2. Ordered pairs, and binary relations

The notion of an ordered pair is fundamental for defining and working with binary relations and functions. The notion has to be defined as some sort of set, and it is natural to somehow do this in terms of our known unordered pair notion. For any sets a, b we define

$$(a, b) = \{\{a\}, \{a, b\}\}.$$

This is called the *ordered pair* a, b. The "ordered" adjective is completely expressed in the following proposition.

Proposition 2.1. *$(a, b) = (c, d)$ iff $a = c$ and $b = d$.*

Proof. The direction $\Leftarrow$ is true on logical grounds. Now suppose that $(a, b) = (c, d)$. Then

$$\bigcap\bigcap(a, b) = \bigcap\bigcap\{\{a\}, \{a, b\}\} = \bigcap(\{a\} \cap \{a, b\}) = \bigcap\{a\} = a,$$

and similarly $\bigcap\bigcap(c, d) = c$, so $a = c$. Now $\bigcup(a, b) = \bigcup\{\{a\}, \{a, b\}\} = \{a\} \cup \{a, b\} = \{a, b\}$, and similarly $\bigcup(c, d) = \{c, d\}$, so

$$(*) \qquad \{a, b\} = \{c, d\}.$$

Now suppose that $b \neq d$. Now by $(*)$, $b \in \{a, b\} = \{c, d\}$, so $b = c$; but also $a = c$, so $a = b$. Then by $(*)$, $d \in \{c, d\} = \{a, b\}$, so $d = b$, contradiction. So $b = d$. □

The property expressed in Proposition 2.1 is really all that is desired in a definition of ordered pair. In principle, any definition that satisfies this property would be satisfactory. See the exercises. The definition given here appears to be the simplest of various possibilities, and it has become standard.

Because of this proposition, we can define

$$1^{\text{st}}(a, b) = a \quad \text{and} \quad 2^{\text{nd}}(a, b) = b$$

for the *first coordinate* and *second coordinate* of an ordered pair (a, b).

A *binary relation* is by definition any collection of ordered pairs. This is a notion which will be specialized in the following chapters to important special cases: functions, equivalence relations, and orderings. For now we introduce some additional notions which will be important in all of the special cases. The adjective "binary" implies that there are relations which are not binary. Indeed there are, but we will introduce them only later; they are not of such central significance as binary relations. Since we deal mainly with binary relations, we will sometimes omit the adjective "binary".

The *domain* of a relation R is the set

$$\text{dmn}(R) = \{a : (a, b) \in R \text{ for some } b\}.$$

Thus for any set a, $a \in \mathrm{dmn}(R)$ iff $(a,b) \in R$ for some b. This is an instance of the "illegal" use of the comprehension axioms. Legally, we could define

$$\mathrm{dmn}(R) = \left\{a \in \bigcup\bigcup R : (a,b) \in R \text{ for some } b\right\}.$$

Under this legal definition, if $a \in \mathrm{dmn}(R)$, then there is a b such that $(a,b) \in R$. Conversely, if $(a,b) \in R$ for some b, then $\{a\} \in \{\{a\},\{,b\}= (a,b) \in R$, so $\{a\} \in \bigcup R$, and $a \in \{a\} \in \bigcup R$, so $a \in \bigcup\bigcup R$; hence $a \in \mathrm{dmn}(R)$ in the sense of the legal definition. So we see that the original definition was ok. In future cases of this sort, where it is relatively easy to "legalize" an illegal use of comprehension we will not give as many details.

Analogous to the domain is the *range* of a relation R; it is defined to be

$$\mathrm{rng}(R) = \{b : (a,b) \in R \text{ for some } a\}.$$

This definition can be justified like the definition of domain was. Both of these definitions will be used mainly when R is a function (see below).

If we want to define a subset of a relation R, it is natural to use an extension of the notation associated with the comprehension axiom:

$$\{(a,b) \in R : \ldots\ldots\}$$

is to be considered as an abbreviation for

$$\{x \in R : \text{there exist } a,b \text{ such that } x = (a,b) \text{ and } \ldots\ldots\}.$$

As with the comprehension axiom, we will sometimes leave off the "$\in R$" and write simply $\{(a,b) : \ldots\ldots\}$. This means of course, that there is some relation R such that the condition $\ldots\ldots$, whatever it is, implies that $(a,b) \in R$.

To proceed, we need another axiom.

Axiom 5. (*Power set*) **For any set A, there is a set B whose members are exactly all of the subsets of A.**

By the extensionality axiom, we can define the *power set* of A to be the set given by this axiom; we denote it by $\mathscr{P}(A)$.

This axiom does not directly have to do with relations, but we need it to justify the following definition. For any sets A, B, we define

$$A \times B = \{(a,b) : a \in A \text{ and } b \in B\}.$$

This set is called the *product* of A and B. To justify it, note that if $a \in A$ and $b \in B$, then both $\{a\}$ and $\{a,b\}$ are subsets of $A \cup B$, and hence they are members of $\mathscr{P}(A \cup B)$. It follows that $(a,b) = \{\{a\},\{a,b\}\}$ is a subset of $\mathscr{P}(A \cup B)$, and hence it is a member of $\mathscr{P}(\mathscr{P}(A \cup B))$. This justifies the definition of $A \times B$.

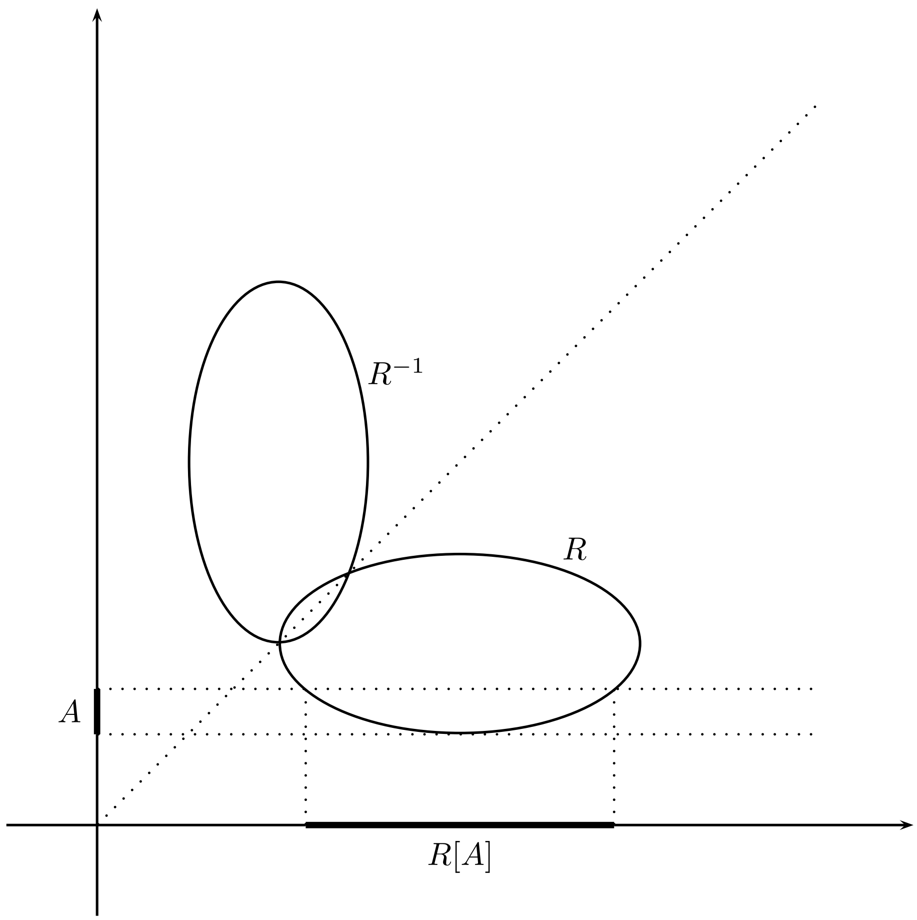
Of course, $A \times B$ is a generalization to arbitrary sets of the set $\mathbb{R} \times \mathbb{R}$, the ordinary real plane. This relation is probably the one most familiar to you.

We introduce two more definitions concerning a relation R:

$$R^{-1} = \{(a,b) : (b,a) \in R\};$$
$$R[A] = \{b \in \text{rng}(R) : (a,b) \in R \text{ for some } a \in A\}.$$

For $R[A]$, we do not assume that $A \subseteq \text{dmn}(R)$. But note that $R[A] = R[A \cap \text{dmn}(R)]$. We call R^{-1} the *inverse* of R. Note that we do not assume that R is a function, much less a one-one function. For one thing, we have not yet introduced the notion of a function. The relation R^{-1} is of interest even if R is not a function.

