Table 3-1.1

$\circ$	f_1	f_2	f_3	f_4
f_1	f_1	f_2	f_3	f_4
f_2	f_2	f_2	f_2	f_2
f_3	f_3	f_3	f_3	f_3
f_4	f_4	f_3	f_2	f_1

plication of the operation τ. Therefore, 1 can be called a *generator* of the algebraic system $\langle M, \tau \rangle$.

EXAMPLE 2 Let $X = \{a, b\}$ and S denote the set of all mappings from X to X. Let us write $S = \{f_1, f_2, f_3, f_4\}$ where

$$f_1(a) = a \quad f_1(b) = b \quad f_2(a) = a \quad f_2(b) = a$$
$$f_3(a) = b \quad f_3(b) = b \quad f_4(a) = b \quad f_4(b) = a$$

Then $\langle S, \circ \rangle$, where $\circ$ denotes the operation of (left) composition of functions, is an algebraic system in which the operation is associative. The composition table for the operation $\circ$ is given in Table 3-1.1. Note that f_1 is the identity element with respect to the operation $\circ$. Furthermore, not all the elements of S are invertible. These remarks hold whether we choose $\circ$ to denote the left or right composition. This example can be generalized to the set of mappings from any set X to X. We shall return to this example in Sec. 3-2.

EXAMPLE 3 Let $X = \{1, 2, 3, 4\}$ and $f: X \to X$ be given by

$$f = \{\langle 1, 2\rangle, \langle 2, 3\rangle, \langle 3, 4\rangle, \langle 4, 1\rangle\}$$

Let the identity mapping on X be denoted f^0. If we form the composite functions $f \circ f = f^2$, $f^2 \circ f = f^3$, $f^3 \circ f = f^4$, and so on, we find that $f^4 = f^0$. Let us denote f by f^1 and consider the set $F = \{f^0, f^1, f^2, f^3\}$. It is clear that the set F is closed under the operation of composition and that $\langle F, \circ \rangle$ is an algebraic system. The operation $\circ$ is both commutative and associative. Also the element f^0 is the identity element with respect to the operation of composition. The result of composition of any two functions of F is given by Table 3-1.2.

EXAMPLE 4 An equivalence relation called "congruence modulo m" on the set of integers was defined in Sec. 2-3.5. Let $m = 4$ and $\mathbf{Z}_4$ denote the set of equiv-

Table 3-1.2

$\circ$	f^0	f^1	f^2	f^3
f^0	f^0	f^1	f^2	f^3
f^1	f^1	f^2	f^3	f^0
f^2	f^2	f^3	f^0	f^1
f^3	f^3	f^0	f^1	f^2

Table 3-1.3

$+_4$	[0]	[1]	[2]	[3]
[0]	[0]	[1]	[2]	[3]
[1]	[1]	[2]	[3]	[0]
[2]	[2]	[3]	[0]	[1]
[3]	[3]	[0]	[1]	[2]

alence classes generated, so that

$$\mathbf{Z}_4 = \{[0], [1], [2], [3]\}$$

where $[j]$ denotes the set of all those integers which are equivalent to j. Let us define an operation $+_4$ on $\mathbf{Z}_4$ given by

$$[i] +_4 [j] = [(i + j) \pmod 4]$$

for all $i, j = 0, 1, 2, 3$. The operation $+_4$ on $\mathbf{Z}_4$ is described in Table 3-1.3. Since the set $\mathbf{Z}_4$ is closed with respect to the operation $+_4$, we have the algebraic system $\langle \mathbf{Z}_4, +_4 \rangle$ in which the operation $+_4$ is commutative and associative. Also $[0]$ is the identity element, and every element of $\mathbf{Z}_4$ is invertible.

It is easy to observe a similarity between Tables 3-1.2 and 3-1.3. One can obtain Table 3-1.3 from Table 3-1.2 by simply relabeling the entries f^0, f^1, f^2, and f^3 by $[0]$, $[1]$, $[2]$, and $[3]$ respectively and the operation $\circ$ by $+_4$. Let us now define a mapping $\psi: F \to \mathbf{Z}_4$ such that

$$\psi(f^j) = [j] \qquad \text{for } j = 0, 1, 2, 3$$

It is clear from the tables that

$$\psi(f^i \circ f^j) = \psi(f^i) +_4 \psi(f^j) \qquad \text{for all } i, j = 0, 1, 2, 3 \qquad (1)$$

This equation shows that the image of ψ for the argument $f^i \circ f^j$ is the same as the result of the operation $+_4$ applied to the images $\psi(f^i)$ and $\psi(f^j)$ of the elements f^i and f^j. In other words, we can say that the mapping ψ preserves the operations $\circ$ and $+_4$. This property of the mapping ψ is demonstrated in Fig. 3-1.2, which clearly shows that the effect of applying the mapping $\circ$ from $F \times F$ to F and then applying the mapping ψ to the result is the same as the effect of the mapping $\psi \times \psi$ applied to $F \times F$ to obtain an ordered pair of $\mathbf{Z}_4 \times \mathbf{Z}_4$ and then applying the mapping $+_4$ to this ordered pair.

The existence of such a mapping shows that the two algebraic systems $\langle F, \circ \rangle$ and $\langle \mathbf{Z}_4, +_4 \rangle$ are not structurally different. They are only different in the names of the elements and the symbols used for the operations. Any result associated

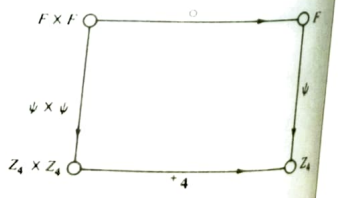

FIGURE 3-1.2

with any one system will also be true for the other system after the labels are changed. We shall now formalize these ideas for any two algebraic systems.

