Sets, Logic and Categories
Solutions to Exercises: Chapter 1

1.1 Show that the empty set is a subset of every set.

Let x be any set. Then for any set z, the implication $(z \in \emptyset) \Rightarrow (z \in x)$ is true, since $(z \in \emptyset)$ is false; thus $\emptyset \subseteq x$.

1.2 Which of the following equations are true? If the equation is not true, is one side a subset of the other?

(a) $\bigcup \mathcal{P} X = X$.

(b) $\mathcal{P} \bigcup X = X$.

(c) $\bigcup \mathcal{P} X = \mathcal{P} \bigcup X$.

(d) $\mathcal{P}(X \times Y) = \mathcal{P} X \times \mathcal{P} Y$.

(e) $\mathcal{P}(X \cup Y) = \mathcal{P} X \cup \mathcal{P} Y$.

(a) True. If $x \in X$, then $\{x\} \in \mathcal{P} X$, and so $x \in \bigcup \mathcal{P} X$. Conversely, if $x \in \bigcup \mathcal{P} X$, then $x \in Y$ for some $Y \in \mathcal{P} X$; then $Y \subseteq X$, and so $x \in X$.

(b) False. If $X = \{\{1\}\}$, then $\bigcup X = \{1\}$, and $\mathcal{P} \bigcup X = \{\emptyset, \{1\}\}$.

It is true that $X \subseteq \mathcal{P} \bigcup X$. For take $x \in X$. Then every element of x is contained in $\bigcup X$, so $x \subseteq \bigcup X$, that is, $x \in \mathcal{P} \bigcup X$.

(c) From (a) and (b) we see that these two sets are not equal but $\bigcup \mathcal{P} X \subseteq \mathcal{P} \bigcup X$.

(d) False. If X has m elements and Y has n elements, then $\mathcal{P}(X \times Y)$ has 2^{mn} elements while $\mathcal{P} X \times \mathcal{P} Y$ has 2^{m+n} elements. If $m = n = 1$, then the second is greater than the first, while if $m = n = 3$, the first is greater than the second. So neither of the sets can be a subset of the other for all choices of X and Y. (You should also try to give specific examples to refute both inclusions.)

(e) False. If $X = \{1\}$ and $Y = \{2\}$, then the set $\{1, 2\}$ belongs to $\mathcal{P}(X \cup Y)$ but not to $\mathcal{P} X$ or $\mathcal{P} Y$.

It is true that $\mathcal{P} X \cup \mathcal{P} Y \subseteq \mathcal{P}(X \cup Y)$. For every subset of either X or Y is a subset of $X \cup Y$.

1.3 Prove that each of the following is *not* a suitable definition of the ordered pair (x,y):

(a) $(x,y) = \{x, \{y\}\}$.

(b) $(x,y) = \{\{x\}, \{y\}\}$.

(a) In this case we would have $(\{1\}, 2) = (\{2\}, 1)$.

(b) In this case $(1, 2) = (2, 1)$.

1.4 Which of the following would be a suitable definition of the ordered triple (x,y,z)?

(a) $(x,y,z) = \{(x,y),(y,z)\}$.

(b) $(x,y,z) = ((x,y),(y,z))$.

(c) $(x,y,z) = \{\{x\},\{x,y\},\{x,y,z\}\}$.

(a) No: $(1,2,1) = (2,1,2)$.

(b) Yes: if $((x_1,y_1),(y_1,z_1)) = ((x_2,y_2),(y_2,z_2))$, then $(x_1,y_1) = (x_2,y_2)$ and $(y_1,z_1) = (y_2,z_2)$, so $x_1 = x_2$, $y_1 = y_2$ and $z_1 = z_2$.

(c) No: $(1,1,2) = (1,2,2)$.

1.5 Let G be a group, and X a set. An *action* of G on X is a function $\mu : X \times G \to X$ satisfying the rules

- $\mu(x,gh) = \mu(\mu(x,g),h)$ for all $g,h \in G$, $x \in X$;
- $\mu(x,1) = x$ for all $x \in X$, where 1 is the identity element of G.