These two operations are illustrated in the following diagram.



Exercises, chapter 2

1. For any sets a, b, let $[\![a,b]\!] = \{\{\emptyset, a\}, \{\{\emptyset\}, b\}\}$. Prove that if $[\![a,b]\!] = [\![c,d]\!]$ then $a = c$ and $b = d$. Thus we could have used $[\![a,b]\!]$ for the notion of ordered pair rather than (a,b).

2. For any set a, b let $\langle\langle a,b\rangle\rangle = \{\{a\}, b\}$. Find sets a, b such that $\langle\langle a,b\rangle\rangle = \langle\langle c,d\rangle\rangle$ while ($a \neq c$ or $b \neq d$). Thus we could not use $\langle\langle a,b\rangle\rangle$ in place of (a,b) for the notion of ordered pair.

3. Show that if R is a relation and A and B are any sets, then $R[A \cup B] = R[A] \cup R[B]$.

4. Show that if R is a relation and A and B are any sets, then $R[A \cap B] \subseteq R[A] \cap R[B]$.

5. Give an example showing that the inclusion in exercise 4 cannot be replaced by equality in general.

6. Show that if R is a relation and A and B are any sets,, then $R[A]\backslash R[B] \subseteq R[A\backslash B]$.

7. Give an example showing that the inclusion in exercise 6 cannot be replaced by equality in general.

8. Prove that if R is a relation, then $R \subseteq \mathrm{dmn}(R) \times \mathrm{rng}(R)$.

9. Give an example showing that the inclusion in exercise 8 cannot be replaced by equality in general.

10. Justify the definition of R^{-1}.

11. Show that for any relation R, $\mathrm{dmn}(R^{-1}) = \mathrm{rng}(R)$ and $\mathrm{rng}(R^{-1}) = \mathrm{dmn}(R)$.

12. Show that for any sets A, B, C, $A \times (B \cup C) = (A \times B) \cup (A \times C)$, $A \times (B \cap C) = (A \times B) \cap (A \times C)$, and $A \times (B\backslash C) = (A \times B)\backslash(A \times C)$. Similarly for operations on the left side.

13. Prove that $(A \times B)^{-1} = B \times A$.

14. If R, S are relations, define

$$R \,|\, S = \{(a, c) : \text{ there is a set } b \text{ such that } (a, b) \in R \text{ and } (b, c) \in S\}.$$

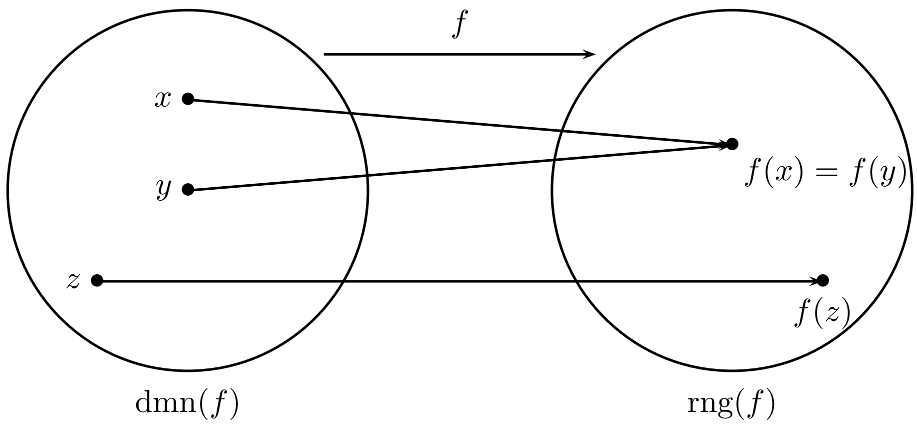
Prove the following, for any relations R, S, T:
(i) $R|(S|T) = (R|S)|T$.
(ii) $(R|S)^{-1} = S^{-1}|R^{-1}$.
(iii) $(R|S) \cap T^{-1} = \emptyset$ iff $(S|T) \cap R^{-1} = \emptyset$.

3. Functions

A *function* is a relation f such that for all x, y, z, if $(x, y) \in f$ and $(x, z) \in f$, then $y = z$. This is a rigorous formulation of the elementary definition of a function as being a rule which assigns to each element of its domain some unique element of its range. We call this unique value the *value* of f at x, and denote it by $f(x)$. Many other notations for this value are common, and we will use some of them from time to time: fx, f_x, f^x, and others. Each element of the domain of f is called an *argument* of f. A function can be pictured as on the top of the next page, although one should understand that it is by no means necessary or even common to have domain and range not intersecting.

If f is a function with domain I, we will sometimes write f as $\langle f(i) : i \in I\rangle$, indicating that f is thought of as a generalized kind of sequence of elements $f(i)$.

The following proposition is fundamental; it gives an equivalent condition for two functions to be equal which is usually the one actually used in practice, rather than directly showing that the two functions have the same members.



Proposition 3.1. *For any functions f and g, the following conditions are equivalent:*
(i) $f = g$.
(ii) $\text{dmn}(f) = \text{dmn}(g)$, *and for all* $a \in \text{dmn}(f)$, $f(a) = g(a)$.

Proof. (i)⇒(ii): this is true on the basis of logic alone (substitution of equals for equals).

(ii)⇒(i): Assume (ii), and suppose that $x \in f$. Since f is a relation, there exist a, b such that $x = (a, b)$. Thus $a \in \text{dmn}(f)$ and $f(a) = b$. So by (ii), also $a \in \text{dmn}(g)$, and $g(a) = f(a) = b$. It follows that $x = (a, b) \in g$. This proves that $f \subseteq g$. By symmetry, $g \subseteq f$, so (i) holds. □

In analogy with Proposition 3.1, we can define a function by specifying its domain D, and then for each $a \in D$ giving a rule R which specifies the value of the function. More rigorously, one sets

$$f = \{(a, b) : a \in D \text{ and } R \text{ gives the value } b\}.$$

As long as this "rigorous" definition can be made really rigorous using the comprehension axioms, this method of defining a function is ok. Sometimes we specify the domain D of a function and then indicate the action of the function by saying that $a \mapsto \ldots$ for each $a \in D$. This is to be taken as shorthand for saying $f = \{(a, \ldots) : a \in D\}$, which in turn is an abbreviation for something more formal if we need to check that the existence of the function follows from the axioms.

We say that f is a *mapping* or function *from A into B* iff f is a function, $\text{dmn}(f) = A$, and $\text{rng}(f) \subseteq B$. We write $f : A \to B$ to abbreviate this. Note that it is not required that $\text{rng}(f) = B$; in case $\text{rng}(f) = B$, we say that f *maps onto* B, or is a *surjection from A to B*. Thus for the same function f we can have many sets B such that $f : A \to B$; any set B which contains the range of f will work. But A is uniquely determined by f; it is the domain of f. There is, unfortunately, conflicting terminology here. Some books use "image" instead of our range, for example.

Our definition of big unions in chapter 1 can now take a more familiar form. If f is a function with domain I, then we define

$$\bigcup_{i\in I} f(i) = \bigcup \mathrm{rng}(f) \quad \text{and} \quad \bigcap_{i\in I} f(i) = \bigcap \mathrm{rng}(f).$$

Proposition 3.2. *Let f be a function with domain I, and let x be any set.*
(i) $x \in \bigcup_{i\in I} f(i)$ iff there is an $i \in I$ such that $x \in f(i)$.
(ii) If I is nonempty, then $x \in \bigcap_{i\in I} f(i)$ iff $x \in f(i)$ for all $i \in I$. □

Proposition 3.3. *Suppose that $f : A \to B$ and X is any set. Then $f[X] = \{f(x) : x \in X \cap \mathrm{dmn}(f)\}$. Hence if $X \subseteq \mathrm{dmn}(f)$, then $f[X] = \{f(x) : x \in X\}$.*

Proof. Note that $\{f(x) : x \in X\cap\mathrm{dmn}(f)\}$ is $\{z : z = f(x)$ for some $x \in X\cap\mathrm{dmn}(f)\}$. This notation will be used in similar situations. To prove the conclusion of the proposition, let z be any set. Then

$$\begin{aligned} z \in f[X] \quad &\text{iff} \quad (x, z) \in f \text{ for some } x \in X \\ &\text{iff} \quad \text{there is an } x \in X \cap \mathrm{dmn}(f) \text{ such that } z = f(x). \end{aligned}$$

Note here that if $(x, z) \in f$ and $x \in X$, then $x \in \mathrm{dmn}(f)$ and $z = f(x)$; conversely, if $x \in X\cap\mathrm{dmn}(f)$ and $z = f(x)$, then $x \in X$ and $(x, z) \in f$. This gives details on the second equivalence. □

We call $f[A]$ the *image* of A under F. Thus the image of $\mathrm{dmn}(f)$ under f is the range of f.

