27A_2 + 1024$$

we obtain

$$A_1 = 1 \qquad A_2 = 2$$

Thus,

$$a_r = (-2)^r + 2(-3)^r + 16 \cdot 4^r$$

is the total solution of the difference equation.

One might question how we can be sure that solutions of the k equations in (10.25) are always unique. It can be shown mathematically that this is indeed the case.† However, in Sec. 10.2, recall that we demonstrated the uniqueness of the solution of a kth-order linear recurrence relation with constant coefficients for any given boundary conditions consisting of k consecutive values $a_{r_0}, a_{r_0+1}, \ldots, a_{r_0+k-1}$. Consequently, the uniqueness of the solution of the recurrence relation guarantees the uniqueness of the solutions of the k equations in (10.25). On the other hand, if we are given the value of the numeric function at k not necessarily consecutive points, although we can set up k equations for the undetermined coefficients $A_1, A_2, \ldots, A_k$ similar to that in (10.25), since the solution of the recurrence relation might not be uniquely specified by such boundary conditions, it is not always the case that these equations can be solved uniquely.

When the characteristic roots of the difference equation are not all distinct, a derivation similar to the foregoing can be carried out. Again, the undetermined coefficients in the homogeneous solution can be determined uniquely by the value of the numeric function at k consecutive points.

10.7 SOLUTION BY THE METHOD OF GENERATING FUNCTIONS

Instead of solving a difference equation for an expression for the value of a numeric function as we did above, we can also determine the generating function of the numeric function from the difference equation. In many cases, once the generating function is determined, an expression for the value of the numeric function can easily be obtained.

Consider the recurrence relation

$$a_r = 3a_{r-1} + 2 \qquad r \geqq 1 \tag{10.26}$$

with the boundary condition $a_0 = 1$. Let us point out that in (10.26) we write down explicitly (for the first time in this chapter) that the recurrence relation is

† See, for example, chap. 3 of Liu [5].

valid only for $r \geqq 1$, a fact we knew implicitly all along. Note that for $r = 0$, the recurrence relation becomes

$$a_0 = 3a_{-1} + 2$$

Since a_{-1} is not defined, we cannot assume that the recurrence relation is valid for $r = 0$. In general, for a given kth-order difference equation that specifies a numeric function, we should know for what values of r the equation is valid. First, we note that the equation is valid only if $r \geqq k$ because, for $r < k$, the equation will involve a_{-i}'s, which are not defined. Furthermore, in many cases, a difference equation arises from a physical problem in which the value of a_i is meaningful only for $i \geqq t$, for some t larger than or equal to 0. In that case, the difference equation is valid only for $r - k \geqq t$.

Multiplying both sides of (10.26) by z^r, we obtain

$$a_r z^r = 3a_{r-1}z^r + 2z^r \qquad r \geqq 1 \tag{10.27}$$

Summing (10.27) for all r, $r \geqq 1$, we obtain

$$\sum_{r=1}^{\infty} a_r z^r = 3 \sum_{r=1}^{\infty} a_{r-1}z^r + 2 \sum_{r=1}^{\infty} z^r$$

Note that

$$\sum_{r=1}^{\infty} a_r z^r = A(z) - a_0$$

$$\sum_{r=1}^{\infty} a_{r-1}z^r = z \sum_{r=1}^{\infty} a_{r-1}z^{r-1} = zA(z)$$

$$\sum_{r=1}^{\infty} z^r = \frac{z}{1-z}$$

We obtain

$$A(z) - a_0 = 3zA(z) + \frac{2z}{1-z}$$

That is,

$$(1 - 3z)A(z) = \frac{2z}{1-z} + 1$$

which simplifies to

$$(1 - 3z)A(z) = \frac{1+z}{1-z}$$

or

$$A(z) = \frac{1+z}{(1-3z)(1-z)}$$

$$= \frac{2}{1-3z} - \frac{1}{1-z}$$

Consequently, we have

$$a_r = 2 \cdot 3^r - 1 \qquad r \geqq 0$$

We state now a general procedure for determining the generating function of the numeric function **a** from the difference equation

$$C_0 a_r + C_1 a_{r-1} + C_2 a_{r-2} + \cdots + C_k a_{r-k} = f(r)$$

which is valid for $r \geqq s$, where $s \geqq k$. Multiplying both sides of this equation by z^r and summing from $r = s$ to $r = \infty$, we obtain

$$\sum_{r=s}^{\infty} (C_0 a_r + C_1 a_{r-1} + C_2 a_{r-2} + \cdots + C_k a_{r-k}) z^r = \sum_{r=s}^{\infty} f(r) z^r$$

Since

$$\sum_{r=s}^{\infty} C_0 a_r z^r = C_0[A(z) - a_0 - a_1 z - a_2 z^2 - \cdots - a_{s-1} z^{s-1}]$$

$$\sum_{r=s}^{\infty} C_1 a_{r-1} z^r = C_1 z[A(z) - a_0 - a_1 z - a_2 z^2 - \cdots - a_{s-2} z^{s-2}]$$

.....

$$\sum_{r=s}^{\infty} C_k a_{r-k} z^r = C_k z^k[A(z) - a_0 - a_1 z - a_2 z^2 - \cdots - a_{s-k-1} z^{s-k-1}]$$

we have

$$\begin{aligned} A(z) = \frac{1}{C_0 + C_1 z + \cdots + C_k z^k} \Bigg[& \sum_{r=s}^{\infty} f(r) z^r \\ & + C_0(a_0 + a_1 z + a_2 z^2 + \cdots + a_{s-1} z^{s-1}) \\ & + C_1 z(a_0 + a_1 z + a_2 z^2 + \cdots + a_{s-2} z^{s-2}) \\ & + \cdots \\ & + C_k z^k(a_0 + a_1 z + a_2 z^2 + \cdots + a_{s-k-1} z^{s-k-1}) \Bigg] \end{aligned}$$

We have more illustrative examples:

Example 10.12 Suppose we toss a coin r times. There are 2^r possible sequences of outcomes. We want to know the number of sequences of outcomes in which heads never appear on successive tosses. Let a_r denote the number of such sequences. To each sequence of $r - 1$ heads and tails in which there are no consecutive heads, we can append a tail to obtain a sequence of r heads and tails in which there are no consecutive heads. To each sequence of $r - 2$ heads and tails in which there are no consecutive heads, we can append a tail and then a head to obtain a sequence of r heads and tails in which there are no consecutive heads. Moreover, these exhaust

all sequences of r heads and tails in which there are no consecutive heads. We thus have the recurrence relation

$$a_r = a_{r-1} + a_{r-2} \tag{10.28}$$

Note that $a_1 = 2$ and $a_2 = 3$. The value of a_0 has no physical significance, and a reasonable choice of its value is 0. In this case, the difference equation is *not* valid for $r = 2$, and is only valid for $r \geqq 3$.

We now proceed to show how we can obtain $A(z)$ from (10.28). Multiplying both sides of (10.28) by z^r and summing from $r = 3$ to $r = \infty$, we obtain

$$\sum_{r=3}^{\infty} a_r z^r = \sum_{r=3}^{\infty} a_{r-1} z^r + \sum_{r=3}^{\infty} a_{r-2} z^r$$

That is,

$$A(z) - a_2 z^2 - a_1 z - a_0 = z[A(z) - a_1 z - a_0] + z^2[A(z) - a_0]$$

which simplifies to

$$A(z) = \frac{z^2 + 2z}{1 - z - z^2} = -1 + \frac{1 + z}{1 - z - z^2}$$

Writing $A(z)$ as

$$A(z) = -1 + \frac{(5 + 3\sqrt{5})/10}{1 - [(1 + \sqrt{5})/2]z} - \frac{(-5 + 3\sqrt{5})/10}{1 - [(1 - \sqrt{5})/2]z}$$

we see that

$$a_r = \begin{cases} 0 & r = 0 \\ \dfrac{5 + 3\sqrt{5}}{10}\left(\dfrac{1 + \sqrt{5}}{2}\right)^r - \dfrac{-5 + 3\sqrt{5}}{10}\left(\dfrac{1 - \sqrt{5}}{2}\right)^r & r \geqq 1 \end{cases}$$