Definition 3-1.1 Let $\langle X, \circ \rangle$ and $\langle Y, * \rangle$ be two algebraic systems of the same type in the sense that both $\circ$ and $*$ are binary (*n*-ary) operations. A mapping $g: X \to Y$ is called a *homomorphism*, or simply morphism, from $\langle X, \circ \rangle$ to $\langle Y, * \rangle$ if for any $x_1, x_2 \in X$

$$g(x_1 \circ x_2) = g(x_1) * g(x_2)$$

If such a function g exists, then it is customary to call $\langle Y, * \rangle$ a homomorphic image of $\langle X, \circ \rangle$, although we must note that $g(X) \subseteq Y$.

The concept of homomorphism is not restricted to algebraic systems with one binary operation. One can extend this definition to any two algebraic systems of the same type. Since in a homomorphism the operations are preserved, we shall see that several properties of the operations are also preserved.

For the algebraic systems $\langle F, \circ \rangle$ and $\langle \mathbf{Z}_4, +_4 \rangle$, the mapping $\psi: F \to \mathbf{Z}_4$ given by Eq. (1) is a homomorphism. Any mapping which satisfies the condition given by Eq. (2) is a homomorphism. In the example of the algebraic systems $\langle F, \circ \rangle$ and $\langle \mathbf{Z}_4, +_4 \rangle$, the mapping is bijective, which is a special case of homomorphism as can be seen from Definition 3-1.2 which follows. It is possible to have more than one homomorphic mapping from one algebraic system to another.

Definition 3-1.2 Let g be a homomorphism from $\langle X, \circ \rangle$ to $\langle Y, * \rangle$. If $g: X \to Y$ is onto, then g is called an *epimorphism*. If $g: X \to Y$ is one-to-one, then g is called a *monomorphism*. If $g: X \to Y$ is one-to-one onto, then g is called an *isomorphism*.

Definition 3-1.3 Let $\langle X, \circ \rangle$ and $\langle Y, * \rangle$ be two algebraic systems of the same type. If there exists an isomorphic mapping $g: X \to Y$, then $\langle X, \circ \rangle$ and $\langle Y, * \rangle$ are said to be *isomorphic*.

In the case when $\langle X, \circ \rangle$ and $\langle Y, * \rangle$ are isomorphic, then the two algebraic systems are structurally indistinguishable in the sense that they differ only in the labels used to denote the elements of the sets and the operations involved. It is easy to see that the inverse of an isomorphism is also an isomorphism. Also all the properties of the operations are preserved in an isomorphism.

Definition 3-1.4 Let $\langle X, \circ \rangle$ and $\langle Y, * \rangle$ be two algebraic systems such that $Y \subseteq X$. A homomorphism g from $\langle X, \circ \rangle$ to $\langle Y, * \rangle$ in such a case is called an *endomorphism*. If $Y = X$, then an isomorphism from $\langle X, \circ \rangle$ to $\langle Y, * \rangle$ is called an *automorphism*.

EXAMPLE 5 Show that the algebraic systems of $\langle F, \circ \rangle$ and $\langle \mathbf{Z}_4, +_4 \rangle$ given in Examples 3 and 4 are isomorphic.

SOLUTION The mapping $\psi: F \to \mathbf{Z}_4$ defined by Eq. (1) is one-to-one onto and is a homomorphism, hence ψ is an isomorphism.

EXAMPLE 6 Let $\langle \mathbf{Z}_4, +_4 \rangle$ and $\langle B, + \rangle$ be the algebraic systems given in Example 4 of Sec. 3-1.2 and Example 6 of Sec. 3-1.1 respectively. Show that $\langle B, + \rangle$ is a homomorphic image of $\langle \mathbf{Z}_4, +_4 \rangle$.

SOLUTION Let a mapping $g: \mathbf{Z}_4 \to B$ be given by

$$g([0]) = g([2]) = 0 \qquad \text{and} \qquad g([1]) = g([3]) = 1$$

It is easy to verify that g is an epimorphism from $\langle \mathbf{Z}_4, +_4 \rangle$ to $\langle B, + \rangle$ because for any $i, j = 0, 1, 2, 3$

$$g([i] +_4 [j]) = g([i]) + g([j]) \qquad ////$$

EXAMPLE 7 Let $S = \{a, b, c\}$ and let $*$ denote a binary operation on S given by Table 3-1.4. Also let $P = \{1, 2, 3\}$ and $\oplus$ be a binary operation on P given by Table 3-1.4. Show that $\langle S, * \rangle$ and $\langle P, \oplus \rangle$ are isomorphic.

SOLUTION Consider a mapping $g: S \to P$ given by

$$g(a) = 3 \qquad g(b) = 1 \qquad \text{and} \qquad g(c) = 2$$

Obviously, g is one-to-one onto. Also a check of entries in Table 3-1.4 shows that g is a homomorphism. ////

EXAMPLE 8 Given the algebraic system $\langle \mathbf{N}, + \rangle$ and $\langle \mathbf{Z}_4, +_4 \rangle$, where $\mathbf{N}$ is the set of natural numbers and $+$ is the operation of addition on $\mathbf{N}$, show that there exists a homomorphism from $\langle \mathbf{N}, + \rangle$ to $\langle \mathbf{Z}_4, +_4 \rangle$.