Proposition 3.4. *Suppose that $f : A \to B$ and Y is any set. Then $f^{-1}[Y] = \{x \in A : f(x) \in Y\}$.*

Proof. For any set x,

$$\begin{aligned} x \in f^{-1}[Y] \quad &\text{iff} \quad (y, x) \in f^{-1} \text{ for some } y \in Y \\ &\text{iff} \quad (x, y) \in f \text{ for some } y \in Y \\ &\text{iff} \quad x \in A \text{ and } f(x) \in Y. \end{aligned}$$

The last equivalence here can be seen in great detail as follows. If $(x, y) \in f$ for some $y \in Y$, then $x \in \mathrm{dmn}(f) = A$, and $f(x) = y \in Y$. On the other hand, if $x \in A$ and $f(x) \in Y$, then $(x, f(x)) \in f$ and $f(x) \in Y$. □

The set $f^{-1}[Y]$ is called the *pre-image* of Y under f.

We now turn to methods of combining several functions to form a new function. The precise definition of function which we have given is essential in understanding the following proposition, which describes a somewhat advanced method of combining functions.

Proposition 3.5. *Suppose that $\mathscr{F}$ is a collection of functions, and $f \cup g$ is a function for all $f, g \in \mathscr{F}$. Then $\bigcup \mathscr{F}$ is a function.*

Proof. Clearly $\bigcup \mathscr{F}$ is a relation. Now suppose that $(x,y),(x,z) \in \bigcup \mathscr{F}$. Then there are $f, g \in \mathscr{F}$ such that $(x,y) \in f$ and $(x,z) \in g$. So $(x,y),(x,z) \in f \cup g$, and $f \cup g$ is a function by hypothesis, so $y = z$. Hence $y = z$. This shows that $\bigcup \mathscr{F}$ is a function. □

Corollary 3.6. *Suppose that $\mathscr{F}$ is a collection of functions, and $f \subseteq g$ or $g \subseteq f$ for all $f, g \in \mathscr{F}$. Then $\bigcup \mathscr{F}$ is a function.* □

We now define the very important *composition* of functions g, f. This is a new function, denoted by $g \circ f$, with domain equal to $f^{-1}[\mathrm{dmn}(g)] = \{a \in \mathrm{dmn}(f) : f(a) \in \mathrm{dmn}(g)\}$, and with value $g(f(a))$ for any such a. Thus more formally (see the discussion above following 3.1),

$$g \circ f = \{(a,b) : a \in \mathrm{dmn}(f) \text{ and } f(a) \in \mathrm{dmn}(g) \text{ and } b = g(f(a))\}.$$

Proposition 3.7. *If f and g are functions, then* $\mathrm{rng}(g \circ f) = g[\mathrm{rng}(f)]$.

Proof. Suppose that b is any set. Then

$$\begin{aligned}
b \in \mathrm{rng}(g \circ f) \quad &\text{iff} \quad \text{there is an } a \in \mathrm{dmn}(g \circ f) \text{ such that } (g \circ f)(a) = b \\
&\text{iff} \quad \text{there is an } a \in \mathrm{dmn}(f) \text{ such that } f(a) \in \mathrm{dmn}(g) \text{ and } g(f(a)) = b \\
&\text{iff} \quad \text{there is a } c \in \mathrm{rng}(f) \text{ such that } c \in \mathrm{dmn}(g) \text{ and } g(c) = b \\
&\text{iff} \quad b \in g[\mathrm{rng}(f)].
\end{aligned}$$
□

Proposition 3.8. *If $f : A \to B$ and $g : B \to C$, then $g \circ f : A \to C$.*

Proof. The hypothesis gives $\mathrm{rng}(f) \subseteq B = \mathrm{dmn}(g)$, so $g \circ f$ has domain A. Finally, $\mathrm{rng}(g \circ f) = g[\mathrm{rng}(f)] \subseteq C$. □

Proposition 3.9. *If f, g, h are functions, then $h \circ (g \circ f) = (h \circ g) \circ f$.*

Proof. For any set a,

$$\begin{aligned}
a \in \mathrm{dmn}(h \circ (g \circ f)) \quad &\text{iff} \quad a \in \mathrm{dmn}(g \circ f) \text{ and } (g \circ f)(a) \in \mathrm{dmn}(h) \\
&\text{iff} \quad a \in \mathrm{dmn}(f) \text{ and } f(a) \in \mathrm{dmn}(g) \text{ and } g(f(a)) \in \mathrm{dmn}(h); \\
a \in \mathrm{dmn}((h \circ g) \circ f) \quad &\text{iff} \quad a \in \mathrm{dmn}(f) \text{ and } f(a) \in \mathrm{dmn}(h \circ g) \\
&\text{iff} \quad a \in \mathrm{dmn}(f) \text{ and } f(a) \in \mathrm{dmn}(g) \text{ and } g(f(a)) \in \mathrm{dmn}(h).
\end{aligned}$$

Thus the two functions have the same domain. For any a in their common domain,

$$(h \circ (g \circ f))(a) = h((g \circ f)(a)) = h(g(f(a))) = (h \circ g)(f(a)) = ((h \circ g) \circ f)(a).$$
□

Proposition 3.10. *For any functions f, g and any set A, $(g \circ f)[A] = g[f[A]]$.*

Proof. Let x be any set. First suppose that $x \in (g \circ f)[A]$. Then there is an $a \in A \cap \mathrm{dmn}(g \circ f)$ such that $x = (g \circ f)(a)$. thus $a \in A$, $a \in \mathrm{dmn}(f)$, $f(a) \in \mathrm{dmn}(g)$, and $x = g(f(a))$. Let $b = f(a)$. then $b \in f[A]$ and $g(b) = x$, so $x \in g[f[A]]$.

Conversely, suppose that $x \in g[f[A]]$. Then there is a $b \in \mathrm{dmn}(g) \cap f[A]$ such that $x = g(b)$. So there is an $a \in \mathrm{dmn}(f) \cap A$ such that $b = f(a)$. Hence $a \in \mathrm{dmn}(f)$ and $f(a) = b \in \mathrm{dmn}(g)$, so $a \in \mathrm{dmn}(g \circ f)$. Also $a \in A$ and $x = g(f(a)) = (g \circ f)(a)$, so $x \in (g \circ f)[A]$. □

Proposition 3.11. *For any functions f, g and any set A, $(g \circ f)^{-1}[A] = f^{-1}[g^{-1}[A]]$.*

Proof. Let $B = \mathrm{dmn}(f)$ and $C = \mathrm{dmn}(g)$. Therefore, $\mathrm{dmn}(g \circ f) = f^{-1}[C]$, and by Proposition 3.4, for any set b,

$$\begin{aligned} b \in (g \circ f)^{-1}[A] \quad & \text{iff} \quad b \in f^{-1}[C] \text{ and } (g \circ f)(b) \in A \\ & \text{iff} \quad b \in B \text{ and } f(b) \in C \text{ and } g(f(b)) \in A \\ & \text{iff} \quad b \in B \text{ and } f(b) \in g^{-1}[A] \\ & \text{iff} \quad b \in f^{-1}[g^{-1}[A]]. \end{aligned}$$ □

Another important notion when talking about functions is that of a one-one function. We say that a function f is *one-one* iff f^{-1} is also a function; also in this case f is said to be *injective*. The following is purely a theorem of logic, but it is very useful.

Proposition 3.12. *For any function f the following conditions are equivalent:*
(i) f is one-one.
(ii) For any $x, y \in \mathrm{dmn}(f)$, if $f(x) = f(y)$, then $x = y$.
(iii) For any $x, y \in \mathrm{dmn}(f)$, if $x \neq y$, then $f(x) \neq f(y)$. □

Proposition 3.13. *If $f : A \to B$ is one-one, then for any $a \in A$ and $b \in B$ the following conditions are equivalent:*
(i) $f(a) = b$.
(ii) b is in the range of f and $f^{-1}(b) = a$.