(a) Prove that

- $\mu(\mu(x,g),g^{-1}) = \mu(\mu(x,g^{-1}),g) = x$ for all $x \in X$, $g \in G$.

(b) Define a relation $\sim$ on X by the rule that $x \sim y$ if and only if $\mu(x,g) = y$ for some $g \in G$. Show that $\sim$ is an equivalence relation.

(c) Show that each of the following equivalence relations on the set of all $m \times n$ real matrices arises from a group action:

- row-equivalence (A and B are row-equivalent if some sequence of elementary row operations transforms A to B);
- equivalence (A and B are equivalent if some sequence of elementary row and column operations transforms A to B);
- conjugacy, for $m = n$ (A and B are conjugate if $B = P^{-1}AP$ for some invertible matrix P);
- congruence, for $m = n$ (A and B are congruent if $B = P^\top AP$ for some invertible matrix P, where $P^\top$ is the transpose of P).

(d) If H is a subgroup of G, and H acts on G by *right multiplication* (that is, $\mu(x,h) = xh$), then the orbits of H are its *left cosets* in G.

(e) If G acts on itself by *conjugation* (that is, $\mu(x,g) = g^{-1}xg$), then the orbits of G are the *conjugacy classes*.

(a) $\mu(\mu(x,g),g^{-1}) = \mu(x,gg^{-1}) = \mu(x,1) = x$. The other equation is similar.

(b) The reflexive law follows from the second axiom for an action; the symmetric law from the result of (a); and the transitive law from the first axiom.

(c) Let $\mathrm{GL}(n,\mathbb{R})$ denote the group of invertible $n \times n$ real matrices. Now the appropriate groups and actions are:

- $G = \mathrm{GL}(m,\mathbb{R})$, $\mu(A,P) = P^{-1}A$ for $P \in G$;
- $G = \mathrm{GL}(m,\mathbb{R}) \times \mathrm{GL}(n,\mathbb{R})$, $mu(A,(P,Q)) = P^{-1}AQ$ for $(P,Q) \in G$;
- $G = \mathrm{GL}(m,\mathbb{R})$, $\mu(A,P) = P^{-1}AP$ for $P \in G$;
- $G = \mathrm{GL}(m,\mathbb{R})$, $\mu(A,P) = P^{\top}AP$ for $P \in G$.

(d) Verifying that this is an action is straightforward. The orbit of the element $x \in G$ is $\{xh : h \in H\} = xH$.

(e) Verifying that this is an action is straightforward. This is the definition of conjugacy classes.

1.6 Suppose that μ is an action of the group G on the set X, as defined in the preceding exercise. Show that, for any $g \in G$, the function $x \mapsto \mu(x,g)$ is a bijection from X to X.

Let f_g be the function $x \mapsto \mu(x,g)$. To show that f_g is injective, suppose that $f_g(x) = f_g(y)$, that is, $\mu(x,g) = \mu(y,g)$. Then

$$x = \mu(x,1) = \mu(\mu(x,g),g^{-1}) = \mu(\mu(y,g),g^{-1}) = \mu(y,1) = y.$$

To show that f_g is surjective, take $z \in X$, and let $x = \mu(x,g^{-1})$. Then

$$z = \mu(z,1) = \mu(\mu(z,g^{-1}),g) = \mu(x,g) = f_g(x),$$

as required.

1.7 Show that the composition of injective functions is injective, and the composition of surjective functions is surjective.

Let $f : X \to Y$ and $g : Y \to Z$ be functions.

Suppose that f and g are injective, and that $(g \circ f)(x) = (g \circ f)(x')$ for $x,x' \in X$. That is, $g(f(x)) = g(f(x'))$. By the injectivity of g, we have $f(x) = f(x')$; and by the injectivity of f, we have $x = x'$. Thus $g \circ f$ is injective.