Since the value of a_0 has no physical significance, we can choose to set a_0 to 1 instead of 0, so that the difference equation in (10.28) is valid for $r \geqq 2$. In that case, we can multiply both sides of (10.28) by z^r and sum from $r = 2$ to $r = \infty$. We obtain

$$\sum_{r=2}^{\infty} a_r z^r = \sum_{r=2}^{\infty} a_{r-1} z^r + \sum_{r=2}^{\infty} a_{r-2} z^r$$

That is,

$$A(z) - a_1 z - a_0 = z[A(z) - a_0] + z^2 A(z)$$

or

$$A(z) = \frac{1 + z}{1 - z - z^2}$$

which yields

$$a_r = \frac{5 + 3\sqrt{5}}{10}\left(\frac{1 + \sqrt{5}}{2}\right)^r - \frac{-5 + 3\sqrt{5}}{10}\left(\frac{1 - \sqrt{5}}{2}\right)^r \qquad r \geqq 0 \qquad \square$$

Example 10.13 We shall solve the difference equation in Example 10.11 by determining first the generation function $A(z)$. For the difference equation

$$a_r - 5a_{r-1} + 6a_{r-2} = 2^r + r \qquad r \geqq 2$$

with the boundary conditions $a_0 = 1$ and $a_1 = 1$, since

$$\sum_{r=2}^{\infty} a_r z^r - 5 \sum_{r=2}^{\infty} a_{r-1} z^r + 6 \sum_{r=2}^{\infty} a_{r-2} z^r = \sum_{r=2}^{\infty} 2^r z^r + \sum_{r=2}^{\infty} r z^r$$

we obtain

$$A(z) - a_0 - a_1 z - 5z[A(z) - a_0] + 6z^2 A(z) = \frac{4z^2}{1 - 2z} + z\left[\frac{1}{(1 - z)^2} - 1\right]$$

which simplifies to

$$A(z) = \frac{1 - 8z + 27z^2 - 35z^3 + 14z^4}{(1 - z)^2(1 - 2z)^2(1 - 3z)}$$

$$= \frac{5/4}{1 - z} + \frac{1/2}{(1 - z)^2} - \frac{3}{1 - 2z} - \frac{2}{(1 - 2z)^2} + \frac{17/4}{1 - 3z}$$

Consequently, we have

$$a_r = \frac{5}{4} + \frac{1}{2}(r + 1) - 3 \cdot 2^r - 2(r + 1)2^r + \frac{17}{4} 3^r$$

$$= \frac{7}{4} + \frac{r}{2} - r2^{r+1} - 5 \cdot 2^r + \frac{17}{4} 3^r$$

We strongly suggest that the reader follow the derivation carefully as a review of the relationship between a numeric function and its generating function. $\square$

Example 10.14 Consider a certain nuclear reaction inside of a reactor containing nuclei and high- and low-energy free particles. There are two kinds of events: (1) A high-energy particle strikes a nucleus and is absorbed, causing it to emit three high-energy particles and one low-energy particle; and (2) a low-energy particle strikes a nucleus and is absorbed, causing it to emit two high-energy particles and one low-energy particle. We assume that every free particle causes an event 1 μs after it is emitted. Suppose a single high-energy particle is injected at time 0 into a system containing only nuclei. We want to determine the numbers of high-energy and low-energy particles in the system

at time equal to r μs. Let a_r and b_r denote the number of high-energy and low-energy particles at time r, respectively. We have

$$\begin{aligned} a_r &= 3a_{r-1} + 2b_{r-1} \\ b_r &= a_{r-1} + b_{r-1} \end{aligned} \tag{10.29}$$

Note that these are two simultaneous difference equations for the numeric functions **a** and **b**. The equations are valid for $r \geqq 1$, and we have the boundary conditions

$$a_0 = 1 \qquad b_0 = 0$$

From (10.29) we obtain

$$\sum_{r=1}^{\infty} a_r z^r = \sum_{r=1}^{\infty} 3a_{r-1}z^r + \sum_{r=1}^{\infty} 2b_{r-1}z^r$$

$$\sum_{r=1}^{\infty} b_r z^r = \sum_{r=1}^{\infty} a_{r-1}z^r + \sum_{r=1}^{\infty} b_{r-1}z^r$$

or

$$\begin{aligned} A(z) - 1 &= 3zA(z) + 2zB(z) \\ B(z) &= zA(z) + zB(z) \end{aligned} \tag{10.30}$$

Solving (10.30) for $A(z)$ and $B(z)$, we obtain

$$A(z) = \frac{1-z}{1-4z+z^2} = \frac{(3+\sqrt{3})/6}{1-(2+\sqrt{3})z} + \frac{(3-\sqrt{3})/6}{1-(2-\sqrt{3})z}$$

$$B(z) = \frac{z}{1-4z+z^2} = \frac{\sqrt{3}/6}{1-(2+\sqrt{3})z} - \frac{\sqrt{3}/6}{1-(2-\sqrt{3})z}$$

It follows that

$$a_r = \frac{3+\sqrt{3}}{6}(2+\sqrt{3})^r + \frac{3-\sqrt{3}}{6}(2-\sqrt{3})^r$$

$$b_r = \frac{\sqrt{3}}{6}(2+\sqrt{3})^r - \frac{\sqrt{3}}{6}(2-\sqrt{3})^r \qquad \square$$

We demonstrated in Example 10.14 how we can solve a set of *simultaneous difference equations* by the technique of generating functions. From a set of simultaneous difference equations relating several numeric functions we can obtain a set of simultaneous (algebraic) equations relating their generating functions. These equations can then be solved for closed-form expressions for the generating functions.

10.8 SORTING ALGORITHMS

Let us return to the simple example in Sec. 10.2 in which we showed that the numeric function $(3^0, 3^1, 3^2, \ldots, 3^r, \ldots)$ can be specified by the recurrence relation

$$a_r = 3a_{r-1} \tag{10.31}$$

together with the boundary condition $a_0 = 1$. Equation (10.31) is a *recursive* specification of the value of **a** at r. It says that if a_{r-1} is known, then a_r can be computed as three times a_{r-1}. However, if a_{r-2} is known, then a_{r-1} can be computed as three times a_{r-2}; if a_{r-3} is known, then a_{r-2} can be computed as three times a_{r-3}; and so on. Consequently, knowing a_0 is sufficient to determine a_r for any r. Recalling our discussion in Sec. 10.1, we realize that such a point of view is not restricted to the description of numeric functions. It also suggests a concise and powerful way to describe computing algorithms. We present in this section some examples from the problem of sorting. As we discussed in Sec. 8.2, by sorting the n numbers that are stored in n registers $x_1, x_2, \ldots, x_n$, we mean to rearrange the contents of the registers so that the rearranged contents of registers $x_1, x_2, \ldots, x_n$ are in ascending order. A sorting procedure consists of a sequence of comparison steps, each of which compares the contents of registers x_i and x_j and places the smaller number in x_i and the larger in x_j. We shall denote a comparison step by $A(x_i, x_j)$.