Proof. (i) is true iff $(a, b) \in f$, and (ii) is true iff $(b, a) \in f^{-1}$, so the equivalence follows. □

We also introduce an *identity function* Id_A for every set A; this is a function with domain A which assigns to each $a \in A$, the element a itself.

The following characterization of one-one functions is frequently useful.

Proposition 3.14. *Let $f : A \to B$, with A nonempty. Then the following conditions are equivalent:*
(i) f is one-one.
(ii) There is a $g : B \to A$ such that $g \circ f = \mathrm{Id}_A$.

Proof. (i)⇒(ii): Fix $a_0 \in A$. Define $g : B \to A$ by letting $g(b) = f^{-1}(a)$ if $a \in \mathrm{rng}(f)$ (hence equivalently if $a \in \mathrm{dmn}(f^{-1})$), and $g(b) = a_0$ otherwise. Then for any $a \in A$ we have $(g \circ f)(a) = g(f(a)) = f^{-1}(f(a)) = a$. So $g \circ f = \mathrm{Id}_A$.

(ii)⇒(i): Assuming (ii), suppose that $a, c \in A$ and $f(a) = f(c)$. Then $a = \mathrm{Id}_A(a) = (g \circ f)(a) = g(f(a)) = g(f(c)) = (g \circ f)(c) = \mathrm{Id}_A(c) = c$. □

The condition here that A is nonempty is necessary. In fact, the empty set is itself a function, and it maps $\emptyset$ into any set. But there is no function which maps a nonempty set into $\emptyset$.

There is an analogous statement to 3.14 for onto functions. To prove it, we need another axiom.

Axiom 6. *(Choice)* **For any sets A, B and any function f mapping A onto B there is a function $g : B \to A$ such that $f \circ g = \mathrm{Id}_B$.**

The name can be justified by noticing that for any $b \in B$, $g(b)$ "chooses" an element $a \in A$ such that $f(a) = b$. In fact, $b = \mathrm{Id}_B(b) = (f \circ g)(b) = f(g(b))$. This axiom is somewhat controversial. It has many equivalent formulations, and much later in these notes we will prove some of these equivalences. Since we will need the axiom occasionally before we get to these proofs, we will put a superscript ^{ch} on results requiring the axiom of choice in their proofs given here. This means that when we get around to proving the equivalents of the axiom of choice we will have to avoid using any result up to that point which has this superscript.

Proposition 3.15^{ch}**.** *Let $f : A \to B$. Then the following conditions are equivalent:*
(i) f maps onto B
(ii) There is a $g : B \to A$ such that $f \circ g = \mathrm{Id}_B$.

Proof. (i)⇒(ii): This is true by the axiom of choice.
(ii)⇒(i): Assume (ii). In order to show that f maps onto B, take any $b \in B$; we want to find $a \in A$ such that $f(a) = b$. Let $a = g(b)$. Then $f(a) = f(g(b)) = (f \circ g)(b) = \mathrm{Id}_B(b) = b$. □

A function $f : A \to B$ which is one-one and onto is called a *bijection.* Functions of this kind are very important, and we will use them shortly in defining the basic notion of size of sets. The following fact along the lines of the preceding two propositions does not require the axiom of choice.

Proposition 3.16. *For any function $f : A \to B$ the following conditions are equivalent:*
(i) f is a bijection.
(ii) f^{-1} is a function mapping B into A.
(iii) There is a function $g : B \to A$ such that $f \circ g = \mathrm{Id}_B$ and $g \circ f = \mathrm{Id}_A$.

Moreover, if g satisfies (iii), then $g = f^{-1}$.

Proof. (i)⇒(ii): Assume (i). If $(x, y), (x, z) \in f^{-1}$, then $(y, x), (z, x) \in f$, and so $y, z \in \mathrm{dmn}(f)$ and $f(y) = x = f(z)$. Hence $y = z$ since f is one-one. So f^{-1} is a function. By facts on relations, we have $\mathrm{dmn}(f^{-1}) = \mathrm{rng}(f) = B$ (since f maps onto B). Also, $\mathrm{rng}(f^{-1}) = \mathrm{dmn}(f) = A$, so $f^{-1} : B \to A$.
(ii)⇒(iii): Assume that $f^{-1} : B \to A$. Thus $f \circ f^{-1} : B \to B$, and for any $b \in B$, $(f \circ f^{-1})(b) = f((f^{-1})(b)) = b$. Hence $f \circ f^{-1} = \mathrm{Id}_B$. Similarly, $f^{-1} \circ f = \mathrm{Id}_A$.
(iii)⇒(i): Assume (iii). If $f(x) = f(y)$, then

$$x = \mathrm{Id}_A(x) = (f^{-1} \circ f)(x) = f^{-1}(f(x)) = f^{-1}(f(y)) = (f^{-1} \circ f)(y) = \mathrm{Id}_A(y) = y,$$

so f is one-one. Given $b \in B$, we have $f(f^{-1}(b)) = (f \circ f^{-1})(b) = \mathrm{Id}_B(b) = b$, so f maps onto B.

Finally, if g is as in (iii), then $f^{-1} : B \to A$ by The equivalence of (i)–(iii), and for any $b \in B$ we have $b = \mathrm{Id}_B(b) = (f \circ g)(b) = f(g(b))$, and hence $f^{-1}(b) = g(b)$ by Proposition 3.13. so g and f^{-1} have the same domain and agree on all arguments in that domain, so they are equal. □

We introduce some more important definitions concerning functions, but for now we do not prove anything about them; but see the exercises.

If f is a function and A is any set, we denote by $f \restriction A$ the *restriction* of f to A:

$$f \restriction A = \{(a, b) \in f : a \in A\}.$$

This is a function whose domain is $A \cap \mathrm{dmn}(f)$; on this domain, it acts just like f does.

For any sets A and B we let ${}^AB$ be the set of all functions mapping A into B. The exponent is put to the left, since we will have an operation on numbers which has the exponent on the right.

If $\langle A_i : i \in I\rangle$ is a system of sets (by which we just mean that A is a function with domain I), we define the *cartesian product* of the system to be

$$\prod_{i\in I} A_i = \{f : f \text{ is a function, } \mathrm{dmn}\, f = I \text{ and } f(i) \in A_i \text{ for all } i \in I\}.$$

A familiar procedure in mathematics is to make a construction from a set A, producing a new set B and a one-one function f from A into B, thereby "extending" A. Then one customarily says something like "we may assume that actually $A \subseteq B$". For example, in constructing the rational numbers from the integers, a mapping from the integers into constructed rationals is given; or in constructing the real numbers from the rationals, a similar overall procedure is used. We want to indicate how this "replacement" process can be done in a rigorous manner. To do this it is convenient to have a standard way of producing new elements from a given set A, and here the special definition of ordered pair can be used. This standard way involves a new axiom, which has much wider uses than just the present situation, as we will see.

Axiom 7. (*Foundation*) **If A is a nonempty set, then A has an element a such that $a \cap A = \emptyset$.**

This axiom excludes some strange kinds of sets and is useful in this way, in order to simplify some constructions; it also can form the basis for building the entire hierarchy of sets, as we will see much later in these notes. First we derive some simple consequences of it, before turning to the way it is used to produce standard new elements from a given set.

Proposition 3.17. *$x \notin x$, for every set x.*

Proof. Suppose that $x \in x$. Then the foundation axiom applied to $\{x\}$ gives a contradiction. □

Proposition 3.18. *For any sets x, y, it is not possible to have $x \in y \in x$.*

Proof. Otherwise, the foundation axiom applied to $\{x, y\}$ gives a contradiction. □

Proposition 3.19. *$(a, b) \notin a$, for all sets a, b.*

Proof. Otherwise, $a \in \{a\} \in \{\{a\}, \{a,b\}\} = (a,b) \in a$. Now we apply the foundation axiom to $A \stackrel{\text{def}}{=} \{a, \{a\}, (a,b)\}$ and obtain an element $x \in A$ such that $x \cap A = \emptyset$. But by considering the three possibilities for x we see that this is a contradiction. $\square$

Proposition 3.19 gives our standard method for producing new elements, given a set a. No matter what b is, (a,b) is an element not in a. Now we can give the promised replacement theorem.