Now suppose that f and g are surjective. Choose $z \in Z$. By injectivity of g, there exists $y \in Y$ with $g(y) = z$; then by injectivity of f, there exists $x \in X$ with $f(x) = y$. Thus, $(g \circ f)(x) = g(f(x)) = g(y) = z$, and $g \circ f$ is surjective.

1.8 Let $X \neq \emptyset$, let $f : X \to Y$ be a function, and let i_X and i_Y be the identity functions on X and Y respectively. Prove that

(a) f is injective if and only if there exists a function $g : Y \to X$ such that $f \circ g = i_X$;

(b) f is surjective if and only if there exists a function $h : Y \to X$ such that $h \circ f = i_Y$.

Where (if anywhere) have you used the Axiom of Choice in this proof?

(a) Let f be injective. Define $g : Y \to X$ as in Theorem 1.8. Then $f \circ g = i_X$.

Conversely, suppose that $f \circ g = i_X$, and let $f(x_1) = f(x_2)$. Then $x_1 = g(f(x_1)) = g(f(x_2)) = x_2$. So f is injective.

(b) Let f be surjective. For each $y \in Y$ choose $x \in X$ with $f(x) = y$ (here the Axiom of Choice is used), and define $h : Y \to X$ by $h(y) = x$. Then $h \circ f = i_Y$.

Conversely, suppose that $h \circ f = i_Y$, and choose $y \in Y$. Then f maps $h(y)$ to y; so f is surjective.

1.9 Let $X \neq \emptyset$ and let $f : X \to Y$ be a function.

(a) Prove that f is injective if and only if $h_1 \circ f = h_2 \circ f$ implies $h_1 = h_2$, for any two functions $h_1, h_2 : Y \to X$.

(b) Prove that f is surjective if and only if $f \circ g_1 = f \circ g_2$ implies $g_1 = g_2$, for any two functions $g_1, g_2 : Y \to X$.

Where (if anywhere) have you used the Axiom of Choice in this proof?

(a) Let f be injective and $h_1 \circ f = h_2 \circ f$. Suppose that $h_1 \neq h_2$. Then there exists $y \in Y$ with $h_1(y) \neq h_2(y)$, whence $f(h_1(y)) \neq f(h_2(y))$ (since f is injective), a contradiction.

Conversely, suppose that f is not injective; let $f(x_1) = f(x_2)$. Define $h_1, h_2 : Y \to X$ by $h_i(y) = x_i$ for all $y \in Y$ ($i = 1, 2$). Then $f(h_1(y)) = f(h_2(y))$ for all $y \in Y$, that is, $h_1 \circ f = h_2 \circ f$, but $h_1 \neq h_2$.

(b) Let f be surjective and $f \circ g_1 = f \circ g_2$. That is, $g_1(f(x)) = g_2(f(x))$ for all $x \in X$. For every $y \in Y$, there exists $x \in X$ with $f(x) = y$; thus, $g_1(y) = g_2(y)$ for all $y \in Y$, or $g_1 = g_2$.

Conversely, suppose that f is not surjective; suppose that $y \in Y$ is such that no element $x \in X$ satisfies $f(x) = y$. Let $g_1 : Y \to X$ be any function, and $g_2 : Y \to X$ a function which agrees with g_1 everywhere except at y, with $g_1(y) \neq g_2(y)$. Then $f \circ g_1 = f \circ g_2$ but $g_1 \neq g_2$.

The Axiom of Choice is not used, but in (b) we do need to assume that X has more than one element.

1.10 Let R be a relation between X and Y. Define the *converse* of R to be the relation between Y and X defined by reversing all the pairs in R:

$$R^* = \{(y, x) : (y, x) \in R\}.$$

Show that the converse of a function f is a function if and only if f is bijective (in which case f^* is the inverse of f).

Suppose that f is a function whose converse is a function. That means, for any $y \in Y$, there is exactly one $x \in X$ such that $(y, x) \in f^*$ (that is, such that $y = f(x)$). This shows that f is both injective and surjective, and f^* is its inverse.