We present first a sorting algorithm from Sec. 8.2; **BUBBLESORT**. However, our description of this algorithm will be different from that in Sec. 8.2 in that we shall show a recursive specification of the algorithm.† We introduce some notations that will enable us to give a "symbolic" specification of the algorithm. Let $S(x_1, x_2, \ldots, x_n)$ denote algorithm **BUBBLESORT** that sorts the numbers in registers $x_1, x_2, \ldots, x_n$ in ascending order, and let $M(x_1, x_2, \ldots, x_n)$ denote the algorithm **LARGEST2** presented in Sec. 8.2 that places in register x_n the largest number among the numbers in registers $x_1, x_2, \ldots, x_n$. We have

$$S(x_1, x_2, \ldots, x_n) \triangleq M(x_1, x_2, \ldots, x_n)S(x_1, x_2, \ldots, x_{n-1}) \tag{10.32}$$

The symbol $\triangleq$ means "is defined to be" and the concatenation of the symbols of procedures means sequential execution of the corresponding procedures from left to right. Note that $M(x_1, x_2, \ldots, x_n)$ can be defined as

$$M(x_1, x_2, \ldots, x_n) \triangleq A(x_1, x_2)A(x_2, x_3) \cdots A(x_{n-2}, x_{n-1})A(x_{n-1}, x_n)$$

which is exactly the specification of algorithm **LARGEST2** in Sec. 8.2. The relation in (10.32) can be written as

$$\begin{aligned} S(x_1, x_2, \ldots, x_n) \triangleq\ & A(x_1, x_2)A(x_2, x_3) \cdots \\ & A(x_{n-2}, x_{n-1})A(x_{n-1}, x_n)S(x_1, x_2, \ldots, x_{n-1}) \end{aligned}$$

† We remind the reader that the bubble sort algorithm presented here is the same as that presented in Sec. 8.2 in that it consists of exactly the same sequence of comparison steps. The only difference is the way the comparison steps are specified. In general, an algorithm can be described in different ways, most notably, *nonrecursively* and *recursively*.

which is a symbolic description of algorithm **BUBBLESORT**. With the boundary condition

$$S(x_1, x_2) \triangleq A(x_1, x_2)$$

our recursive specification of the algorithm is completed.

As an explicit example, we note that

$$\begin{aligned} S(x_1, x_2, x_3, x_4) &\triangleq A(x_1, x_2)A(x_2, x_3)A(x_3, x_4)S(x_1, x_2, x_3) \\ &\triangleq A(x_1, x_2)A(x_2, x_3)A(x_3, x_4)A(x_1, x_2)A(x_2, x_3)S(x_1, x_2) \\ &\triangleq A(x_1, x_2)A(x_2, x_3)A(x_3, x_4)A(x_1, x_2)A(x_2, x_3)A(x_1, x_2) \end{aligned}$$

Let a_n denote the number of comparison steps the bubble sort algorithm takes to sort n numbers. According to (10.32), we have

$$a_n = (n - 1) + a_{n-1} \tag{10.33}$$

Solving the recurrence relation in (10.33) and using the boundary condition $a_1 = 0$, we obtain†

$$a_n = \frac{n(n-1)}{2}$$

We present now another sorting algorithm known as the *Bose-Nelson algorithm*, due to R. C. Bose and R. J. Nelson (see Bose and Nelson [2]), which we shall denote $T(x_1, x_2, \ldots, x_n)$. To simplify our discussion, we assume that the number of numbers to be sorted is a power of 2. Let $P[(x_1, x_2, \ldots, x_n), (x_{n+1}, x_{n+2}, \ldots, x_{2n})]$ denote an algorithm that will arrange the numbers in registers $x_1, x_2, \ldots, x_{2n}$ in ascending order given that the numbers in registers $x_1, x_2, \ldots, x_n$ have already been arranged in ascending order, as have the numbers in registers $x_{n+1}, x_{n+2}, \ldots, x_{2n}$. Clearly,

$$\begin{aligned} T(x_1, x_2, \ldots, x_{2n}) \triangleq\ & T(x_1, x_2, \ldots, x_n)T(x_{n+1}, x_{n+2}, \ldots, x_{2n}) \\ & P[(x_1, x_2, \ldots, x_n), (x_{n+1}, x_{n+2}, \ldots, x_{2n})] \end{aligned} \tag{10.34}$$

In other words, to sort the $2n$ numbers in registers $x_1, x_2, \ldots, x_{2n}$, we can first sort the n numbers in $x_1, x_2, \ldots, x_n$, and the n numbers in $x_{n+1}, x_{n+2}, \ldots, x_{2n}$, and then combine these two sorted sequences into one. The algorithm $P[(x_1, x_2, \ldots, x_n), (x_{n+1}, x_{n+2}, \ldots, x_{2n})]$ can be defined recursively as:

$$\begin{aligned} & P[(x_1, x_2, \ldots, x_n), (x_{n+1}, x_{n+2}, \ldots, x_{2n})] \\ & \quad \triangleq P[(x_1, x_2, \ldots, x_{n/2}), (x_{n+1}, x_{n+2}, \ldots, x_{3n/2})] \\ & \qquad P[(x_{n/2+1}, x_{n/2+2}, \ldots, x_n), (x_{3n/2+1}, x_{3n/2+2}, \ldots, x_{2n})] \\ & \qquad P[(x_{n/2+1}, x_{n/2+2}, \ldots, x_n), (x_{n+1}, x_{n+2}, \ldots, x_{3n/2})] \end{aligned} \tag{10.35}$$

† See Prob. 10.10.

Note that $P[(x_1, x_2, \ldots, x_{n/2}), (x_{n+1}, x_{n+2}, \ldots, x_{3n/2})]$ places the smallest $n/2$ numbers, among the $2n$ numbers in the registers, in registers $x_1, x_2, \ldots, x_{n/2}$ in ascending order. Similarly, $P[(x_{n/2+1}, x_{n/2+2}, \ldots, x_{2n}), (x_{3n/2+1}, x_{3n/2+2}, \ldots, x_{2n})]$ places the largest $n/2$ numbers, among the $2n$ numbers in the registers, in registers $x_{3n/2+1}, x_{3n/2+2}, \ldots, x_{2n}$ in ascending order. Finally, $P[(x_{n/2+1}, x_{n/2+2}, \ldots, x_n), (x_{n+1}, x_{n+2}, \ldots, x_{3n/2})]$ places the remaining n of the $2n$ numbers in registers $x_{n/2+1}, x_{n/2+2}, \ldots, x_{3n/2}$ in ascending order. Applying the relations in (10.34) and (10.35) repeatedly, we can express $T(x_1, x_2, \ldots, x_n)$, $n = 2^r$, as a sequence of procedures of the forms $T(x_i, x_j)$ and $P[(x_i), (x_j)]$. However, both of $T(x_i, x_j)$ and $P[(x_i), (x_j)]$ are equal to $A(x_i, x_j)$.

For example, we note that

$$T(x_1, x_2, x_3, x_4) \triangleq T(x_1, x_2)T(x_3, x_4)P[(x_1, x_2), (x_3, x_4)]$$

Since

$$T(x_1, x_2) \triangleq A(x_1, x_2)$$

$$T(x_3, x_4) \triangleq A(x_3, x_4)$$

and

$$\begin{aligned} P[(x_1, x_2), (x_3, x_4)] &\triangleq P[(x_1), (x_3)]P[(x_2), (x_4)]P[(x_2), (x_3)] \\ &\triangleq A(x_1, x_3)A(x_2, x_4)A(x_2, x_3) \end{aligned}$$

we have

$$T(x_1, x_2, x_3, x_4) \triangleq A(x_1, x_2)A(x_3, x_4)A(x_1, x_3)A(x_2, x_4)A(x_2, x_3)$$

as a complete specification of an algorithm that sorts four numbers stored in the registers x_1, x_2, x_3, x_4.

Let c_r denote the number of comparison steps that $P[(x_1, x_2, \ldots, x_{2^r}), (x_{2^r+1}, x_{2^r+2}, \ldots, x_{2^{r+1}})]$ takes. According to the relation in (10.35), we have

$$c_r = 3c_{r-1} \tag{10.36}$$

Solving (10.36) with the boundary condition $c_0 = 1$, we obtain

$$c_r = 3^r$$

Let d_r denote the number of comparison steps the procedure $T(x_1, x_2, \ldots, x_{2^r})$ takes. According to the relation in (10.34), we have

