A.1. Sets

We shall use the words 'set,' 'class,' 'collection,' and 'family' interchangeably, although we give preference to 'set.' If S is a set and x is an object in the set S, we shall say that x is a **member of** S, that x is an **element of** S, that x **belongs to** S, or simply that x is in S. If S has only a finite number of members, $x_1, \ldots, x_n$, we shall often describe S by displaying its members inside braces:

$$S = \{x_1, \ldots, x_n\}.$$

Thus, the set S of positive integers from 1 through 5 would be

$$S = \{1, 2, 3, 4, 5\}.$$

If S and T are sets, we say that S is a **subset of** T, or that S is **contained in** T, if each member of S is a member of T. Each set S is a subset of itself. If S is a subset of T but S and T are not identical, we call S a **proper subset** of T. In other words, S is a proper subset of T provided that S is contained in T but T is not contained in S.

If S and T are sets, the **union of** S **and** T is the set $S \cup T$, consisting of all objects x which are members of either S or T. The **intersection of** S **and** T is the set $S \cap T$, consisting of all x which are members of both S and T. For any two sets, S and T, the intersection $S \cap T$ is a subset of the union $S \cup T$. This should help to clarify the use of the word 'or' which will prevail in this book. When we say that x is either in S or in T, we do not preclude the possibility that x is in both S and T.

In order that the intersection of S and T should always be a set, it is necessary that one introduce the **empty set,** i.e., the set with no members. Then $S \cap T$ is the empty set if and only if S and T have no members in common.

We shall frequently need to discuss the union or intersection of several sets. If $S_1, \ldots, S_n$ are sets, their **union** is the set $\bigcup\limits_{j=1}^{n} S_j$ consisting of all x which are members of at least one of the sets $S_1, \ldots, S_n$. Their **intersection** is the set $\bigcap\limits_{j=1}^{n} S_j$, consisting of all x which are members of each of the sets $S_1, \ldots, S_n$. On a few occasions, we shall discuss the union or intersection of an infinite collection of sets. It should be clear how such unions and intersections are defined. The following example should clarify these definitions and a notation for them.

Example 1. Let R denote the set of all real numbers (the real line). If t is in R, we associate with t a subset S_t of R, defined as follows: S_t consists of all real numbers x which are not less than t.

(a) $S_{t_1} \cup S_{t_2} = S_t$, where t is the smaller of t_1 and t_2.

(b) $S_{t_1} \cap S_{t_2} = S_t$, where t is the larger of t_1 and t_2.

(c) Let I be the unit interval, that is, the set of all t in R satisfying $0 \leq t \leq 1$. Then

$$\bigcup_{t \text{ in } I} S_t = S_0$$

$$\bigcap_{t \text{ in } I} S_t = S_1.$$

A.2. Functions

A **function** consists of the following:

(1) a set X, called the domain of the function;

(2) a set Y, called the co-domain of the function;

(3) a rule (or correspondence) f, which associates with each element x of X a single element $f(x)$ of Y.

If (X, Y, f) is a function, we shall also say f **is a function from** X **into** Y. This is a bit sloppy, since it is not f which is the function; f is the rule of the function. However, this use of the same symbol for the function and its rule provides one with a much more tractable way of speaking about functions. Thus we shall say that f is a function from X into Y, that X is the domain of f, and that Y is the co-domain of f—all this meaning that (X, Y, f) is a function as defined above. There are several other words which are commonly used in place of the word 'function.' Some of these are 'transformation,' 'operator,' and 'mapping.' These are used in contexts where they seem more suggestive in conveying the role played by a particular function.

If f is a function from X into Y, the **range** (or **image**) of f is the set of all $f(x)$, x in X. In other words, the range of f consists of all elements y in Y such that $y = f(x)$ for some x in X. If the range of f is all of Y, we say that f is a **function from** X **onto** Y, or simply that f is **onto.** The range of f is often denoted $f(X)$.

Example 2. (a) Let X be the set of real numbers, and let $Y = X$. Let f be the function from X into Y defined by $f(x) = x^2$. The range of f is the set of all non-negative real numbers. Thus f is not onto.

(b) Let X be the Euclidean plane, and $Y = X$. Let f be defined as follows: If P is a point in the plane, then $f(P)$ is the point obtained by rotating P through 90° (about the origin, in the counterclockwise direction). The range of f is all of Y, i.e., the entire plane, and so f is onto.

(c) Again let X be the Euclidean plane. Coordinatize X as in analytic geometry, using two perpendicular lines to identify the points of X with ordered pairs of real numbers (x_1, x_2). Let Y be the x_1-axis, that is, all points (x_1, x_2) with $x_2 = 0$. If P is a point of X, let $f(P)$ be the point obtained by projecting P onto the x_1-axis, parallel to the x_2-axis. In other words, $f((x_1, x_2)) = (x_1, 0)$. The range of f is all of Y, and so f is onto.