Conversely, if f is a bijection, then it has an inverse function, which is its converse (as a relation).

1.11 Let X and Y be finite sets, with m and n elements respectively. How many elements are there in each of the following sets?

(a) $\mathcal{P}X$.

(b) $X \times Y$.

(c) The set of relations from X to Y.

(d) The set $\prod_{y \in Y} X_y$, where $X_y = X$ for all $y \in Y$.

(a) 2^m; (b) mn; (c) 2^{mn}; (d) m^n.

1.12 Show that

(a) any finite partially ordered set has a minimal element;

(b) any two (strict) total orders on a finite set are isomorphic;

(c) any (strict) partial order on a finite set X is contained in a (strict) total order on X.

(a) Choose $x_0 \in X$. If it is not minimal, choose $x_1 < x_0$, and so on. This descending chain has no repetitions, so must terminate in a minimal element.

(b) The unique minimal element of a finite totally ordered set is its least element. Given two finite totally ordered sets of the same size, match their least elements, and proceed by induction.

(c) Let R be a non-strict partial order on X, where X is finite. Suppose that R is not total, so that there exist incomparable elements $a, b \in X$. Let

$$R^+ = R \cup \{(x,y) \in X \times X : (x,a), (b,y) \in R\}.$$

Case analysis shows that R^+ is a partial order containing R. After finitely many steps of this kind, we reach a total order.

1.13 Let R be a reflexive and transitive relation on a set X.

(a) Define a relation S on X by

$$S = \{(x,y) : (x,y), (y,x) \in R\}.$$

Show that S is an equivalence relation on X.

(b) Define a relation T on the set X/S of S-classes in X by

$$T = \{(S(x), S(y)) : (x,y) \in R\}.$$

Show that T is a non-strict order on X/S.

(a) For all $x \in X$, we have $(x,x),(x,x) \in R$, so $(x,x) \in S$; that is, S is reflexive. It is clearly symmetric. Suppose that $(x,y),(y,z) \in S$. Then $(x,y),(y,x),(y,z),(z,y) \in R$. As R is transitive, the first and third pairs show that $(x,z) \in R$, while the fourth and second show that $(z,x) \in R$. Hence $(x,z) \in S$, and S is transitive.

(b) First, observe that if $(S(x),S(y)) \in T$, then $(x',y') \in R$ for all $x' \in S(x)$, $y' \in S(y)$. For se have $(x',x),(x,y),(y,y') \in R$; now apply the transitive law twice. So the definition of T is independent of the choice of equivalence class representatives.

For all $x \in X$, we have $(x,x) \in R$, so $(S(x),S(x)) \in T$. So T is reflexive. A similar argument shows that it is transitive. Now suppose that $(S(x),S(y)),(S(y),S(x)) \in T$. Then $(x,y),(y,x) \in R$, whence $(x,y) \in S$ and $S(x) = S(y)$. So T is antisymmetric.

1.14

(a) Show that the cartesian product of finitely many copies of $\mathbb{N}$ is countable.

(b) Let X be a countable set. Show that the set X^* of all finite sequences of elements of X is countable.

(c) Prove that the set of *algebraic numbers* (those which satisfy some polynomial equation with integer coefficients) is countable. Prove that the set of *transcendental numbers* (those real numbers which are not algebraic) is uncountable.

(a) We show that $\mathbb{N}^n$ is countable by induction on n. The assertion is clearly true for $n = 1$. Suppose that $\mathbb{N}^n$ is countable. Then

$$\mathbb{N}^{n+1} = \mathbb{N}^n \times \mathbb{N}$$

is the cartesian product of two countable sets, so is countable.

(b) The set of n-tuples of elements of X is countable, by part (a). So X^* is the union of countably many countable sets (namely X^n for each natural number n), so is countable.

(c) A polynomial equation of degree n is specified by $n+1$ coefficients. By part (b), the set of equations is countable. But an equation of degree n has at most n real roots. So the set of algebraic numbers is the union of countably many countable sets, hence countable.