$$d_r = 2d_{r-1} + c_{r-1}$$

or

$$d_r = 2d_{r-1} + 3^{r-1} \tag{10.37}$$

Solving (10.37), we obtain

$$d_r = B2^r + 3^r$$

From the boundary condition $d_0 = 0$, we determine that $B = -1$. Thus,

$$d_r = 3^r - 2^r$$

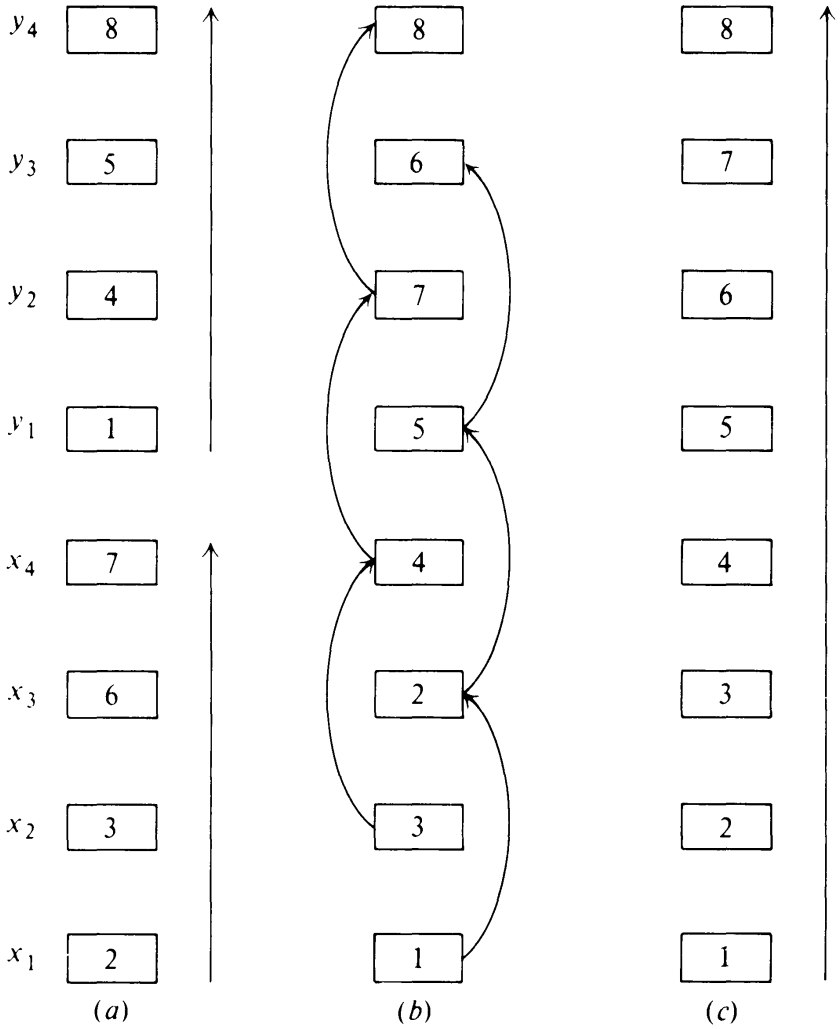

Figure 10.1

Another algorithm is known as the *odd-even merge*, due to K. E. Batcher (see Batcher [1]), which we shall denote $U(x_1, x_2, \ldots, x_n)$. Again, we assume that n is a power of 2. Let

$$U(x_1, x_2, \ldots, x_{2n}) \triangleq U(x_1, x_2, \ldots, x_n)U(x_{n+1}, x_{n+2}, \ldots, x_{2n})$$
$$B[(x_1, x_2, \ldots, x_n), (x_{n+1}, x_{n+2}, \ldots, x_{2n})] \tag{10.38}$$

where the algorithm B is defined to be

$$\begin{aligned} &B[(x_1, x_2, x_3, x_4, \ldots, x_{2n}), (y_1, y_2, y_3, y_4, \ldots, y_{2n})] \\ &\quad \triangleq B[(x_1, x_3, x_5, \ldots, x_{2n-1}), (y_1, y_3, y_5, \ldots, y_{2n-1})] \\ &\qquad B[(x_2, x_4, x_6, \ldots, x_{2n}), (y_2, y_4, y_6, \ldots, y_{2n})] \\ &\qquad A(x_2, x_3)A(x_4, x_5)A(x_6, x_7) \cdots A(x_{2n}, y_1) \\ &\qquad A(y_2, y_3)A(y_4, y_5) \cdots A(y_{2n-2}, y_{2n-1}) \end{aligned} \tag{10.39}$$

It is not immediately obvious that the algorithm defined in (10.39) indeed merges two sorted sequences of numbers into one. We present first an illustrative example before we proceed to verify the validity of the algorithm. Suppose eight numbers, 2, 3, 6, 7, 1, 4, 5, 8, are stored in registers $x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4$, as shown in Fig. 10.1*a*, such that the numbers in registers x_1, x_2, x_3, x_4 are in

ascending order, and the numbers in registers y_1, y_2, y_3, y_4 are in ascending order. According to (10.39),

$$B[(x_1, x_2, x_3, x_4), (y_1, y_2, y_3, y_4)]$$
$$\triangleq B[(x_1, x_3), (y_1, y_3)]B[x_2, x_4), (y_2, y_4)]A(x_2, x_3)A(x_4, y_1)A(y_2, y_3)$$

$B[(x_1, x_3), (y_1, y_3)]$ arranges the numbers in x_1, x_3, y_1, y_3 in ascending order, and $B[(x_2, x_4), (y_2, y_4)]$ arranges the numbers in x_2, x_4, y_2, y_4 in ascending order as shown in Fig. 10.1*b*. Finally, $A(x_2, x_3)$ arranges the numbers in x_2, x_3; $A(x_4, y_1)$ arranges the numbers in x_4, y_1; and $A(y_2, y_3)$ arranges the numbers in y_2, y_3 in ascending order as shown in Fig. 10.1*c*.

To verify the validity of (10.39), we want first to show that after the execution of $B[(x_1, x_3, x_5, \ldots, x_{2n-1}), (y_1, y_3, y_5, \ldots, y_{2n-1})]$ and $B[(x_2, x_4, x_6, \ldots, x_{2n}), (y_2, y_4, y_6, \ldots, y_{2n})]$:

1. The number in x_1 is the smallest of the $4n$ numbers.
2. The number in y_{2n} is the largest of the $4n$ numbers.
3. The two numbers in x_{2i} and x_{2i+1}, $1 \leqq i \leqq n-1$, are the $2i$th and the $(2i+1)$st smallest of the $4n$ numbers; the two numbers in x_{2n} and y_1 are the $2n$th and the $(2n+1)$st smallest of the $4n$ numbers; and the two numbers in y_{2i} and y_{2i+1}, $1 \leqq i \leqq n-1$, are the $(2n+2i)$th and $(2n+2i+1)$st smallest of the $4n$ numbers.

If this is indeed the case, the comparisons $A(x_2, x_3)A(x_4, x_5) \cdots A(y_{2n-2}, y_{2n-1})$ will yield a sorted sequence of $4n$ numbers.

The reader can easily ascertain that (1) and (2) are obvious. To show that (3) is true, we shall only show that the two numbers in x_{2i} and x_{2i+1}, $1 \leqq i \leqq n-1$, are the $2i$th and the $(2i+1)$st smallest of the $4n$ numbers. The other two cases can be shown in a similar manner.† Consider the number that is in x_{2i} after the execution. This number is larger than exactly $i-1$ of the numbers that were in $x_2, x_4, x_6, \ldots, x_{2n}, y_2, y_4, y_6, \ldots, y_{2n}$. Assume that this number is larger than k of the numbers that were in $x_2, x_4, x_6, \ldots, x_{2n}$, and is larger than $i-k-1$ of the numbers that were in $y_2, y_4, y_6, \ldots, y_{2n}$. That is, this number is larger than the numbers that were in $x_2, x_4, x_6, \ldots, x_{2k}$ and those that were in $y_2, y_4, y_6, \ldots, y_{2(i-k-1)}$. In other words, this number is larger than the numbers that were in $x_1, x_2, x_3, x_4, \ldots, x_{2k}$ and those that were in $y_1, y_2, y_3, y_4, \ldots, y_{2(i-k-1)}$. Now, if this number was in $x_{2(k+1)}$, it is also larger than x_{2k+1}, and it might or might not be larger than the number that was in $y_{2(i-k-1)+1}$. On the other hand, if this number was in $y_{2(i-k)}$, it is also larger than $y_{2(i-k)-1}$, and it might or might not be larger than the number that was in x_{2k+1}. Consequently, we can conclude that this number is larger than $2i-1$ or $2i$ of the $4n$ numbers in the registers.