SOLUTION Define $g: \mathbf{N} \to \mathbf{Z}_4$ given by

$$g(a) = [a (\text{mod } 4)] \qquad \text{for any } a \in \mathbf{N}$$

For $a, b \in \mathbf{N}$, let $g(a) = [i]$ and $g(b) = [j]$; then

$$g(a + b) = [(i + j) (\text{mod } 4)] = [i] +_4 [j] = g(a) +_4 g(b)$$

Observe that $g(0) = [0]$; that is, the mapping g also preserves the identity element. ////

Let $\langle X, \circ \rangle$ be an algebraic system in which $\circ$ is a binary operation on X. Let us assume that E is an equivalence relation on X. The equivalence relation E is said to have the *substitution property* with respect to the operation $\circ$ iff for any $x_1, x_2 \in X$.

$$(x_1 \, E \, x_1') \wedge (x_2 \, E \, x_2') \Rightarrow (x_1 \circ x_2) \, E \, (x_1' \circ x_2') \qquad (3)$$

where $x_1', x_2' \in X$. Implication (3) states that in $x_1 \circ x_2$ if we substitute for x_1 any other element of X which is equivalent to it and similarly for x_2 any other element of X which is equivalent to x_2, then the resulting element is equivalent to $x_1 \circ x_2$; that is, the equivalence relationship is unaltered by such substitutions. The substitution property can be defined for any n-ary operation and also for any number of operations.

Table 3-1.4

$*$	a	b	c
a	a	b	c
b	b	b	c
c	c	b	c

$\oplus$	1	2	3
1	1	2	1
2	1	2	2
3	1	2	3

Definition 3-1.5 Let $\langle X, \circ \rangle$ be an algebraic system and E be an equivalence relation on X. The relation E is called a *congruence relation* on $\langle X, \circ \rangle$ if E satisfies the substitution property with respect to the operation $\circ$.

This definition of congruence relation can be generalized to any algebraic system, whenever the substitution property is satisfied with respect to all the operations of that system.

With the help of congruence relations it is possible to construct new and simpler algebraic systems from a given algebraic system. Also the concept of congruence relation is closely connected to that of homomorphism, as will be shown now.

Let $\langle X, \circ \rangle$ be an algebraic system in which $\circ$ is a binary operation on X. Let E be a congruence relation on $\langle X, \circ \rangle$. Since E is an equivalence relation on X, it partitions X into equivalence classes. The set of equivalence classes is the quotient set X/E. Let $x_1, x_2, y_1, y_2 \in X$ with $x_1 \, E \, y_1$ and $x_2 \, E \, y_2$. Since E is a congruence relation, $(x_1 \circ x_2) \, E \, (y_1 \circ y_2)$, so that $[x_1 \circ x_2] = [y_1 \circ y_2]$ where $[x_i]$ is used to denote the equivalence class in which x_i is a member. Corresponding to the operation $\circ$ on X, we now define a binary operation $*$ on X/E such that for any $x_1, x_2 \in X$

$$[x_1] * [x_2] = [x_1 \circ x_2] \qquad (4)$$

It is easy to see that the operation $*$ is well defined on X/E because $[x_1 \circ x_2]$ is independent of the elements chosen to represent the equivalence classes $[x_1]$ and $[x_2]$. We therefore have an algebraic system $\langle X/E, * \rangle$ called the *quotient algebra* which is of the same type as the algebra $\langle X, \circ \rangle$. In fact, several properties of the operation $\circ$ are preserved by the operation $*$.

In order to show a connection between a homomorphism and congruence relation, we first show that corresponding to a congruence relation E on $\langle X, \circ \rangle$ we can define a homomorphism g_E from $\langle X, \circ \rangle$ onto the quotient algebra $\langle X/E, * \rangle$ that is given by $g_E(x) = [x]$ for any $x \in X$. Obviously,

$$g_E(x_1 \circ x_2) = [x_1 \circ x_2] = [x_1] * [x_2] = g_E(x_1) * g_E(x_2) \qquad (5)$$

The homomorphism g_E is called the *natural homomorphism* associated with the congruence relation E.

As a next step, let us consider a homomorphism f from $\langle X, \circ \rangle$ onto $\langle Y, \oplus \rangle$. Now we show that corresponding to f we can define a congruence relation E_f on $\langle X, \circ \rangle$ by

$$x_1 \, E_f \, x_2 \Leftrightarrow f(x_1) = f(x_2) \qquad \text{for any } x_1, x_2 \in X \qquad (6)$$

This definition guarantees that E_f is an equivalence relation for any mapping

6 If $f: S \rightarrow T$ is a homomorphism from $\langle S, * \rangle$ to $\langle T, \Delta \rangle$ and $g: T \rightarrow P$ is also a h morphism from $\langle T, \Delta \rangle$ to $\langle P, \nabla \rangle$, then $g \circ f: S \rightarrow P$ is a homomorphism from $\langle S$ to $\langle P, \nabla \rangle$.

3-2 SEMIGROUPS AND MONOIDS

Several examples of algebraic systems were given in the previous section. In t section and also in the remainder of this chapter, we study certain abstract al braic systems. An abstract algebraic system is defined in terms of an arbitr set and a number of operations on the set. These operations are assumed to ha properties which are taken as axioms of the system. No other properties a assumed. Any valid conclusion which follows from these axioms is a theorem the system. Such theorems are true for any algebraic system for which the axio hold.

We shall restrict ourselves to algebraic systems with one binary operatio If no restriction is placed on the binary operation, we get an algebraic syste which is too general to be of much use. As a next step, we consider an algebrai system consisting of a set and an associative binary operation on the set. Afte studying such a system, we consider those systems which possess an associativ binary operation and an identity element. These algebraic systems are calle semigroups and monoids respectively and have useful applications in the area of computer arithmetic, formal languages, and sequential machines.