If the set of transcendental numbers were countable, then the set of all real numbers would be the union of two countable sets, whence countable, which it is not. So the set of transcendental numbers is uncountable.

This is Cantor's proof of the existence of transcendental numbers: an uncountable set cannot be empty!

1.15

(a) Show that there is a bijection between $\mathbb{R}$ and the open interval $(0,1)$.

(b) Show that there is a bijection between the interval $(0,1)$ and the interior of the unit square.

(c) Deduce that $\mathbb{C}$ has the same cardinality as $\mathbb{R}$.

(a) The function $x \mapsto \arctan(\pi(x-\frac{1}{2}))$ is a bijection.

(b) Represent numbers in $(0,1)$ by their decimal expansions, with the convention that terminating decimals are represented by infinite decimals which are zero (rather than nine) from some point on. Now to the pair (x,y) of real numbers, where $x = 0.x_1x_2\ldots$ and $y = 0.y_1y_2\ldots$, corresponds the real number $z = 0.x_1y_2x_2y_2\ldots$. This function is clearly one-to-one, and its image contains all real numbers except those whose decimal expansion from some point on reads $090909\ldots$. There are only countably many of these (they are all rational). So the function can be adjusted to give a bijection between the interval and the square.

(c) By (a), the set $\mathbb{C}$ has the same cardinality as the interior of the unit square, hence as the open unit interval, hence as $\mathbb{R}$.

1.16 Let $(X,<)$ be a countable totally ordered set. Suppose that

(a) X is *dense*, that is, if $x<y$, then there exists z with $x<z<y$.

(b) X has no least or greatest element.

Prove that X is order-isomorphic to $\mathbb{Q}$.

Enumerate $X = (x_0, x_1, \ldots)$ and $\mathbb{Q} = (q_0, q_1, \ldots)$.

Now define, inductively, a map $f : X \to \mathbb{Q}$ as follows:

(a) $f(x_0) = q_0$.

(b) Suppose that $f(x_0), \ldots, f(x_{n-1})$ have been defined. Then the n points $x_0, \ldots, x_{n-1}$ divide X into $n+1$ intervals (including two semi-infinite intervals); x_n lies in one of these intervals, say (x_i, x_j). Now the corresponding interval $(f(x_i), f(x_j))$ in $\mathbb{Q}$ is non-empty. Choose the rational number q_h with smallest index in this interval, and define $f(x_n) = q_h$.

The function f defined in this way is certainly one-to-one. It is order-preserving: for the construction ensures that the order relation holding between $f(x_n)$ and each $f(x_i)$ for $i<n$ is the same as that between x_n and x_i. The difficult part is to show that f is onto.

Suppose, for a contradiction, that f is not onto. Let q_m be the rational with smallest index which is not in the image of f. Then $q_0, \ldots, q_{m-1}$ are all in the image of f. Choose n such that

$$\{q_0, \ldots, q_{m-1}\} \subseteq \{f(x_0), \ldots, f(x_{n-1})\}.$$

Now q_m lies in one of the $n+1$ intervals into which the line is divided by the n points $f(x_0),\ldots,f(x_{n-1})$; say $q_m \in (f(x_i),f(x_j))$. The corresponding interval (x_i,x_j) in X is non-empty (by the denseness of X). Let x_r be the point of X in this interval with smallest index. When we come to define $f(x_r)$, we find $x_r \in (x_i,x_j)$, so we must choose $f(x_r)$ to be the rational with smallest index in $(f(x_i),f(x_j))$. But this is q_m, since all of $q_0,\ldots,q_{m-1}$ have already been chosen. Thus, $f(x_r) = q_m$, contrary to the assumption that q_m is not in the image of f.

If q_m is in one of the semi-infinite intervals at either end, then the argument is similar, but using the fact that X has no least or greatest element instead of the denseness of X.

Thus f is onto, and so is an order-isomorphism.

1.17 Use the same method to prove that any countable totally ordered set is isomorphic to a subset of $\mathbb{Q}$.

Define the function f as in the preceding question. As before, it is one-to-one and order-preserving. This is all that is required; we don't have to prove that it is onto (and, of course, it need not be).