In a similar way, let us consider the number that is in x_{2i+1} after the execu-

† We do not really have three cases in (3). It is only because of the notation we have chosen that we have to make three parallel statements.

tion. Assume that this number is larger than k of the numbers that were in $x_1, x_3, \ldots, x_{2n-1}$, and is larger than $i - k$ of the numbers that were in $y_1, y_3, \ldots, y_{2n-1}$. That is, this number is larger than the numbers that were in $x_1, x_3, \ldots, x_{2k-1}$, and is larger than the numbers that were in $y_1, y_3, \ldots, y_{2(i-k)-1}$. Thus, this number is larger than the numbers that were in $x_1, x_2, x_3, \ldots, x_{2k-2}, x_{2k-1}$, and those that were in $y_1, y_2, y_3, \ldots, y_{2(i-k)-2}, y_{2(i-k)-1}$. Now, if this number was in x_{2k+1}, it is also larger than the number that was in x_{2k}, and it might or might not be larger than the number that was in $y_{2(i-k)}$. If this number was in $y_{2(i-k)+1}$, it is larger than the number that was in $y_{2(i-k)}$, and it might or might not be larger than the number that was in x_{2k}. Consequently, we can conclude that this number is larger than $2i - 1$ or $2i$ of the $4n$ numbers in the registers.

We now complete the proof of (3) by noting that if the number in x_{2i} is the $2i$th smallest number, then the number in x_{2i+1} must be the $(2i + 1)$st smallest number, and vice versa.

Let c_r denote the number of comparison steps the algorithm $B[(x_1, x_2, \ldots, x_{2^r}), (y_1, y_2, \ldots, y_{2^r})]$ takes. According to the relation in (10.39), we have

$$c_r = 2c_{r-1} + (2^r - 1) \tag{10.40}$$

Solving (10.40), with the boundary condition $c_0 = 1$, we obtain

$$c_r = r2^r + 1$$

Let d_r denote the number of comparison steps the algorithm $U(x_1, x_2, \ldots, x_{2^r})$ takes. According to the relation in (10.38), we have

$$d_r = 2d_{r-1} + c_{r-1}$$

or

$$d_r = 2d_{r-1} + (r - 1)2^{r-1} + 1 \tag{10.41}$$

Solving the recurrence relation in (10.41), with the boundary condition $d_0 = 0$, we obtain

$$d_r = (r^2 - r + 4)2^{r-2} - 1$$

*10.9 MATRIX MULTIPLICATION ALGORITHMS

As another illustration of recursive algorithms, we consider a problem on matrix multiplication. Let **A** and **B** be two 2×2 matrices.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

The product **A** and **B**, denoted **C**, can be computed as

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

where

$$\begin{aligned} c_{11} &= a_{11}b_{11} + a_{12}b_{21} \\ c_{12} &= a_{11}b_{12} + a_{12}b_{22} \\ c_{21} &= a_{21}b_{11} + a_{22}b_{21} \\ c_{22} &= a_{21}b_{12} + a_{22}b_{22} \end{aligned}$$

We note immediately that such a computation requires four addition operations and eight multiplication operations. Let d_1 denote the cost of an addition operation, and d_2 denote the cost of a multiplication operation. The total cost of multiplying two 2×2 matrices is

$$4d_1 + 8d_2$$

Once again, one may ask if there is a "better" algorithm to multiply two 2×2 matrices? As it turns out, an alternative algorithm can do so. (Later, we shall determine if this alternative algorithm is indeed a better algorithm.) The alternative algorithm computes the first seven intermediate results:

$$\begin{aligned} m_1 &= (a_{11} + a_{22})(b_{11} + b_{22}) \\ m_2 &= (a_{12} - a_{22})(b_{21} + b_{22}) \\ m_3 &= (a_{11} - a_{21})(b_{11} + b_{12}) \\ m_4 &= (a_{11} + a_{12})b_{22} \\ m_5 &= (a_{21} + a_{22})b_{11} \\ m_6 &= a_{11}(b_{12} - b_{22}) \\ m_7 &= a_{22}(b_{21} - b_{11}) \end{aligned} \tag{10.42}$$

These intermediate results can then be used to compute the entries of the matrix **C**:

$$\begin{aligned} c_{11} &= m_1 + m_2 - m_4 + m_7 \\ c_{12} &= m_4 + m_6 \\ c_{21} &= m_5 + m_7 \\ c_{22} &= m_1 - m_3 - m_5 + m_6 \end{aligned}$$

A quick count indicates that such an algorithm requires 18 addition operations and 7 multiplication operations. Thus, the total cost is

$$18d_1 + 7d_2$$

It is far from obvious that such an algorithm is superior to the first algorithm, which is the classical matrix multiplication algorithm. As a matter of fact, one would immediately point out that the first algorithm requires 12 arithmetic operations, while the second one requires 25 arithmetic operations. On the other hand, an advocate of the second algorithm might argue that the second algorithm will be less expensive than the first if a multiplication operation is significantly more expensive than an addition operation. One might point out that such an argument is unrealistic because, in practice, costs for multiplication and addition are about the same. However, before we conclude prematurely that this is an empty exercise in algorithm design, let us consider the more general problem of multiplying two $n \times n$ matrices **A** and **B**. To simplify our discussion, we shall assume that n is a power of 2. One recalls that computing the product of two $n \times n$ matrices by the classical multiplication algorithm requires n^3 multiplication operations and $n^2(n-1)$ addition operations, that is, a total of $2n^3 - n^2$ arithmetic operations. (The ijth entry of the product is computed as $\sum_{k=1}^{n} a_{ik} b_{kj}$, which requires n multiplication operations and $(n-1)$ addition operations.) We recognize that the second algorithm presented above can be used to multiply two $n \times n$ matrices. Given two $n \times n$ matrices **A** and **B**, we can partition them into $(n/2) \times (n/2)$ submatrices:

$$\mathbf{A} = \left[\begin{array}{c:c} \mathbf{a}_{11} & \mathbf{a}_{12} \\ \hdashline \mathbf{a}_{21} & \mathbf{a}_{22} \end{array}\right]$$

$$\mathbf{B} = \left[\begin{array}{c:c} \mathbf{b}_{11} & \mathbf{b}_{12} \\ \hdashline \mathbf{b}_{21} & \mathbf{b}_{22} \end{array}\right]$$

The product **AB** can then be expressed as

$$\mathbf{C} = \mathbf{AB} = \left[\begin{array}{c:c} \mathbf{c}_{11} & \mathbf{c}_{12} \\ \hdashline \mathbf{c}_{21} & \mathbf{c}_{22} \end{array}\right]$$

where $\mathbf{a}_{11}$, $\mathbf{a}_{12}$, $\mathbf{a}_{21}$, $\mathbf{a}_{22}$, $\mathbf{b}_{11}$, $\mathbf{b}_{12}$, $\mathbf{b}_{21}$, $\mathbf{b}_{22}$, $\mathbf{c}_{11}$, $\mathbf{c}_{12}$, $\mathbf{c}_{21}$, and $\mathbf{c}_{22}$† are all $(n/2) \times (n/2)$ submatrices. We can then compute

$$\mathbf{m}_1 = (\mathbf{a}_{11} + \mathbf{a}_{22})(\mathbf{b}_{11} + \mathbf{b}_{22})$$

$$\mathbf{m}_2 = (\mathbf{a}_{12} - \mathbf{a}_{22})(\mathbf{b}_{21} + \mathbf{b}_{22})$$

$$\mathbf{m}_3 = (\mathbf{a}_{11} - \mathbf{a}_{21})(\mathbf{b}_{11} + \mathbf{b}_{12})$$

$$\mathbf{m}_4 = (\mathbf{a}_{11} + \mathbf{a}_{12})\mathbf{b}_{22}$$

$$\mathbf{m}_5 = (\mathbf{a}_{21} + \mathbf{a}_{22})\mathbf{b}_{11}$$

$$\mathbf{m}_6 = \mathbf{a}_{11}(\mathbf{b}_{12} - \mathbf{b}_{22})$$

$$\mathbf{m}_7 = \mathbf{a}_{22}(\mathbf{b}_{21} - \mathbf{b}_{11})$$

† We use boldface type to indicate that these are matrices.