3-2.1 Definitions and Examples

Definition 3-2.1 Let S be a nonempty set and $\circ$ be a binary operation on S. The algebraic system $\langle S, \circ \rangle$ is called a *semigroup* if the operation $\circ$ is associative. In other words $\langle S, \circ \rangle$ is a semigroup if for any $x, y, z \in S$,

$$(x \circ y) \circ z = x \circ (y \circ z)$$

A semigroup can be described by giving the composition table of the operation $\circ$ when S is finite and when its order is not large; otherwise, it can be described by means of some rule for the operation $\circ$ on S. Several examples of semigroups are given here. These include some of the examples given in the previous section as special cases. Note that a semigroup may or may not have an identity element with respect to the operation $\circ$. If there is an identity element, then we have the following definition

Definition 3-2.2 A semigroup $\langle M, \circ \rangle$ with an identity element with respect to the operation $\circ$ is called a *monoid*. In other words, an algebraic system $\langle M, \circ \rangle$ is called a monoid if for any $x, y, z \in M$,

$$(x \circ y) \circ z = x \circ (y \circ z)$$

and there exists an element $e \in M$ such that for any $x \in M$

$$e \circ x = x \circ e = x$$

It has already been shown in Sec. 2-4.4 that an identity element for any binary operation, if it exists, is unique. Therefore, a monoid has a unique or a distinguished element called the *identity* of the monoid. We shall sometimes represent a monoid as $\langle M, \circ, e \rangle$ to emphasize the fact that e is a distinguished element of such a monoid.

For a monoid $\langle M, * \rangle$, the existence of the identity element guarantees that no two columns or rows of the composition table are identical, because if e is the identity, then for $a_i, a_j \in M$ and $a_i \neq a_j$, $e * a_i = a_i \neq e * a_j = a_j$. Similarly, because $a_i * e = a_i$, no two rows are identical. This property will be used later in this section.

EXAMPLE 1 Let X be a nonempty set and X^X be the set of all mappings from X to X. Let $\circ$ denote the operation of composition of these mappings; i.e., for $f, g \in X^X$, $f \circ g$ given by $(f \circ g)(x) = f(g(x))$ for all $x \in X$ is in X^X. The algebra $\langle X^X, \circ \rangle$ is a monoid, because the operation of composition is associative and the identity mapping $f(x) = x$ for all $x \in X$ is the identity of the operation. (See also Example 3, Sec. 3-1.2.)

EXAMPLE 2 In Example 1, if we let $B(X)$ denote the set of all relations from X to X and we let the operation $\circ$ mean the composition of relations on $B(X)$, then we have a monoid $\langle B(X), \circ \rangle$ in which the identity relation is the identity of the monoid.

EXAMPLE 3 Let S be a nonempty set and $\rho(S)$ be its power set. The algebras $\langle \rho(S), \cup \rangle$ and $\langle \rho(S), \cap \rangle$ are monoids with the identities $\varnothing$ and S respectively.

EXAMPLE 4 Let $\mathbf{N}$ be the set of natural numbers. Then $\langle \mathbf{N}, + \rangle$ and $\langle \mathbf{N}, \times \rangle$ are monoids with the identities 0 and 1 respectively. On the other hand, if E denotes the set of positive even numbers, then $\langle E, + \rangle$ and $\langle E, \times \rangle$ are semigroups but not monoids.

EXAMPLE 5 Let $\mathbf{I}$ be the set of integers and $\mathbf{Z}_m$ be the set of equivalence classes generated by the equivalence relation "congruence modulo m" for any positive integer m. The algebraic systems $\langle \mathbf{Z}_m, +_m \rangle$ and $\langle \mathbf{Z}_m, \times_m \rangle$ are monoids in which the operations $+_m$ and $\times_m$ are defined in terms of the operations $+$ and $\times$ on $\mathbf{I}$ as follows. For any $[i], [j] \in \mathbf{Z}_m$

$$[i] +_m [j] = [(i + j) \pmod m]$$
$$[i] \times_m [j] = [(i \times j) \pmod m]$$

The composition tables for $m = 5$ and 6 are given in Table 3-2.1 (see also Example 4, Sec. 3-1.2). In this table $[i]$ is simply written as i. It may be noted that $[0]$ and $[1]$ are the identity elements with respect to the operations $+_m$ and $\times_m$ respectively. Note also that in the case of $m = 5$, the elements $[1]$, $[2]$, $[3]$, and $[4]$ have inverses with respect to the operation $\times_5$, while in the case of $m = 6$ only $[1]$ and $[5]$ are invertible.

The operation $\oplus$ is associative and commutative. Also the partition consisting of single elements of S is the identity for the operation $\oplus$ on $\Pi(S)$.

If in a semigroup (monoid) $\langle S, * \rangle$ the operation $*$ is commutative, then the semigroup (monoid) is called *commutative*. Similarly, if in a monoid $\langle M, *, e \rangle$ every element is invertible, then the monoid is called a *group*. We discuss groups in the next section.

In a monoid $\langle M, * \rangle$, the powers of any particular element, say $a \in M$, are defined as

$$a^0 = e \qquad a^1 = a \qquad a^2 = a * a \qquad \cdots \qquad a^{j+1} = a^j * a \qquad \text{for } j \in \mathbf{N}$$

By using the generalized associative law one can write

$$a^{j+k} = a^j * a^k = a^k * a^j \qquad \text{for all } j, k \in \mathbf{N}$$

A monoid $\langle M, *, e \rangle$ is said to be *cyclic* if there exists an element $a \in M$ such that every element of M can be written as some powers of a, that is, as a^n for some $n \in \mathbf{N}$. In such a case, the cylic monoid is said to be generated by the element a, and the element a is called the *generator* of the cyclic monoid. A cyclic monoid is commutative because for any $b, c \in M$, $b = a^m$ and $c = a^n$ for some m, $n \in \mathbf{N}$, so that

$$b * c = a^m * a^n = a^{m+n} = a^n * a^m = c * b$$

The monoid given in Example 1 of Sec. 3-1.2 is a finite cyclic monoid. On the other hand, $\langle \mathbf{N}, + \rangle$ is an infinite cyclic monoid generated by $1 \in \mathbf{N}$.