Theorem 3.20. *Suppose that f is a one-one function from A into B. Then there exist C and g such that $A \subseteq C$, g is a one-one function from C onto B, and $f(a) = g(a)$ for all $a \in A$.*

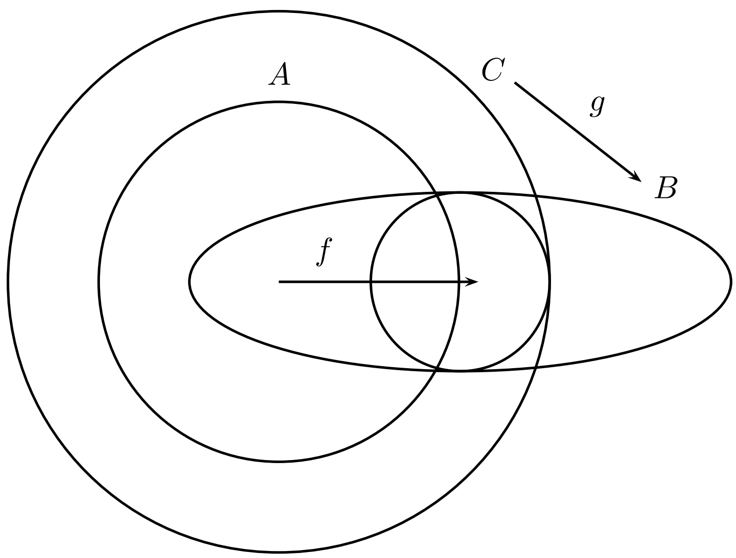
Namely, one can take

$$C = A \cup \{(A,b) : b \in B \backslash \operatorname{rng}(f)\},$$

while g has the property that for any $c \in C$,

$$g(c) = \begin{cases} f(c) & \text{if } c \in A, \\ b & \text{if } c \text{ has the form } (A,b). \end{cases}$$

Proof. The proof is illustrated by the following diagram.



Let $E = B \backslash \operatorname{rng} f$, let D be the set of all ordered pairs (A,b) with $b \in E$, and let $C = A \cup D$. Thus C is as specified in the second part of the theorem. Now note:

(1) $A \cap D = 0$.

In fact, if $x \in A \cap D$, then x must have the form (A,b) for some $b \in B$, and then we have $(A,b) \in A$, which contradicts Theorem 3.19.

Because of (1), the following unambigously defines a function g with domain C: for any $c \in C$, let

$$g(c) = \begin{cases} f(c) & \text{if } c \in A, \\ b & \text{if } c \in D, \text{ and } c \text{ has the form } (A,b). \end{cases}$$

More precisely, to justify this definition using our axioms, let

$$g = f \cup \{(c, b) \in D \times B : c = (A, b)\}.$$

This gives the property of g stated in the second part of the theorem. The assertions of the theorem are now easy to check. $\square$

Exercises, chapter 3

1. Prove that if $f : A \to B$ and $X \subseteq A$, then $X \subseteq f^{-1}[f[X]]$.

2. Prove that if $f : A \to B$, then the following conditions are equivalent:
   (i) f is one-one.
   (ii) $X = f^{-1}[f[X]]$ for every $X \subseteq A$.

3. Prove that if $f : A \to B$ and $Y \subseteq B$, then $f[f^{-1}[Y]] \subseteq Y$.

4. Prove that if $f : A \to B$, then the following conditions are equivalent:
   (i) f maps onto B.
   (ii) $f[f^{-1}[Y]] = Y$ for every $Y \subseteq B$.

5. Prove that if $f : A \to B$ and $X \subseteq A$, then $f[X] = f[f^{-1}[f[X]]]$.

6. Prove that if $f : A \to B$ and $Y \subseteq B$, then $f^{-1}[Y] = f^{-1}[f[f^{-1}[Y]]]$.

7. Show that if $f : A \to B$ and $X \subseteq A$, then

$$B\backslash f[A\backslash X] = \{b \in B : f^{-1}[\{b\}] \subseteq X\}.$$

8. Prove that if $f : A \to B$ and $Y \subseteq B$, then $f^{-1}[B\backslash Y] = A\backslash f^{-1}[Y]$.

9. Prove that if $f : A \to B$ and $\langle Y_i : i \in I\rangle$ is a system of subsets of B, then $f^{-1}[\bigcup_{i\in I} Y_i] = \bigcup_{i\in I} f^{-1}[Y_i]$.

10. (Assuming usual knowledge of real numbers.) Define functions f, g, h as follows:
    $\mathrm{dmn}(f) = \{0, 1, 2, \ldots\}$; $f(m) = m + 1$ for every m in $\mathrm{dmn}(f)$.
    $\mathrm{dmn}(g) = \{x : x \text{ is a positive real number}\}$; $g(x) = \sqrt{x}$ for every $x \in \mathrm{dmn}(g)$.
    $\mathrm{dmn}(h) = \{x : x \text{ is a real number}\}$; $h(x) = x^2$ for every $x \in \mathrm{dmn}(h)$.

Describe the domains and the action of the six possible ways of composing these functions.

11. Define functions f, g, h so that the 27 different ways of composing the functions, including repeated functions, are all distinct.

12. Prove that if A and B are sets, then $\mathscr{P}(A) \cup \mathscr{P}(B) \subseteq \mathscr{P}(A \cup B)$.

13. Give an example in which equality in exercise 13 does not hold.

14. Suppose that A is a set, and f maps $\mathscr{P}(A)$ into $\mathscr{P}(A)$. Assume that

$$X \cap f(Y) \neq \emptyset \text{ iff } f(X) \cap Y \neq \emptyset, \text{ for all } X, Y \in \mathscr{P}(A).$$

Show that if $\langle B_i : i \in I \rangle$ is a system of subsets of A, then

$$f\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} f(B_i).$$

15. Suppose that $f : A \to B$. Show that the following conditions are equivalent:
   (i) f maps onto B.
   (ii) For all C, g, h, if $g : B \to C$, $h : B \to C$, and $g \circ f = h \circ f$, then $g = h$.

16. Suppose that $f : B \to C$. Show that the following conditions are equivalent:
   (i) f is one-one.
   (ii) For all A, g, h, if $g : A \to B$, $h : A \to B$, and $f \circ g = f \circ h$, then $g = h$.

17. Let A and B be given. Define $p_0 : A \times B \to A$ and $p_1 : A \times B \to B$ by setting $p_0(a, b) = a$ and $p_1(a, b) = b$ for all $a \in A$ and $b \in B$.

Show that for any C, q_0, q_1, if $q_0 : C \to A$ and $q_1 : C \to B$, then there is a unique $f : C \to A \times B$ such that $p_0 \circ f = q_0$ and $p_1 \circ f = q_1$.

18. Let A and B be given. Set $C = (A \times \{0\}) \cup (B \times \{1\})$. Define $i_0 : A \to C$ by $i_0(a) = (a, 0)$ for all $a \in A$, and define $i_1 : B \to C$ by $i_1(b) = (b, 1)$ for all $b \in B$.

Show that for any D, j_0, j_1, if $j_0 : A \to D$ and $j_1 : B \to D$, then there is a unique $f : C \to D$ such that $f \circ i_0 = j_0$ and $f \circ i_1 = j_1$.

19. Justify the definition of ${}^A B$.

20^{ch}. Show that if $\langle X_i : i \in I \rangle$ is a system of nonempty sets, then $\prod_{i \in I} X_i \neq \emptyset$. Hint: Consider the set $\{(x, i) : i \in I, x \in X_i\}$.

21^{ch}. If $\langle \langle A_{ij} : i \in I \rangle : j \in J \rangle$ is a system of sets, then

$$\prod_{i \in I} \bigcup_{j \in J} A_{ij} = \bigcup_{f \in {}^I J} \prod_{i \in I} A_{if(i)}.$$

22^{ch}. If $\langle \langle A_{ij} : i \in I \rangle : j \in J \rangle$ is a system of sets, then

$$\bigcap_{i \in I} \bigcup_{j \in J} A_{ij} = \bigcup_{f \in {}^I J} \bigcap_{i \in I} A_{if(i)}.$$

23. Let $\langle A_i : i \in I \rangle$ be a system of sets. For each $j \in I$, define $p_j : \prod_{i \in I} A_i \to A_j$ by setting $p_j(x) = x(j)$ for all $x \in \prod_{i \in I} A_i$

Show that for any B and any system $\langle q_i : i \in I \rangle$ such that $q_i : B \to A_i$ for all $i \in I$, there is a unique $f : B \to \prod_{i \in I} A_i$ such that $p_j \circ f = q_j$ for all $j \in I$.