Note that these are repetitions of the relations in (10.42). We merely replaced the roman letters by boldface ones to emphasize that $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_7, \mathbf{a}_{11}, \ldots, \mathbf{a}_{22}, \mathbf{b}_{11}, \ldots, \mathbf{b}_{22}$ are all matrices. Of course, in these equations all multiplications are multiplications of $(n/2) \times (n/2)$ matrices, and all additions are additions of $(n/2) \times (n/2)$ matrices. Finally, we can compute

$$\mathbf{c}_{11} = \mathbf{m}_1 + \mathbf{m}_2 - \mathbf{m}_4 + \mathbf{m}_7$$

$$\mathbf{c}_{12} = \mathbf{m}_4 + \mathbf{m}_6$$

$$\mathbf{c}_{21} = \mathbf{m}_5 + \mathbf{m}_7$$

$$\mathbf{c}_{22} = \mathbf{m}_1 - \mathbf{m}_3 - \mathbf{m}_5 + \mathbf{m}_6$$

We note that the total cost of multiplying **A** and **B** is 7 matrix multiplications and 18 matrix additions (of $(n/2) \times (n/2)$ matrices). If we use the classical method to carry out multiplication of $(n/2) \times (n/2)$ matrices, the total number of arithmetic operations will be

$$7 \times \left[2\left(\frac{n}{2}\right)^3 - \left(\frac{n}{2}\right)^2\right] + 18 \times \left(\frac{n}{2}\right)^2 = \frac{7}{4}n^3 + \frac{11}{4}n^2$$

which is superior to $2n^3 - n^2$ for large n. However, we can do better than that by employing our new multiplication algorithm recursively to compute the products of the submatrices. In other words, to compute the product of two $n \times n$ matrices, we need to perform 7 multiplications and 18 additions of $(n/2) \times (n/2)$ matrices, while each of the multiplication operations of $(n/2) \times (n/2)$ matrices, can be performed with seven multiplications and 18 additions of $(n/4) \times (n/4)$ matrices, and so on. Let f_r denote the total number of arithmetic operations required to multiply two $2^r \times 2^r$ matrices. We have

$$f_r = 7f_{r-1} + 18 \cdot 2^{2r-2} \qquad r \geqq 1 \tag{10.43}$$

with the boundary condition $f_0 = 1$. Solving (10.43), we obtain

$$f_r = 7 \cdot 7^r - \tfrac{18}{3} \cdot 4^r$$

Thus, for $n = 2^r$,

$$f_r = 7 \cdot 7^{\lg n} - \tfrac{18}{3} \cdot n^2 = 7n^{2.81} - \tfrac{18}{3} \cdot n^2$$

Since the time complexity of the algorithm presented above is $\Theta(n^{2.81})$, the time complexity of the matrix multiplication problem is $O(n^{2.81})$.

10.10 REMARKS AND REFERENCES

See Levy and Lessman [4] on the subject of finite-difference equations. See also chaps. 2 and 3 of Liu and Liu [6]. The sorting algorithms discussed in Sec. 10.8 have two distinct features: (1) They are *nonadaptive* in the sense that the sequence of comparison steps in an algorithm is predetermined and does not vary in

accordance with the outcome of any particular comparison step. (2) No additional registers for storing intermediate results is needed. We note that in a nonadaptive algorithm, some of the comparison steps can be carried out simultaneously. [For example, $A(x_1, x_2)$ and $A(x_3, x_4)$ can be carried out simultaneously while $A(x_1, x_2)$ and $A(x_1, x_3)$ cannot be.] Consequently, there is the possibility of speeding up the computation by parallel processing. The advantage of algorithms that do not use additional storage registers is obvious when the number of items to be sorted is large. For a most complete discussion on sorting algorithms and related topics, see chap. 5 of Knuth [3].

The matrix multiplication algorithm presented in Sec. 10.9 is due to Strassen [8]. See Prob. 10.38 for an algorithm that uses 7 multiplication operations and 15 addition operations. Note, however, that the time complexity of such an algorithm would still be $\Theta(n^{2.81})$. To reduce the exponent of n from 2.81 to a smaller number, one's immediate reaction would be to search for a multiplication algorithm for 3×3 matrices that uses 21 or fewer multiplication operations. However, no such algorithm has so far been discovered. Surprisingly, multiplication algorithms for larger matrices that lead to improvement on the exponent of n have been discovered. For example, Pan [7] shows that we can multiply two 48×48 matrices using 47,216 multiplication operations, which reduces the exponent to 2.78.

1. Batcher, K. E.: Sorting Networks and Their Applications, *AFIPS Proc. of the 1968 SJCS*, **32**: 307–314 (1968).
2. Bose, R. C., and R. J. Nelson: A Sorting Problem, *Journal of ACM*, **9**: 282–296 (1962).
3. Knuth, D. E.: "The Art of Computer Programming, Vol. 3, Sorting and Searching," Addison-Wesley Publishing Company, Reading, Mass., 1973.
4. Levy, H., and F. Lessman: "Finite Difference Equations," The Macmillan Company, New York, 1961.
5. Liu, C. L.: "Introduction to Combinatorial Mathematics," McGraw-Hill Book Company, New York, 1968.
6. Liu, C. L., and J. W. S. Liu: "Linear Systems Analysis," McGraw-Hill Book Company, New York, 1975.
7. Pan, V. Ya.: Strassen's Algorithm Is Not Optimal, *Proc. 19th Annual Symposium on Foundations of Computer Science*, 166–176 (1978).
8. Strassen, V.: Gaussian Elimination Is Not Optimal, *Numerische Mathematik*, **13**: 354–356 (1969).

PROBLEMS

10.1 Solve the following recurrence relations:

(*a*) $a_r - 7a_{r-1} + 10a_{r-2} = 0$, given that $a_0 = 0$ and $a_1 = 3$.

(*b*) $a_r - 4a_{r-1} + 4a_{r-2} = 0$, given that $a_0 = 1$ and $a_1 = 6$.

10.2 Solve the following recurrence relations:

(*a*) $a_r - 7a_{r-1} + 10a_{r-2} = 3^r$, given that $a_0 = 0$ and $a_1 = 1$.

(*b*) $a_r + 6a_{r-1} + 9a_{r-2} = 3$, given that $a_0 = 0$ and $a_1 = 1$.

(*c*) $a_r + a_{r-1} + a_{r-2} = 0$, given that $a_0 = 0$ and $a_1 = 2$.

10.3 Solve the following recurrence relations:

(*a*) $a_r - a_{r-1} - a_{r-2} = 0$, given that $a_0 = 1$ and $a_1 = 1$.

(*b*) $a_r - 2a_{r-1} + 2a_{r-2} - a_{r-3} = 0$, given that $a_0 = 2$, $a_1 = 1$, and $a_2 = 1$.

10.4 Given that $a_0 = 0$, $a_1 = 1$, $a_2 = 4$, and $a_3 = 12$ satisfy the recurrence relation

$$a_r + C_1 a_{r-1} + C_2 a_{r-2} = 0$$

determine a_r.

10.5 The solution of the recurrence relation

$$C_0 a_r + C_1 a_{r-1} + C_2 a_{r-2} = f(r)$$

is

$$3^r + 4^r + 2$$

Given that $f(r) = 6$ for all r, determine C_0, C_1, and C_2.

10.6 The solution of the recurrence relation

$$a_r = Aa_{r-1} + B3^r \qquad r \geqq 1$$

is

$$a_r = C2^r + D3^{r+1} \qquad r \geqq 0$$

Given that $a_0 = 19$ and $a_1 = 50$, determine the constants A, B, C, and D.

10.7 Let

$$4a_r + C_1 a_{r-1} + C_2 a_{r-2} = f(r) \qquad r \geqq 2$$

be a second-order linear recurrence with constant coefficients. For some boundary conditions a_0 and a_1, the solution of the recurrence is

$$1 - 2r + 3 \cdot 2^r$$

Determine a_0, a_1, C_1, C_2, and $f(r)$. (The solution is not unique.)

10.8 Consider the recurrence relation

$$a_r = a_{r-1} - a_{r-2}$$

(*a*) Solve the recurrence relation, given that $a_1 = 1$ and $a_2 = 0$.
(*b*) Can you solve the recurrence relation if it is given that $a_0 = 0$ and $a_3 = 0$?
(*c*) Repeat part (*b*) if it is given that $a_0 = 1$ and $a_3 = 2$.