The cyclic semigroups or monoids considered so far were all generated by a single element. One can generalize this situation by considering a finite set of elements as the set of generators of a semigroup. For this purpose, let X be a set of n elements called the set of generators and let $*$ be a binary operation on X which is associative. The set of all elements generated by the elements of X by repeated application of $*$ is called the semigroup generated by X. We shall show in Sec. 3-2.2 that a semigroup generated by n generators is a homomorphic image of the free semigroup X^* generated by X.

EXAMPLE 6 Show that the semigroup $\langle X, * \rangle$ in which $X = \{a, b, p, q\}$ and the operation $*$ is given by Table 3-2.2 is generated by the set $\{a, b\}$.

SOLUTION Note that $a * b = p$ and $a * a = q$. Therefore, every element of S involving p or q can be written in terms of a and b. ////

Table 3-2.2

$*$	a	b	p	q
a	q	p	b	a
b	b	b	b	b
p	p	p	p	p
q	a	b	p	q

3-2.2 Homomorphism of Semigroups and Monoids

The concept of homomorphism for algebraic systems was introduced in Sec. 3-1.2. Now we apply this concept to semigroups and monoids. Homomorphisms of semigroups and monoids have useful applications in the economical design of sequential machines and in formal languages.

Definition 3-2.3 Let $\langle S, * \rangle$ and $\langle T, \Delta \rangle$ be any two semigroups. A mapping $g: S \to T$ such that for any two elements $a, b \in S$,

$$g(a * b) = g(a) \Delta g(b) \tag{1}$$

is called a *semigroup homomorphism*.

As before, a semigroup homomorphism is called a semigroup monomorphism, epimorphism, or isomorphism depending on whether the mapping is one-to-one, onto, or one-to-one onto respectively. Two semigroups $\langle S, * \rangle$ and $\langle T, \Delta \rangle$ are said to be isomorphic if there exists a semigroup isomorphic mapping from S to T.

We have already seen in Example 8, Sec. 3-1.2, that there exists a semigroup homomorphism g from $\langle \mathbf{N}, + \rangle$ to $\langle \mathbf{Z}_m, +_m \rangle$ in which the identity of $\langle \mathbf{N}, + \rangle$ is mapped into the identity $[0]$ of $\langle \mathbf{Z}_m, +_m \rangle$.

Let us now examine some of the implications of Eq. (1). For this purpose, let us assume that $\langle S, * \rangle$ is a semigroup and $\langle T, \Delta \rangle$ is an algebraic structure of the same type, that is, Δ is a binary operation on T but it is not necessarily associative. If there exists an onto mapping $g: S \to T$ such that for any $a, b \in S$ Eq. (1) is satisfied, then we can show that Δ must be associative and hence $\langle T, \Delta \rangle$ must be a semigroup. In order to see this, let $a, b, c \in S$

$$\begin{aligned} g((a * b) * c) &= g(a * b) \Delta g(c) \\ &= (g(a) \Delta g(b)) \Delta (g(c)) \end{aligned}$$

On the other hand,

$$g(a * (b * c)) = g((a * b) * c)$$

but $g(a * (b * c))$ can be shown to be equal to $g(a) \Delta (g(b) \Delta g(c))$ by a similar argument. Hence Δ is associative, and $\langle T, \Delta \rangle$ must be a semigroup. This result shows that Eq. (1) preserves the semigroup character because it preserves associativity.

Next, note that if g is a semigroup homomorphism from $\langle S, * \rangle$ to $\langle T, \Delta \rangle$, then for any element $a \in S$ which is idempotent, we must have $g(a)$ idempotent, because

$$g(a * a) = g(a) = g(a) \Delta g(a)$$

The property of idempotency is preserved under the semigroup homomorphism. In a similar manner commutativity is also preserved.

If $\langle S, * \rangle$ is a semigroup with an identity e, that is, $\langle S, *, e \rangle$ is a monoid and g is a homomorphism from $\langle S, * \rangle$ to a semigroup $\langle T, \Delta \rangle$, then for any $a \in S$,

$$g(a * e) = g(e * a) = g(a) \Delta g(e) = g(e) \Delta g(a) = g(a)$$

corresponding to a in the composition table of $\langle S, * \rangle$. Since $f_a = g(a)$, every row of such a table determines the image under the homomorphism g. ////

Let $g(S)$ be the image set of S under the homomorphism g of Theorem 3-2.3 such that $g(S) \subseteq S^S$. For $a, b \in S$, $g(a) = f_a$ and $g(b) = f_b$ are in $g(S)$. Also

$$f_a \circ f_b = g(a) \circ g(b) = g(a * b) = f_{a*b} \in g(S)$$

implying that $g(S)$ is closed under the operation of composition, and $g: S \to g(S)$ is an epimorphism from S onto $g(S)$.