24. Let $\langle A_i : i \in I \rangle$ be a system of sets. Define

$$B = \{(a, i) : i \in I, a \in A_i\}.$$

For each $j \in I$, define $s_j : A_j \to B$ by setting $s_j(a) = (a, j)$ for any $a \in A_j$.

Show that for any C and any system $\langle t_i : i \in I \rangle$ such that $q_i : A_i \to C$ for all $i \in I$, there is a unique $f : B \to C$ such that $f \circ s_i = t_i$ for all $i \in I$.

25. Assume that $f : A \to A$ and $g : A \to A$. Show that the following relation is a function (not necessarily with domain A):

$$(A \times A) \backslash ([(A \times A) \backslash f] | [(A \times A) \backslash g]).$$

(See exercise 15 in chapter 2 for the definition of $R|S$ for relations R, S.)

26. Suppose that $f : A \to B$, f maps onto B, $g : A \to C$, and for all $x, y \in A$, if $f(x) = f(y)$ then $g(x) = g(y)$. Show that there is an $h : B \to C$ such that the diagram

$$\begin{tikzcd} A \arrow[rr, "f"] \arrow[dr, "g"'] & & B \arrow[dl, dotted, "h"] \\ & C & \end{tikzcd}$$

commutes, i.e., $g = h \circ f$.

27. Suppose that $f : A \to B$, and f maps onto B. Show that $\langle f^{-1}[\{b\}] : b \in B \rangle$ is one-one and maps B into $\mathscr{P}A$.

28. Suppose that the following diagram commutes.

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "h"'] & Y \arrow[d, "g"] \\ T \arrow[r, "k"'] & Z \end{tikzcd}$$

That is, assume that $g \circ f = k \circ h$. Show then that the following three conditions are equivalent:

(i) $f[h^{-1}[\{t\}]] = g^{-1}[\{k(t)\}]$ for all $t \in T$.

(ii) $h[f^{-1}[\{y\}]] = k^{-1}[\{g(y)\}]$ for all $y \in Y$.

(iii) If $y \in Y$, $t \in T$, and $g(y) = k(t)$, then there is an $x \in X$ such that $y = f(x)$ and $t = h(x)$.

4. Equivalence relations and partitions

A construction frequently found in mathematics is a rigorous expression of abstraction, in which certain relatively concrete objects are lumped together because of some common features. Thus one gathers together all green objects to make up the abstract concept of green. More mathematically, one considers all lines in the plane which have the same slope to determine a direction. We want to give in this chapter the general framework for this process; later in the notes we will encounter several important special cases. The basic notions for this discussion are equivalence relations and partitions. For equivalence relations, we have several special properties of binary relations which are combined to give the notion:

• A binary relation R is *symmetric* iff for all x, y, if $(x, y) \in R$, then $(y, x) \in R$.

• A binary relation R is *transitive* iff for all x, y, z, if $(x, y) \in R$ and $(y, z) \in R$, then $(x, z) \in R$.

• A binary relation R is *reflexive on* A iff for all $x \in A$, if $(x, x) \in R$.

• An *equivalence relation on* A is a relation $R \subseteq A \times A$ which is symmetric, transitive, and reflexive on A.

A subtlety which is sometimes overlooked is that the notions of symmetry and transitivity are properties of the relation R alone, while those of reflexivity and equivalence relation refer not only to R, but to some set A. Frequently when talking about equivalence relations, we use some suggestive symbol like $\sim$ or $\equiv$ instead of R, and write $a \sim b$ or $a \equiv b$ instead of $(a, b) \in \sim$ or $(a, b) \in \equiv$.

Here are some examples of equivalence relations.

(1) Define $a \sim b$ iff a and b are integers, and $a - b$ is divisible by 5. Then $\sim$ is an equivalence relation on the set of all integers, as is easily checked. (Temporarily we assume knowledge of the integers, although they have not yet been introduced within our system of set theory.) This equivalence relation is a special case of the general modular relation studied in algebra and number theory.

(2) Define $(a, b) \equiv (c, d)$ iff a, b, c, d are integers, $b, d \neq 0$, and $ad = bc$. Then $\equiv$ is an equivalence relation on the set $\{(a, b) : a, b \text{ are integers and } b \neq 0\}$. This equivalence relation is used to define the rational numbers from the integers.

(3) Let $f : A \to B$, and define $a \cong b$ iff $a, b \in A$ and $f(a) = f(b)$. Then $\cong$ is an equivalence relation on A. This relation is important in the general theory of algebras, being the general equivalent of the notion of normal subgroup or ideal in a suitable framework.

(4) Let $A \sim B$ iff A and B are similar triangles in the plane. This is an equivalence relation on the set of all triangles in the plane. It can be taken as the exact notion of similarity.

If $\sim$ is an equivalence relation on A, and $a \in A$, we define the *equivalence class of a under* $\sim$ to be

$$[a]_\sim = \{b \in A : a \sim b\}.$$

The most important properties of this notion are given in the following theorem.

Proposition 4.1. *If $\sim$ is an equivalence relation on A and $a, b \in A$, then:*
(i) $a \in [a]_\sim$.
(ii) $a \sim b$ iff $[a]_\sim = [b]_\sim$.
(iii) if $[a]_\sim \neq [b]_\sim$, then $[a]_\sim \cap [b]_\sim = \emptyset$.

Proof. (i) holds since $a \sim a$. (ii): first assume that $a \sim b$. Suppose that $x \in [a]_\sim$. Then $a \sim x$, so $b \sim a \sim x$ and so $b \sim x$; so $x \in [b]_\sim$. This proves $\subseteq$. The other inclusion is proved similarly.

Conversely, assume that $[a]_\sim = [b]_\sim$. Then $b \in [b]_\sim = [a]_\sim$, so $a \sim b$.

(iii): if $[a]_\sim \cap [b]_\sim \neq \emptyset$, choose $x \in [a]_\sim \cap [b]_\sim$. So $a \sim x$ and $b \sim x$, hence $a \sim x \sim b$ and so $a \sim b$. So by (ii), $[a]_\sim = [b]_\sim$. □

Note that property (ii) in this proposition gives the essential mathematical procedure of "identifying" objects which are equivalent. They are not actually identified, but all similar objects are gathered together into one class.

This notion of equivalence class suggests an alternative way of looking at equivalence relations. This alternative way is very important in actual applications of equivalence relations.

A *partition* of a set A is a collection $\mathscr{A}$ of subsets of A such that each member of $\mathscr{A}$ is nonempty, the members of $\mathscr{A}$ are pairwise disjoint, and $\bigcup_{X \in \mathscr{A}} X = A$. Now we can state the important theorem which says that equivalence relations and partitions are equivalent concepts:

Theorem 4.2. (Equivalence-Partition theorem) *There is a one-one correspondence between equivalence relations on a set A and partitions of A. Specifically:*

(i) For each equivalence relation $\sim$ on A, the set

$$A/\sim \stackrel{\text{def}}{=} \{[a]_\sim : a \in A\}$$

is a partition of A.

(ii) If $\mathscr{A}$ is a partition of A, then

$$\equiv_{\mathscr{A}} \stackrel{\text{def}}{=} \{(a, b) \in A \times A : \exists X \in \mathscr{A}(a, b \in X)\}$$

is an equivalence relation on A.

(iii) If $\sim$ is an equivalence relation on A, then $\equiv_{A/\sim}$ is equal to $\sim$.

(iv) If $\mathscr{A}$ is a partition of A, then $A/\equiv_{\mathscr{A}}$ is equal to $\mathscr{A}$.

Proof. (i): Assume that $\sim$ is an equivalence relation on A. Clearly $[a]_\sim \subseteq A$ for each $a \in A$. Since $a \in [a]_\sim$ by 4.1(i), each equivalence class is nonempty. By 4.1(iii), the members of $A/\sim$ are pairwise disjoint. Finally, $\bigcup(A/\sim) = A$ by the first statements of this proof. Hence $A/\sim$ is a partition of A.