1.18 For $n > 0$, define a function $f : \mathcal{P}_n(\mathbb{N}) \to \mathbb{N}$ by the rule

$$f(\{x_0,x_1,\ldots,x_{n-1}\}) = \binom{x_0}{1} + \binom{x_1}{2} + \cdots + \binom{x_{n-1}}{n},$$

where $x_0 < x_1 < \cdots < x_{n-1}$. Prove that f is a bijection.

First we make the following observation:

$$f(\{x-n+1,x-n+2,\ldots,x\}) = \binom{x+1}{n} - 1.$$

For this, we use the standard identity

$$\binom{y}{r-1} + \binom{y}{r} = \binom{y+1}{r}$$

for binomial coefficients. Now, if we add one to the left-hand side of the first equation, the first two terms become

$$\binom{x-n+2}{1} + \binom{x-n+2}{2} = \binom{x-n+3}{2};$$

this term then adds to the next term $\binom{x-n+3}{3}$ to give $\binom{x-n+4}{3}$; the process continues like a row of dominoes until we have a single term $\binom{x+1}{n}$.

Now we show that f is a bijection by showing that there is a unique solution $(x_0,\ldots,x_{n-1})$ of the equation $f(\{x_0,\ldots,x_{n-1}\}) = N$ with $x_0 < \ldots < x_{n-1}$. The proof is by induction on n. Suppose that

$$\binom{y}{n} \le N < \binom{y+1}{n}.$$

Then we must choose $x_{n-1} = y$, since if $x_{n-1} = x$ then the maximum possible value of $f(\{x_0, \dots, x_{n-1}\})$ would be $\binom{x+1}{n} - 1$, by our earlier calculation. Then we have to choose $x_0, \dots, x_{n-2}$ so that

$$f(\{x_0, \dots, x_{n-2}\}) = N - \binom{y}{n}.$$

By the inductive hypothesis, there is a unique solution; and this solution satisfies $x_{n-2} < y$, since

$$N - \binom{y}{n} < \binom{y+1}{n} - \binom{y}{n} = \binom{y}{n-1}.$$

The result is proved.

The induction begins since for $n = 1$ the function f is simply given by $f(\{x\}) = x$.

1.19 Prove that the following two statements are equivalent.

(a) The cartesian product of any family of non-empty sets is non-empty.

(b) Let P be a partition of X. Then there is a subset Y of X which contains exactly one element from each member of P.

Assume (a) (the Axiom of Choice), and let P be a partition of X. Let F be the identity function on P. Then $F(p) = p \neq \emptyset$ for all $p \in P$. Let f be a choice function, and $Y = \{f(p) : p \in P\}$. Then, for every $p \in P$, $Y \cap p = \{f(p)\}$.

Conversely, assume (b), and let F be any function on X such that $F(X) \neq \emptyset$ for all $x \in X$. Let $Z = \{(x,y) : y \in F(x), x \in X\}$. Now $P = \{\{x\} \times F(x) : x \in X\}$ is a partition of Z. Choose a set Y meeting every set of this partition in just one point. Thus, for each $x \in X$, there is a unique $y \in F(x)$ such that $(x,y) \in Y$. Now Y is itself a choice function for F.

1.20 Use the Axiom of Choice to show that, if there is a surjection from Y to X, then there is an injection from X to Y.

Let g be a surjection from Y to X. Let F be the function from X to $\mathcal{P}Y$ given by

$$F(x) = \{y \in Y : g(y) = x\}.$$

By assumption, $F(x) \neq \emptyset$ for all $x \in X$. Let f be a choice function for F. Then $f(x) \in F(x) \subseteq Y$ for all $x \in X$, that is, f is a function from X to Y. Now clearly if $x_1 \neq x_2$, then $F(x_1)$ and $F(x_2)$ are disjoint, so $f(x_1) \neq f(x_2)$; so f is an injection.