EXAMPLE 2 Let $\langle S, * \rangle$ be a semigroup in which $S = \{a, b, c\}$ and the operation $*$ is given in Table 3-2.4. Let us define a mapping $g: S \to S^S$ given by $g(a) = f_a$, $g(b) = f_b$, and $g(c) = f_c$ where $f_a, f_b, f_c \in S^S$. Also

$$\begin{aligned} f_a(a) &= a & f_a(b) &= b & f_a(c) &= c \\ f_b(a) &= b & f_b(b) &= c & f_b(c) &= a \\ f_c(a) &= c & f_c(b) &= a & f_c(c) &= b \end{aligned}$$

Obviously there are 3^3 elements in S^S and $\langle S^S, \circ \rangle$ is a monoid. However, $g: S \to S^S$ is a homomorphism from $\langle S, * \rangle$ to $\langle S^S, \circ \rangle$ and $g(S) = \{f_a, f_b, f_c\}$. The composition table for $\langle g(S), \circ \rangle$ can be obtained from Table 3-2.4 by writing f_a, f_b, f_c for a, b and c respectively and by replacing $*$ by $\circ$.

The mapping $g: S \to g(S)$ is onto, so that g is an epimorphism from $\langle S, * \rangle$ onto $\langle g(S), \circ \rangle$. This fact does not guarantee that for any two elements $a, b \in S$ such that $a \neq b$ we will also have $g(a) \neq g(b)$. In fact, if any rows in the composition table are identical, then the functions defined by these rows will be equal. This idea suggests that if $\langle S, * \rangle$ is a monoid, then no two rows of the composition table can be identical. In that case, the mapping $g: S \to g(S)$ is one-to-one and onto; that is, g is an isomorphism.

Theorem 3-2.4 Let $\langle M, * \rangle$ be a monoid. Then there exists a subset $T \subseteq M^M$ such that $\langle M, * \rangle$ is isomorphic to the monoid $\langle T, \circ \rangle$.

It is easy to see that the identity element of $\langle T, \circ \rangle$ is $g(e)$ where e is the identity of M.

Theorem 3-2.4 establishes an isomorphism between any finite monoid and a monoid of functions under the operation of composition. Our next theorem establishes the existence of a homomorphism from the free semigroup with n generators to any semigroup with n generators.

Table 3-2.4

$*$	a	b	c
a	a	b	c
b	b	c	a
c	c	a	b

Theorem 3-2.5 Let X be a set containing n elements, let X^* denote the free semigroup generated by X, and let $\langle S, \oplus \rangle$ be any other semigroup generated by any n generators; then there exists a homomorphism g: $X^* \to S$.

PROOF Let Y be the set of n generators of S. Let g: $X \to Y$ be a one-to-one mapping given by $g(x_i) = y_i$ for $i = 1, 2, \ldots, n$. Now, for any string

$$\alpha = x_1 x_2 \ldots x_m$$

of X^*, we define

$$g(\alpha) = g(x_1) \oplus g(x_2) \oplus \cdots \oplus g(x_m)$$

From this definition it follows that for a string $\alpha\beta \in X^*$,

$$g(\alpha\beta) = g(\alpha) \oplus g(\beta)$$

so that g is the required homomorphism. ////

We shall now show that every semigroup homomorphism induces a congruence relation and conversely.

Theorem 3-2.6 Let $\langle S, * \rangle$ and $\langle T, \Delta \rangle$ be two semigroups and g be a semigroup homomorphism from $\langle S, * \rangle$ to $\langle T, \Delta \rangle$. Corresponding to the homomorphism g, there exists a congruence relation R on $\langle S, * \rangle$ defined by

$$x \, R \, y \qquad \text{iff} \qquad g(x) = g(y) \qquad \text{for } x, y \in S$$

PROOF It is easy to see that R is an equivalence relation on S. Let x_1, x_2, x_1', $x_2' \in S$ such that $x_1 \, R \, x_1'$ and $x_2 \, R \, x_2'$. From

$$g(x_1 * x_2) = g(x_1) \, \Delta \, g(x_2) = g(x_1') \, \Delta \, g(x_2') = g(x_1' * x_2')$$

it follows that R is a congruence relation on $\langle S, * \rangle$. ////

Our next theorem permits us to define a homomorphism from a semigroup to its quotient semigroup corresponding to a given congruence relation defined on the semigroup.

Theorem 3-2.7 Let $\langle S, * \rangle$ be a semigroup and R be a congruence relation on $\langle S, * \rangle$. The quotient set S/R is a semigroup $\langle S/R, \oplus \rangle$ where the operation $\oplus$ corresponds to the operation $*$ on S. Also, there exists a homomorphism from $\langle S, * \rangle$ onto $\langle S/R, \oplus \rangle$ called the *natural homomorphism*.

PROOF For any $a \in S$, let $[a]$ denote the equivalence class corresponding to the congruence relation R. For $a, b \in S$ define an operation $\oplus$ on S/R given by

$$[a] \oplus [b] = [a * b]$$

The associativity of the operation $*$ guarantees the associativity of the operation $\oplus$ on S/R, so that $\langle S/R, \oplus \rangle$ is a semigroup. Next, define a mapping g: $S \to S/R$ given by

$$g(a) = [a] \qquad \text{for any } a \in S$$

Then for $a, b \in S$,

$$g(a * b) = [a * b] = [a] \oplus [b] = g(a) \oplus g(b)$$

so that g is a homomorphism from $\langle S, * \rangle$ onto $\langle S/R, \oplus \rangle$. ////

EXAMPLE 3 Let $A = \{0, 1\}$ and A^* be the free semigroup generated by A by the operation of concatenation. Show that the relation R defined for $x, y \in A^*$ such that $x \, R \, y$ iff x and y contain the same number of 1s is a congruence relation. Suggest a homomorphism which induces R on A^*.