(ii): Assume that $\mathscr{A}$ is a partition of A. Clearly then $\equiv_{\mathscr{A}} \subseteq A \times A$. Suppose that $a \equiv_{\mathscr{A}} b$. Then there is an $X \in \mathscr{A}$ such that $a, b \in X$; it follows that also $b \equiv_{\mathscr{A}} a$. Suppose that $a \equiv_{\mathscr{A}} b \equiv_{\mathscr{A}} c$. Then there exist $X, Y \in \mathscr{A}$ such that $a, b \in X$ and $b, c \in Y$. So $b \in X \cap Y$. Since distinct members of $\mathscr{A}$ are disjoint, it follows that $X = Y$. Thus

$a, c \in X$, so $a \equiv_{\mathscr{A}} c$. Finally, if $a \in A$, we can choose $X \in \mathscr{A}$ with $a \in X$, since $\bigcup_{Y \in \mathscr{A}} Y = A$. Hence $a \equiv_{\mathscr{A}} a$. We have thus checked that $\equiv_{\mathscr{A}}$ is an equivalence relation on A.

(iii): Assume that $\sim$ is an equivalence relation on A. Suppose that $a \equiv_{A/\sim} b$. Then there is a member X of $A/\sim$ such that $a, b \in X$. By the definition of $A/\sim$, there is a $c \in A$ such that $X = [c]_\sim$. Hence $a, b \in [c]_\sim$. So $c \sim a$ and $c \sim b$, so we easily get $a \sim b$. This shows that $\equiv_{A/\sim}$ is a subset of $\sim$.

Conversely, suppose that $a \sim b$. Then $a, b \in [a]_\sim \in A/\sim$, so $a \equiv_{A/\sim} b$. This shows that $\sim$ is a subset of $\equiv_{A/\sim}$, and finishes the proof of (iii).

(iv): Suppose that $\mathscr{A}$ is a partition of A. Let $X \in A/\equiv_{\mathscr{A}}$. Say $X = [a]_{\equiv_{\mathscr{A}}}$. Since $\bigcup_{Y \in \mathscr{A}} Y = A$, choose $Y \in \mathscr{A}$ such that $a \in Y$. We claim that $X = Y$; this will show that $X \in \mathscr{A}$. To prove that $X = Y$, let $x \in X$. So $x \in [a]_{\equiv_{\mathscr{A}}}$. Hence $a \equiv_{\mathscr{A}} x$, and this means that there is a $Z \in \mathscr{A}$ such that $a, x \in Z$. Since we already know that $a \in Y \in \mathscr{A}$, and the elements of $\mathscr{A}$ are pairwise disjoint, it follows that $Y = X$. So $x \in Y$. This shows that $X \subseteq Y$. Conversely, suppose that $y \in Y$. Then $a, y \in Y$, so $a \equiv_{\mathscr{A}} y$, and hence $y \in [a]_{\equiv_{\mathscr{A}}} = X$. So we have shown that $X \subseteq Y$, finishing the proof of our claim that $X = Y$. As mentioned, this proves that $X \in \mathscr{A}$, and finishes the proof that $A/\equiv_{\mathscr{A}}$ is a subset of $\mathscr{A}$.

Conversely, suppose that $X \in \mathscr{A}$. Since the members of $\mathscr{A}$ are nonempty, choose $a \in X$. We claim that $X = [a]_{\equiv_{\mathscr{A}}}$; this will prove that $X \in A/\equiv_{\mathscr{A}}$. To prove the claim, suppose that $x \in X$. then $a, x \in X$, so $a \equiv_{\mathscr{A}} x$, and hence $x \in [a]_{\equiv_{\mathscr{A}}}$. Conversely, suppose that $x \in [a]_{\equiv_{\mathscr{A}}}$. Thus $a \equiv_{\mathscr{A}} x$, so there is a $Y \in \mathscr{A}$ such that $a, x \in Y$. Since $a \in X$, it follows that $X = Y$. Hence $x \in X$. This finishes the proof of the claim, which finishes the proof that $\mathscr{A}$ is a subset of $A/\equiv_{\mathscr{A}}$, which finishes the proof of (iv). □

Exercises, chapter 4

1. The definitions of symmetric, transitive, and reflexive on A give rise to 8 possibilities for their truth and falsity; give examples for these eight possibilities. (One for all three being true, one for all three being false, etc.)

2. Prove that if R is symmetric and transitive, then R is reflexive on $\mathrm{dmn}(R)$.

3. Let A be a set, R an equivalence relation on A, and E the collection of all equivalence classes of A under R. For each equivalence relation S on A such that $R \subseteq S$ define

$$F(S) = \{x : \exists a, b \in A((a, b) \in S \text{ and } x = ([a], [b]))\}.$$

Here $[a]$ and $[b]$ refer to equivalence classes under R.

Show that F is a one-one function mapping the set $\mathscr{A}$ of all equivalence relations on A which contain R onto the set of all equivalence relations on E.

4. Suppose that $\mathscr{E}$ is a nonempty set of equivalence relations on a set A. Prove that $\bigcap \mathscr{E}$ is also an equivalence relation on A.

5. Prove that if R and S are equivalence relations on A, then $R \subseteq S$ iff every member of A/R is a subset of some member of A/S. We are using the notation of Theorem 4.2 here.

6. (Continuing exercise 4) Suppose that $\mathscr{E}$ is a nonempty set of equivalence relations on a set A. Prove that the partition corresponding to $\bigcap \mathscr{E}$ is $\{\bigcap_{R \in \mathscr{E}} [a]_R : a \in A\}$.

7. Let $A = \{a, b, c, d\}$ with a, b, c, d distinct. Let R be the equivalence relation on A defined as follows:

$$\begin{aligned} R = \{&(a,a),(b,b),(c,c),(d,d), \\ &(a,b),(b,a),(c,d),(d,c)\} \end{aligned}$$

Let $f : A \to A$ be defined by:

$$f(a) = b; \quad f(b) = c; \quad f(c) = d; \quad f(d) = a.$$

Let E be the set of all equivalence classes of A under R. Prove that there does not exist $g : E \to E$ such that $g([x]) = [f(x)]$ for all $x \in A$.

8. Show that for every equivalence relation $\sim$ on A there is a set B and a function $f : A \to B$ such that $\sim$ is the relation $\{(a, b) : a, b \in A \text{ and } f(a) = f(b)\}$.

5. Ordering

We introduce the major kinds of orderings which will be used later in these notes:

• A *quasi-ordering* is a pair $(A, \leq)$ consisting of a nonempty set A and a relation $\leq$ contained in $A \times A$ such that $\leq$ is reflexive on A and transitive.

• A relation R is *irreflexive* on A iff there is no element $a \in A$ such that $(a, a) \in R$.

• A *partial ordering* or *poset* is a pair $(A, <)$ consisting of a nonempty set A and a relation $<$ contained in $A \times A$ such that $<$ is irreflexive on A and transitive.

• A relation R is *antisymmetric* iff for any x, y, if $(x, y) \in R$ and $(y, x) \in R$ then $x = y$.

• A *weak partial ordering* or *weak poset* is a pair $(A, \leq)$ consisting of a nonempty set A and a relation $\leq$ contained in $A \times A$ such that $\leq$ is reflexive on A, antisymmetric, and transitive.

• A *linear ordering* or *simple ordering* is a partial ordering $(A, <)$ such that for any $a, b \in A$, one of $a < b$, $a = b$, $b < a$ holds.

• If $(A, <)$ is a simple ordering and $X \subseteq A$, then an element $a \in X$ is the *smallest* or *least* element of X iff $a < b$ for any element $b \in X$ which is different from X. Clearly this least element of X is unique if it exists.

• A *well ordering* is a simple ordering $(A, <)$ such that every nonempty subset of A has a least element.

In all of these cases we sometimes refer to the set A as being a quasi-ordered set, or partially ordered set, etc., or to the relation $<$ as being a well-ordering, a simple ordering, etc. These usages are to be taken as shorthand for the full notation above.

We give some examples of orderings.

1. If A is any collection of sets, then $(A, <)$ is a partial ordering, where we define $a < b$ iff $a, b \in A$ and $a \subset b$. Also $(A, \leq)$ is a weak partial ordering, where we define $a \leq b$ iff $a, b \in A$ and $a \subseteq b$.

2. If $(A, \leq)$ is a weak partial ordering, then $(A \times A, \preceq)$ is a quasiordering, where we define $(a, b) \preceq (c, d)$ iff $a, b, c, d \in A$ and $a \leq c$. Note that $\preceq$ is not antisymmetric if A has more than one element, and so $(A \times A, \preceq)$ is not a weak partial ordering.

3. If A is the usual set of integers and $<$ is the usual order relation on A, then $(A, <)$ is a simple ordering. Similarly for A the collection of real numbers.