(d) Let X be the set of real numbers, and let Y be the set of positive real numbers. Define a function f from X into Y by $f(x) = e^x$. Then f is a function from X onto Y.

(e) Let X be the set of positive real numbers and Y the set of all real numbers. Let f be the natural logarithm function, that is, the function defined by $f(x) = \log x = \ln x$. Again f is onto, i.e., every real number is the natural logarithm of some positive number.

Suppose that X, Y, and Z are sets, that f is a function from X into Y, and that g is a function from Y into Z. There is associated with f and g a function $g \circ f$ from X into Z, known as the **composition** of g and f. It is defined by

$$(g \circ f)(x) = g(f(x)).$$

For one simple example, let $X = Y = Z$, the set of real numbers; let f, g, h be the functions from X into X defined by

$$f(x) = x^2, \qquad g(x) = e^x, \qquad h(x) = e^{x^2}$$

and then $h = g \circ f$. The composition $g \circ f$ is often denoted simply gf; however, as the above simple example shows, there are times when this may lead to confusion.

One question of interest is the following. Suppose f is a function from X into Y. When is there a function g from Y into X such that $g(f(x)) = x$ for each x in X? If we denote by I the **identity function** on X, that is, the function from X into X defined by $I(x) = x$, we are asking the following: When is there a function g from Y into X such that $g \circ f = I$? Roughly speaking, we want a function g which 'sends each element of Y back where it came from.' In order for such a g to exist, f clearly must be 1:1, that is, f must have the property that if $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$. If f is 1:1, such a g does exist. It is defined as follows: Let y be an element of Y. If y is in the range of f, then there is an element x in X such that $y = f(x)$; and since f is 1:1, there is exactly one such x. Define $g(y) = x$. If y is not in the range of f, define $g(y)$ to be any element of X. Clearly we then have $g \circ f = I$.

Let f be a function from X into Y. We say that f is **invertible** if there is a function g from Y into X such that

(1) $g \circ f$ is the identity function on X,
(2) $f \circ g$ is the identity function on Y.

We have just seen that if there is a g satisfying (1), then f is 1:1. Similarly, one can see that if there is a g satisfying (2), the range of f is all of Y, i.e., f is onto. Thus, if f is invertible, f is 1:1 and onto. Conversely, if f is 1:1

and onto, there is a function g from Y into X which satisfies (1) and (2). Furthermore, this g is unique. It is the function from Y into X defined by this rule: if y is in Y, then $g(y)$ is the one and only element x in X for which $f(x) = y$.

If f is invertible (1:1 and onto), the **inverse** of f is the unique function f^{-1} from Y into X satisfying

(1′) $f^{-1}(f(x)) = x$, for each x in X,
(2′) $f(f^{-1}(y)) = y$, for each y in Y.

Example 3. Let us look at the functions in Example 2.

(a) If $X = Y$, the set of real numbers, and $f(x) = x^2$, then f is not invertible. For f is neither 1:1 nor onto.

(b) If $X = Y$, the Euclidean plane, and f is 'rotation through 90°,' then f is both 1:1 and onto. The inverse function f^{-1} is 'rotation through $-90°$,' or 'rotation through 270°.'

(c) If X is the plane, Y the x_1-axis, and $f((x_1, x_2)) = (x_1, 0)$, then f is not invertible. For, although f is onto, f is not 1:1.

(d) If X is the set of real numbers, Y the set of positive real numbers, and $f(x) = e^x$, then f is invertible. The function f^{-1} is the natural logarithm function of part (e): $\log e^x = x$, $e^{\log y} = y$.

(e) The inverse of this natural logarithm function is the exponential function of part (d).

Let f be a function from X into Y, and let f_0 be a function from X_0 into Y_0. We call f_0 a **restriction** of f (or a restriction of f to X_0) if

(1) X_0 is a subset of X,
(2) $f_0(x) = f(x)$ for each x in X_0.

Of course, when f_0 is a restriction of f, it follows that Y_0 is a subset of Y. The name 'restriction' comes from the fact that f and f_0 have the same rule, and differ chiefly because we have restricted the domain of definition of the rule to the subset X_0 of X.

If we are given the function f and any subset X_0 of X, there is an obvious way to construct a restriction of f to X_0. We define a function f_0 from X_0 into Y by $f_0(x) = f(x)$ for each x in X_0. One might wonder why we do not call this *the* restriction of f to X_0. The reason is that in discussing restrictions of f we want the freedom to change the co-domain Y, as well as the domain X.

Example 4. (a) Let X be the set of real numbers and f the function from X into X defined by $f(x) = x^2$. Then f is not an invertible function, but it is if we restrict its domain to the non-negative real numbers. Let X_0 be the set of non-negative real numbers, and let f_0 be the function from X_0 into X_0 defined by $f_0(x) = x^2$. Then f_0 is a restriction of f to X_0.