SOLUTION It is easy to see that R is a congruence relation on $\langle A^*, \circ \rangle$ where $\circ$ denotes concatenation. Consider the semigroup $\langle N, + \rangle$ and a mapping g: $A^* \to \mathbf{N}$ such that for any $x \in A^*$, $g(x) = n$ where n is the number of 1s in x. Naturally, for any $x, y \in A^*$,

$$g(xy) = g(x) + g(y)$$

so that g is a homomorphism from $\langle A^*, \circ \rangle$ to $\langle \mathbf{N}, + \rangle$. Now for $x, y \in A^*$,

$$g(x) = g(y) \Leftrightarrow x \, R \, y$$

so that the congruence relation R is induced by the homomorphism g. ////

3-2.3 Subsemigroups and Submonoids

The concepts of subalgebras introduced in Sec. 3-1.2 can be applied to semigroups and monoids.

Definition 3-2.5 Let $\langle S, * \rangle$ be a semigroup and $T \subseteq S$. If the set T is closed under the operation, $*$, then $\langle T, * \rangle$ is said to be a *subsemigroup* of $\langle S, * \rangle$. Similarly, let $\langle M, *, e \rangle$ be a monoid and $T \subseteq M$. If T is closed under the operation $*$ and $e \in T$, then $\langle T, *, e \rangle$ is said to be a submonoid of $\langle M, *, e \rangle$.

The definition of subsemigroup requires that the subsemigroup also be a semigroup under the same operation as the original semigroup. The semigroup $\langle S, * \rangle$ itself is a trivial subsemigroup of $\langle S, * \rangle$. For any $a \in S$, the set consisting of all powers of a under the operation $*$ is a subsemigroup; i.e., if $T = \{a, a^2, a^3, \ldots\}$ where $a^2 = a * a$, etc., then $\langle T, * \rangle$ is a cyclic semigroup which is a subsemigroup of $\langle S, * \rangle$ generated by the element a. Similarly, we can consider the subsemigroups generated by any two elements $a, b \in S$, and so on.

EXAMPLE 1 For the semigroup $\langle \mathbf{N}, \times \rangle$, let T be the set of multiples of a positive integer m; then $\langle T, \times \rangle$ is a subsemigroup of $\langle \mathbf{N}, \times \rangle$.

EXAMPLE 2 For the semigroup $\langle \mathbf{N}, + \rangle$, the set E of all the even nonnegative integers is a subsemigroup $\langle E, + \rangle$ of $\langle \mathbf{N}, + \rangle$.

EXAMPLE 3 In Example 1 of Sec. 3-2.2, the semigroup $\langle \{0, 1\}, * \rangle$ is a subsemigroup of $\langle S, * \rangle$, but not a submonoid.

Theorem 3-2.8 For any commutative monoid $\langle M, * \rangle$, the set of idempotent elements of M forms a submonoid.

PROOF Since the identity element $e \in M$ is idempotent, $e \in S$, where S is the set of idempotents of M. Let $a, b \in S$, so that

$$a * a = a \quad \text{and} \quad b * b = b$$

$$\begin{aligned} \text{Now } (a * b) * (a * b) &= (a * b) * (b * a) \qquad (\langle M, * \rangle \text{ is commutative}) \\ &= a * (b * b) * a \\ &= a * b * a \\ &= a * a * b = a * b \end{aligned}$$

Hence $a * b \in S$ and $\langle S, * \rangle$ is a submonoid. ////

EXAMPLE 4 Recall the monoid $\langle \mathbf{Z}_m, \times_m \rangle$ for $m = 5$ and 6 given in Example 5 of Sec. 3-2.1. The sets of left invertibles of $\mathbf{Z}_5$ and $\mathbf{Z}_6$ are $\{[1], [2], [3], [4]\}$ and $\{[1], [5]\}$ respectively, which are submonoids. A similar result holds in general. ////

From any two given semigroups, we shall generate another semigroup called the direct product of the semigroups.

Definition 3-2.6 Let $\langle S, * \rangle$ and $\langle T, \Delta \rangle$ be two semigroups. The *direct product* of $\langle S, * \rangle$ and $\langle T, \Delta \rangle$ is the algebraic system $\langle S \times T, \circ \rangle$ in which the operation $\circ$ on $S \times T$ is defined by

$$\langle s_1, t_1 \rangle \quad \langle s_2, t_2 \rangle = \langle s_1 * s_2, t_1 \Delta t_2 \rangle$$

for any $\langle s_1, t_1 \rangle$ and $\langle s_2, t_2 \rangle \in S \times T$.

From the definition it follows that $\langle S \times T, \circ \rangle$ is a semigroup because the binary operation $\circ$ on $S \times T$ is defined in terms of the operations $*$ and Δ which are both associative. Therefore, the direct product of any two semigroups is a semigroup.

If $\langle S, * \rangle$ and $\langle T, \Delta \rangle$ are both commutative semigroups, then their direct product is also commutative.

If $\langle S, * \rangle$ and $\langle T, \Delta \rangle$ are monoids with e_S and e_T as their identity elements respectively, then their direct product $\langle S \times T, \circ \rangle$ is also a monoid with $\langle e_S, e_T \rangle$ as the identity element, because

$$\langle e_S, e_T \rangle \circ \langle s, t \rangle = \langle e_S * s, e_T \Delta t \rangle = \langle s, t \rangle$$

and

$$\langle s, t \rangle \circ \langle e_S, e_T \rangle = \langle s * e_S, t \Delta e_T \rangle = \langle s, t \rangle$$

It is easy to verify that if z_S and z_T are any zeros of $\langle S, * \rangle$ and $\langle T, \Delta \rangle$ respectively, then $\langle z_S, z_T \rangle$ is a zero of $\langle S \times T, \circ \rangle$. Similarly, if $s \in S$ and $t \in T$ have inverses, then $\langle s^{-1}, t^{-1} \rangle$ is the inverse of $\langle s, t \rangle$.