10.9 (*a*) Determine the particular solution for the difference equation

$$a_r - 3a_{r-1} + 2a_{r-2} = 2^r$$

(*b*) Determine the particular solution for the difference equation

$$a_r - 4a_{r-1} + 4a_{r-2} = 2^r$$

10.10 (*a*) Determine the particular solution for the difference equation

$$a_r - 2a_{r-1} = f(r)$$

where $f(r) = 7r$.

(*b*) Repeat part (*a*) for $f(r) = 7r^2$.
(*c*) Determine the particular solution for the difference equation

$$a_r - a_{r-1} = 7r$$

(*d*) Repeat part (*c*) if $f(r) = 7r^2$.
(*e*) Let

$$C_0 a_r + C_1 a_{r-1} + \cdots + C_k a_{r-k} = f(r)$$

be a difference equation with a characteristic root 1. Let $f(r) = r^t$. What can be said about the general form of the particular solution $a_r^{(p)}$?

10.11 (*a*) Solve the recurrence relation

$$a_r + 3a_{r-1} + 2a_{r-2} = f(r)$$

where

$$f(r) = \begin{cases} 1 & r = 2 \\ 0 & \text{otherwise} \end{cases}$$

with the boundary condition $a_0 = a_1 = 0$.

(*b*) Repeat part (*a*) for

$$f(r) = \begin{cases} 1 & r = 5 \\ 0 & \text{otherwise} \end{cases}$$

(*c*) Consider the recurrence relation

$$C_0 a_r + C_1 a_{r-1} + C_2 a_{r-2} + \cdots + C_k a_{r-k} = f(r)$$

Let $\hat{a}_r$ denote the solution of the recurrence relation for $f(r) = \hat{f}(r)$ with the boundary conditions $\hat{a}_0 = \hat{a}_1 = \hat{a}_2 = \cdots = \hat{a}_{k-1} = 0$. Let $\tilde{a}_r$ denote the solution of the recurrence relation for $f(r) = \tilde{f}(r)$ with the boundary conditions $\tilde{a}_0 = \tilde{a}_1 = \tilde{a}_2 = \cdots = \tilde{a}_{k-1} = 0$. Given that $\hat{f}(r) = 0$ for $r < k$, and

$$\tilde{f}(r) = \begin{cases} 0 & 0 \leqq r \leqq l - 1 \\ \hat{f}(r - l) & r \geqq l \end{cases}$$

for some fixed l, what can we conclude about $\hat{a}_r$ and $\tilde{a}_r$?

10.12 (*a*) Consider the recurrence relation

$$C_0 a_r + C_1 a_{r-1} + C_2 a_{r-2} + \cdots + C_k a_{r-k} = f(r)$$

Let $\hat{a}_r$ denote the solution of the recurrence relation for $f(r) = \hat{f}(r)$ with the boundary conditions $\hat{a}_0 = \hat{a}_1 = \hat{a}_2 = \cdots = \hat{a}_{k-1} = 0$. Let $\tilde{a}_r$ denote the solution of the recurrence relation for $f(r) = \tilde{f}(r)$ with the boundary conditions $\tilde{a}_0 = \tilde{a}_1 = \tilde{a}_2 = \cdots = \tilde{a}_{k-1} = 0$. Show that $\bar{a}_r = \hat{a}_r + \tilde{a}_r$ is the solution of the recurrence relation for $f(r) = \hat{f}(r) + \tilde{f}(r)$ with the boundary conditions $\bar{a}_0 = \bar{a}_1 = \bar{a}_2 = \cdots = \bar{a}_{k-1} = 0$, provided that $\hat{f}(r) = \tilde{f}(r) = 0$ for $r < k$.

(*b*) Solve the recurrence equation

$$a_r + 5a_{r-1} + 6a_{r-2} = f(r)$$

where

$$f(r) = \begin{cases} 0 & r = 0, 1, 5 \\ 6 & \text{otherwise} \end{cases}$$

given that $a_0 = a_1 = 0$.

10.13 Gossip is spread among r people via telephone. Specifically, in a telephone conversation between A and B, A tells B all the gossip he has heard, and B reciprocates. Let a_r denote the minimum number of telephone calls among r people so that all gossip will be known to everyone.

(*a*) Show that $a_2 = 1$, $a_3 = 3$, and $a_4 = 4$.

(*b*) Show that

$$a_r \leqq a_{r-1} + 2$$

(*c*) Show that

$$a_r \leqq 2r - 4 \qquad \text{for } r \geqq 4$$

[Indeed, it can be shown that $a_r = 2r - 4$. See B. Baker and R. Shostak, Gossips and Telephones. *Discrete Mathematics*, **2**: 191–193, (1972).]

10.14 Let a_r denote the number of partitions of a set of r elements. Show that

$$a_{r+1} = \sum_{i=0}^{r} \binom{r}{i} a_i$$

where $a_0 = 1$.

10.15 Consider an air-traffic-control system in which the desired altitude of an aircraft, a_r, is computed by a computer every second and is compared with the actual altitude of the aircraft, b_{r-1}, determined by a tracking radar 1 second earlier. Depending on whether a_r is larger or smaller than b_{r-1}, the altitude of the aircraft will be changed accordingly. Specifically, the change in altitude at the rth second, $b_r - b_{r-1}$, is proportional to the difference $a_r - b_{r-1}$. That is,

$$b_r - b_{r-1} = K(a_r - b_{r-1})$$

where K is a proportional constant.

(*a*) Determine b_r, given that $a_r = 1000\left(\frac{3}{2}\right)^2$, $K = 3$, and $b_0 = 0$.

(*b*) Determine b_r, given that

$$a_r = \begin{cases} 1000\left(\frac{3}{2}\right)^r & 0 \leqq r \leqq 9 \\ 1000\left(\frac{3}{2}\right)^{10} & r \geqq 10 \end{cases}$$

$K = 3$, and $b_0 = 0$.

10.16 *The Tower of Hanoi problem.* r circular rings of tapering sizes are slipped onto a peg with the largest ring at the bottom, as shown in Fig. 10P.1. These rings are to be transferred one at a time onto another peg, and there is a third peg available on which rings can be left temporarily. If, during the course of transferring the rings, no ring may ever be placed on top of a smaller one, in how many moves can these rings be transferred with their relative positions unchanged?

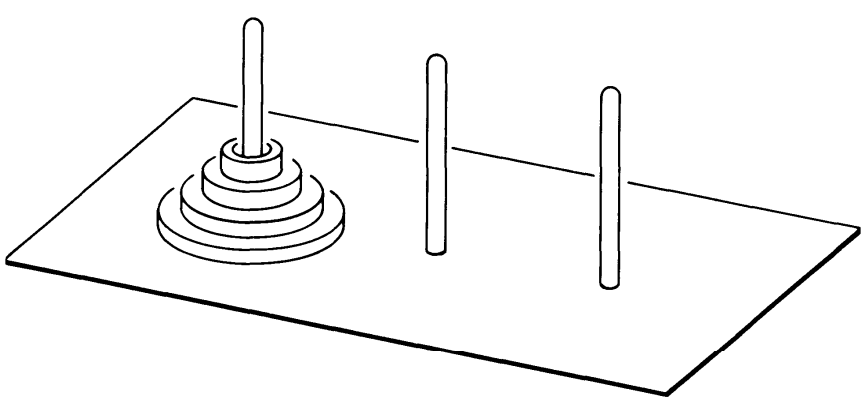

Figure 10P.1

10.17 Consider the multiplication of bacteria in a controlled environment. Let a_r denote the number of bacteria there are on the rth day. We define the rate of growth on the rth day to be $a_r - 2a_{r-1}$. If it is known that the rate of growth doubles every day, determine a_r, given that $a_0 = 1$.

10.18 Consider the operation of a factory whose average profit in every two successive months is equal to the average new order in that period. Let a_r denote the new order received and b_r denote the monthly profit in the rth month.