4. If A is the set of nonnegative integers and $<$ is the usual ordering of A, then $(A, <)$ is a well-ordering.

Recall from the introduction that we are using prior knowledge of mathematics, like knowledge of integers and real numbers, just to illustrate notions. Some of these notions, such as that of natural numbers, will be introduced formally later.

We now give some elementary properties of these notions.

Proposition 5.1. *(i) If $(A, <)$ is a partial ordering, define $a \leq b$ iff $a, b \in A$ and ($a < b$ or $a = b$). Then $(A, \leq)$ is a weak partial ordering.*

(ii) If $(A, \leq)$ is a weak partial ordering, define $a < b$ iff $a, b \in A$ and ($a \leq b$ and $a \neq b$). Then $(A, <)$ is a partial ordering.

(iii) The procedures described in (i) and (ii) are inverses of one another.

Proof. Left to an exercise.

Proposition 5.2. *Suppose that $(A, \leq)$ is a quasiordering. Define $a < b$ iff $a, b \in A$, $a \leq b$, and $b \not\leq a$. Then $(A, <)$ is a partial ordering.*

Proof. Since $a \leq a$ for any $a \in A$, it follows that $a \not< a$. Suppose that $a < b < c$. Hence $a \leq b \leq c$, so $a \leq c$. Suppose that $c \leq a$. Then $b \leq c \leq a$, so $b \leq a$, contradicting $a < b$. Hence $a < c$, as desired. □

Proposition 5.3. *Suppose that $(A, \leq)$ is a quasiordering. Define $a \equiv b$ iff $a \leq b \leq a$. Then $\equiv$ is an equivalence relation on A. There is a partial order $<$ on the set B of all equivalence classes such that for all $a, b \in A$, $[a]_\equiv < [b]_\equiv$ iff $[a]_\equiv \neq [b]_\equiv$ and $a \leq b$.*

Proof. For any $a \in A$ we have $a \leq a \leq a$, so $a \equiv a$. If $a \equiv b$, then $b \leq a \leq b$, so $b \equiv a$. Suppose that $a \equiv b \equiv c$. Then $a \leq b \leq a$ and $b \leq c \leq b$, so $a \leq b \leq c$ and hence $a \leq c$, and $c \leq b \leq a$ and hence $c \leq a$. So $a \equiv c$.

This shows that $\equiv$ is an equivalence relation on A. Now we define

$$\begin{aligned} < \text{ is the set } \{(r, s) : r, s \in B,\ r \neq s \text{ and there are } a, b \in A \\ \text{such that } r = [a]_\equiv,\ s = [b]_\equiv, \text{ and } a < b\}. \end{aligned}$$

Suppose that $a, b \in A$. If $[a]_\equiv < [b]_\equiv$, then by definition, $[a]_\equiv \neq [b]_\equiv$ and there exist $c, d \in A$ such that $[a]_\equiv = [c]_\equiv$, $[b]_\equiv = [d]_\equiv$, and $c < d$. Hence $a \equiv c$ and $b \equiv d$. So $a \leq c \leq a$ and $b \leq d \leq b$, hence $a \leq c \leq d \leq b$, so that $a \leq b$.

Conversely, suppose that $[a]_\equiv \neq [b]_\equiv$ and $a \leq b$. Then $[a]_\equiv < [b]_\equiv$ by definition.

Clearly $<$ is a subset of $B \times B$, and is irreflexive on B. Suppose that $[a]_\equiv < [b]_\equiv < [c]_\equiv$. Then $[a]_\equiv \neq [b]_\equiv \neq [c]_\equiv$, and $a \leq b \leq c$. Hence $a \leq c$. If $a = c$, then $a \equiv b$, so $[a]_\equiv = [b]_\equiv$, contradiction. So $a \neq c$, and hence $a < c$. □

The following theorem shows that example 1 above is typical for partial orderings, and could replace the more abstract definition for many purposes.

Theorem 5.4. *Any partial ordering is isomorphic to a partial ordering of the kind described in example 1. That is, if $(A, <)$ is a partial ordering, then there exist a partial ordering $(B, \prec)$ and a bijection f from A onto B such that $\prec$ consists of all pairs $b, c \in B$ such that $b \subset c$, and for any $a, a' \in A$, $a < a'$ iff $f(a) \subset f(a')$.*

Proof. For any $a \in A$, let $f(a) = \{c \in A : c \leq a\}$, where $\leq$ is defined as in Proposition 5.1(a): $c \leq a$ iff ($c < a$ or $c = a$). Suppose that $a, a' \in A$. If $a < a'$, then for any $c \in f(a)$ we have $c \leq a$, and hence $c \leq a'$, so that $c \in f(a')$. Also, $a' \leq a'$ but $a' \not\leq a$, so $a \in f(a') \backslash f(a)$. Thus $f(a) \subset f(a')$.

Conversely, suppose that $f(a) \subset f(a')$. Since $a \leq a$, we have $a \in f(a)$, and hence $a \in f(a')$, and so $a \leq a'$. There is some $c \in f(a') \backslash f(a)$, so that $c \leq a'$ but $c \not\leq a$. Hence $a \neq a'$, so $a < a'$.

Now f is one-one, since if $a \neq a'$, then $a \not\leq a'$ or $a' \not\leq a$. Say by symmetry that $a \not\leq a'$. Then $a \in f(a) \backslash f(a')$, so that $f(a) \neq f(a')$.

Thus the conclusion of the theorem holds, with $B = \mathrm{rng}(f)$. □

The following characterization of well-orderings is important for later purposes.

Theorem 5.5. *Suppose that $(A, <)$ is a simple ordering. Then the following conditions are equivalent:*

(i) $(A, <)$ is a well-ordering.

(ii) $(A, <)$ satisfies the principle of complete induction. That is, for any subset X of A, the following hypothesis (a) leads to the conclusion that $X = A$:

(a) For all $a \in A$, if $x \in X$ for all $x < a$, then $a \in X$.

Proof. (i)⇒(ii): Assume (i), and suppose that (ii)(a) holds, but suppose that $X \neq A$. Let a be the least element of $A \backslash X$. Then $x \in X$ for all $x < a$, so by (a), $a \in X$, contradiction. Hence $X = A$.

(ii)⇒(i): Assume (ii). Suppose that Y is a nonempty subset of A, but Y does not have a least element. Let $X = A \backslash Y$; we are going to apply (ii)(a) to X. Suppose that $a \in A$, and $x \in X$ for all $x < a$. If $a \in Y$, then it is the least element of Y, since $x \in X = A \backslash Y$ for all $x < a$; contradiction. So $a \in A \backslash Y = X$, verifying (a). Hence by (ii), $A = X$, and so Y is empty, contradiction. □

1. Prove Proposition 5.1.

2. Let $(A, \leq)$ be a weak partial ordering, and let B be any set, with b some element of B. Define $\preceq$ to be the collection of all ordered pairs (f, g) such that $f, g : B \to A$ and $f(b) \leq g(b)$. Show that $({}^{B}A, \preceq)$ is a quasi-ordering

3. With the notation of exercise 2, show that if both A and B have at least two elements, then $({}^{B}A, \preceq)$ is not a weak partial ordering.

4. Let $(A, <)$ be a partial ordering, and form $\leq$ as in Proposition 5.1(i). (Thus $(A, \leq)$ is a weak partial order.) Now take a set B and an element b of it, and form the quasi-order as in exercise 2. Apply the procedure of Theorem 5.3 to $({}^{B}A, \preceq)$ to form a partial order $(C, \ll)$. Show that $(A, <)$ is isomorphic to $(C, \ll)$. (See 5.4 for the notion of isomorphism).

5. Let $(A, <)$ and $(B, \prec)$ be partial orders. Define $(a, b) \ll (c, d)$ iff $a, c \in A$, $b, d \in B$, $a < c$, and $b \prec d$. Show that $(A \times B, \ll)$ is a partial order.

6. Give examples of well-orders $(A, <)$ and $(B, \prec)$ such that $(A \times B, \ll)$, as defined in exercise 5, is not a linear order.

7. Let $(A, <)$ and $(B, \prec)$ be partial orders. Redefine $(a, b) \ll (c, d)$ iff $a, c \in A$, $b, d \in B$, and either $a < c$, or $a = c$ and $b \prec d$. Show that $(A \times B, \ll)$ is a partial order. (This is the dictionary, or lexicographic, order.)

8. With the definition in exercise 7, show that $(A \times B, \ll)$ is a simple order if both $(A, <)$ and $(B, \prec)$ are, and it is a well-order if both $(A, <)$ and $(B, \prec)$ are.