EXERCISES 3-2

1 Find the zeros of the semigroups $\langle P(X), \cap \rangle$ and $\langle P(X), \cup \rangle$ where X is any given set and $P(X)$ is its power set. Are these monoids? If so, what are the identities?

2 Let the alphabet $V = \{a, b\}$ and A be the set including Λ of all sequences on V beginning with a. Show that $\langle A, \circ, \Lambda \rangle$ is a monoid.

3 Show that the set $\mathbf{N}$ of natural numbers is a semigroup under the operation $x * y = \max\{x, y\}$. Is it a monoid?

4 Let $S = \{a, b\}$. Show that the semigroup $\langle S^S, \circ \rangle$ is not commutative.

5 Let $\langle S, * \rangle$ be a semigroup and $z \in S$ be a left zero. Show that for any $x \in S$, $x * z$ is also a left zero.

6 An element $a \in S$, where $\langle S, * \rangle$ is a semigroup, is called a *left-cancellable* element if for all $x, y \in S$, $a * x = a * y \Rightarrow x = y$. Show that if a and b are left-cancellable, then $a * b$ is also left-cancellable.

7 Show that every finite semigroup has an idempotent.

8 Show that a semigroup with more than one idempotent cannot be a group. Give an example of a semigroup which is not a group.

9 Show that the set of all the invertible elements of a monoid form a group under the same operation as that of the monoid.

10 Find all the subsemigroups of the semigroup $\langle X, * \rangle$ given in Example 6, Sec. 3-2.1.

11 In a monoid, show that the set of left-invertibles (right-invertibles) form a submonoid.

12 Find all the subsemigroups of $\langle \mathbf{Z}_6, \times_6 \rangle$ from Table 3-2.1. Then show that the subsemigroup of a monoid may be a monoid without being a submonoid.

13 Let $\mathbf{I}$ be the set of integers and $\cdot$ denote the operation of multiplication so that $\langle \mathbf{I}, \cdot, 1 \rangle$ is a monoid. Show that $\langle \{0\}, \cdot \rangle$ is a subsemigroup but not a submonoid.

14 Let $g: S \rightarrow T$ be an isomorphism of semigroups $\langle S, * \rangle$ and $\langle T, \Delta \rangle$. Show that if z is a zero of S, then $g(z)$ must be a zero of $\langle T, \Delta \rangle$.

15 Show that every monoid $\langle M, *, e \rangle$ is isomorphic to a submonoid of $\langle M^M, \circ, \Delta \rangle$ where Δ is the identity mapping of M. (*Hint*: See Theorem 3-2.4.)

16 Let $V = \{a, b\}$ be an alphabet. Show that $\langle V^*, \circ, \Lambda \rangle$ is an infinite monoid.

17 Let $\langle T, * \rangle$ be a subsemigroup of $\langle S, * \rangle$. T is called a *left ideal* of S if $S * T \subseteq T$. Similarly define a *right ideal*. If T is both a left and right ideal, then it is called an *ideal*. Show that in $\langle \mathbf{I}, \cdot \rangle$, where $\mathbf{I}$ is the set of integers under multiplication $\cdot$, the set of multiples of an integer n is an ideal.

3-3 GRAMMARS AND LANGUAGES

The basic machine instructions of a digital computer are very primitive compared with the complex operations that must be performed in various disciplines such as engineering, commerce, and mathematics. Although a complex procedure can be programmed in machine language, it is desirable to use a high-level language that contains instructions similar to those required in a particular application. For example, in a payroll application, one wants to manipulate employee records in a master file, generate complex reports, and perform rather simple arithmetic operations on certain data. A language such as COBOL which has high-level commands that manipulate records and generate reports is a definite asset to a programmer.

While high-level programming languages reduce much of the drudgery of machine language programming, they also introduce new problems. A program (compiler) which converts a program to some object language must be written. Also, programming languages must be defined. Sometimes it is the existence of a particular compiler that provides the precise definition of a language. The specification of a language involves the definition of the following:

1 The set of symbols (or alphabet) that can be used in programs

2 The set of all correct programs

3 The "meaning" of all correct programs

In this section we shall be concerned with the first two items of programming languages.

A language L can be considered a subset of the free monoid (see Sec. 3-2.1). The language consisting of the free monoid is interesting since it is too large. Our definition of a language L or sentences over some finite alphabet V, so that $L \subseteq V^*$.

How can a language be represented? A language consisting of infinite set of sentences. Finite languages can be specified by merating all their sentences. However, for infinite languages meration is not possible. On the other hand, any description should be finite. One method of specification which satisfies this a generative device called a *grammar*. A grammar is basically or productions which specify the syntax of the language. It imposes structure on the sentences of a language. The study stitutes an important subarea of computer science called area emerged in the mid-1950s as a result of the efforts of Noam gave a mathematical model of a grammar in connection with languages. In 1960, the concept of a grammar became significant because the syntax of ALGOL 60 was described by a grammar.

A second method of language specification is to have a acceptor, determine whether a given sentence belongs to the approach is discussed further in Chap. 6, along with some very important relationships that exist between grammars and accep

In this section we are concerned with a grammar as a mathfor defining languages and as a device for giving some useful tences in a language. The problem of syntactic analysis will be

3-3.1 Discussion of Grammars

It was mentioned earlier that a grammar imposes a structure of a language. For a sentence in English such a structure is de of subject, predicate, phrase, noun, and so on. On the other hand the structure is given in terms of procedures, statements, exp any case, it may be desirable to describe all such structures and of all the correct or admissible sentences in a language. For ex have a set of correct sentences in English or a set of valid ALGOL grammatical structure of a language helps us determine whether sentence does or does not belong to the set of correct sentences. T