(*a*) Given that $a_r = 2^r$ for all $r \geqq 0$ and $b_0 = 0$, determine b_r.

(*b*) Repeat part (*a*) for

$$a_r = \begin{cases} 2^r & 0 \leqq r \leqq 9 \\ 2^{10} & r \geqq 10 \end{cases}$$

10.19 A particle is moving in the horizontal direction. The distance it travels in each second is equal to two times the distance it traveled in the previous second. Let a_r denote the position of the particle in the rth second. Determine a_r, given that $a_0 = 3$ and $a_3 = 10$.

10.20 Consider the following algorithm for sorting r numbers for $r \geqq 2$.

(*a*) Use $2r - 3$ comparisons to determine the largest and second largest of the r numbers.

(*b*) Recursively, sort the remaining $r - 2$ numbers.

Let a_r denote the number of comparisons used for sorting r numbers. Determine a_r.

10.21 Let a_r denote the number of edges in a complete graph on r vertices.

(*a*) Derive a recurrence relation for a_r in terms of a_{r-1}.

(*b*) Solve the recurrence relation.

10.22 Let a_r be the number of subsets of the set $\{1, 2, \ldots, r\}$ that do not contain two consecutive numbers. Determine a_r.

Hint: Among the a_r subsets, how many of them do not contain the number r? How many of them contain the number r?

10.23 Consider the problem of covering a rectangular strip of length n with two types of domino. A blue domino has length 2, while an orange domino has length 1. Assuming you have sufficient dominoes of each type, how many different ways can the strip be covered?

10.24 How many r-digit binary sequences that have no adjacent 0s are there?

10.25 Let a_r denote the total assets of a bank at the end of the rth month which is equal to the sum of the total deposit in the rth month and 1.1 times the total assets at the end of the $(r - 1)$st month. Given that the total deposit is a constant 100 (in thousands of dollars), determine a_r if $a_0 = 0$.

10.26 Let a_r denote the total dollar assets of a company in the rth year. Clearly, $a_r - a_{r-1}$ is the increase in assets during the rth year. If the increase in assets during each year is always five times the increase during the previous year, what are the total assets in the rth year? It is given that $a_0 = 3$ and $a_1 = 7$.

10.27 Let a_r denote the number of nonoverlapping regions into which the interior of a convex r-gon is divided by its diagonals. Suppose no three diagonals meet at one point.

(*a*) Show that

$$a_r - a_{r-1} = \frac{(r-1)(r-2)(r-3)}{6} + r - 2 \qquad r \geqq 3$$

and $a_0 = a_1 = a_2 = 0$.

(*b*) Determine $A(z)$.

(*c*) Determine a_r from $A(z)$.

10.28 There are two kinds of particles inside a nuclear reactor. In every second, an α particle will split into three β particles, and a β particle will split into an α particle and two β particles. If there is a single α particle in the reactor at $t = 0$, how many particles are there altogether at $t = 100$?

10.29 How many spanning trees does the ladder graph in Fig. 10P.2 have?

Hint: Let a_r denote the number of spanning trees the ladder graph of r steps has. Let b_r denote the number of spanning trees that include the first step the ladder graph of r steps has. Express a_r in terms of a_{r-1} and b_r. Express b_r in terms of a_{r-1} and b_{r-1}.

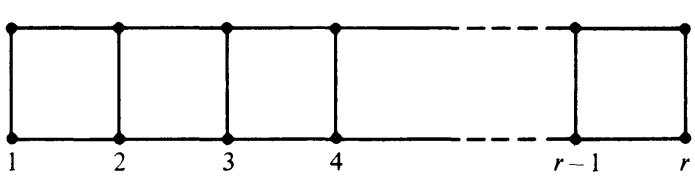

Figure 10P.2

10.30 Let d_r denote the number of ways of permuting r integers $\{1, 2, \ldots, r\}$ so that the integer i will not be in the ith position for $1 \leqq i \leqq r$.

(*a*) Use a combinatorial argument to show that

$$d_r = (r-1)(d_{r-1} + d_{r-2})$$

(*b*) Show that

$$d_r = r!\left[1 - \frac{1}{1!} + \frac{1}{2!} - \cdots + (-1)^r \frac{1}{r!}\right]$$

satisfies the recurrence relation in part (*a*).

10.31 A sequence of binary digits is fed to a counter at the rate of 1 digit/s. The counter is designed to register 1 s in the input sequence. However, it is so slow that it is locked for exactly 7 s following each registration, during which time the input digits are ignored. Let a_r denote the number of binary sequences of length r at the end of which the counter is not locked. Find the difference equation which a_r satisfies, the boundary conditions, and the generating function $A(z)$.

10.32 For the graph in Fig. 10P.3, determine the number of directed paths of length n that start from vertex a and end at vertex d.

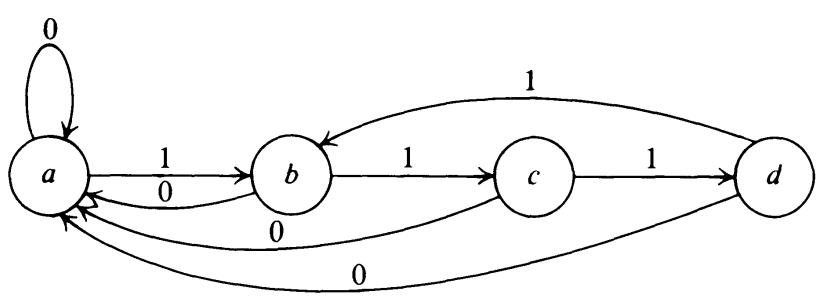

Figure 10P.3

10.33 (*a*) Let $A(z)$ denote the generating function of a numeric function **a**. Show that

$$a_0 = \lim_{z \to 0} A(z)$$

(This result is known as the *initial-value theorem.*)

(*b*) Check the initial-value theorem for the numeric function **a**, where $a_r = 2^{r+1}$ for all r.

10.34 Solve the difference equation

$$a_r^2 - 2a_{r-1}^2 = 1$$

given that $a_0 = 2$.

Hint: Let $b_r = a_r^2$.

10.35 Solve the difference equation

$$ra_r + ra_{r-1} - a_{r-1} = 2^r$$

given that $a_0 = 273$.

Hint: Let $b_r = ra_r$.

10.36 (*a*) Solve the difference equation

$$a_r^2 - 2a_{r-1} = 0$$

given that $a_0 = 4$.

Hint: Let $b_r = \lg a_r$, where lg denotes logorithm base 2.

(*b*) Solve the difference equation

$$a_r = \sqrt{a_{r-1} + \sqrt{a_{r-2} + \sqrt{a_{r-3} + \sqrt{\cdots}}}}$$

given that $a_0 = 4$.

Hint: Compute a_r^2.

10.37 Solve the difference equation

$$a_r - ra_{r-1} = r! \qquad \text{for } r \geqq 1$$

given that $a_0 = 2$.

Hint: Let $b_r = a_r/r!$.

10.38 The product of two 2×2 matrices **A** and **B** can be computed using 7 multiplications and 15 additions. We compute first the following intermediate results:

$$\begin{array}{lll}
s_1 = a_{21} + a_{22} & m_1 = s_2 s_6 & t_1 = m_1 + m_2 \\
s_2 = s_1 - a_{11} & m_2 = a_{11} b_{11} & t_2 = t_1 + m_4 \\
s_3 = a_{11} - a_{21} & m_3 = a_{12} b_{21} & \\
s_4 = a_{12} - s_2 & m_4 = s_3 s_7 & \\
s_5 = b_{12} - b_{11} & m_5 = s_1 s_5 & \\
s_6 = b_{22} - s_5 & m_6 = s_4 b_{22} & \\
s_7 = b_{22} - b_{12} & m_7 = a_{22} s_8 & \\
s_8 = s_6 - b_{21} & &
\end{array}$$

The elements of the product matrix **C** can then be computed as:

$$\begin{aligned}
c_{11} &= m_2 + m_3 \\
c_{12} &= t_1 + m_5 + m_6 \\
c_{21} &= t_2 - m_7 \\
c_{22} &= t_2 + m_5
\end{aligned}$$

Confirm that these equations yield the correct results.