Discrete Mathematics

Supervision 8 – Solutions with Commentary

Marcelo Fiore   Ohad Kammar   Dima Szamozvancev

11. On surjections and injections

11.1. Basic exercises

1. Give two examples of functions that are surjective, and two examples of functions that are not.

   Surjective. Absolute value function to the naturals $|-| \colon \mathbb{Z} \twoheadrightarrow \mathbb{N}$; natural log function $\ln \colon \mathbb{R}_0^+ \twoheadrightarrow \mathbb{R}$; first projection function from the Cartesian product of two (nonempty) sets $\pi_1 \colon A \times B \twoheadrightarrow A$.

   Not surjective. Integer squaring function on the naturals: $(-)^2 \colon \mathbb{N} \to \mathbb{N}$ (only returns perfect squares); constant function $c_b \colon A \to B$ with value $b \in B$ (always outputs b if B is not the singleton set); successor function $(-) + 1 \colon \mathbb{N} \to \mathbb{N}$ (0 is not the successor of any number).

2. Give two examples of functions that are injective, and two examples of functions that are not.

   Injective. The inclusion/injection function $\iota \colon S \rightarrowtail A$ for any subset S of A; exponential function $x \mapsto e^x \colon \mathbb{R} \rightarrowtail \mathbb{R}$; perfect hash function.

   Not injective. Integer squaring function: $(-)^2 \colon \mathbb{Z} \to \mathbb{Z}$ (since $x^2 = (-x)^2$); quotient function $q(a) = [a]_E \colon A \to A/E$ for an equivalence relation E (related elements map to the same equivalence class); $\sin(x) \colon [0, 2\pi] \to [-1, 1]$ since $\sin(0) = \sin(2\pi) = 0$.

11.2. Core exercises

1. Explain and justify the phrase *injections can be undone*.

   Every injection (from a non-empty domain) has a retraction which "undoes" its effect. If $i \colon A \rightarrowtail B$ is an injection, every b in B is mapped to by at most $a \in A$; thus, a retraction can be defined as

   $$r(b) = \begin{cases} a & \text{if } \exists a \in A.\ i(a) = b \\ a_0 & \text{otherwise} \end{cases}$$

   where a_0 is any element of A. This is total, since every b is either mapped to the source a for which $i(a) = b$, or to the fixed a_0. It is also functional, since there may only be at most one a for which $i(a) = b$. By construction, $r \circ i = \mathrm{id}_A$, so the two form a section-retraction pair.

   The implication holds in the other direction as well: every section $s \colon A \to B$ (with a retraction $r \colon B \to A$) is an injection. To see this, consider $a, a' \in A$ and assume $s(a) = s(a')$. But since $r \circ s = \mathrm{id}_A$, we have that $r(s(a)) = r(s(a'))$ implies $a = a'$, so s must be an injection.

2. Show that $f \colon A \to B$ is a surjection if and only if for all sets C and functions $g, h \colon B \to C$, $g \circ f = h \circ f$ implies $g = h$.

($\Rightarrow$) Let $f : A \twoheadrightarrow B$ be a surjection: for all $b \in B$ there exists an $a \in A$ such that $f(a) = b$. Furthermore, let $g, h \colon B \to C$ be functions and assume ① $g \circ f = h \circ f$. We need to show that $g = h$, that is, for all $b \in B$, $g(b) = h(b)$. But by assumption any $b \in B$ is equal to $f(a)$ for some $a \in A$, so the condition is equivalent to $g(f(a)) = h(f(a))$, which is just ①.

($\Leftarrow$) We show the contrapositive: if $f : A \twoheadrightarrow B$ is not surjective, then there exists a C and functions $g, h \colon B \to C$ such that $g \circ f = h \circ f$ but $g \neq h$. If f is not surjective, there exists a $b_0 \in B$ such that for all $a \in A$, $f(a) \neq b_0$. We can therefore choose two functions g and h such that they match on the range of f, but differ on b_0. For example, take C to be B with a new distinguished element $\star$ added: $C = B \cup \{ \star \}$. Let $g : B \to B \cup \{ \star \}$ be the inclusion $g(b) = b$, and let $h(b_0) = \star$ and $h(b) = b$ for all $b \neq b_0$. Then, $g \circ f = h \circ f$, since g and h defined to be equal for all elements in the range of f, but they differ on the element b_0 not "covered" by f, hence $g \neq h$.

♪ The ($\Leftarrow$) direction can be presented as a non-contrapositive argument as well. Let $f : A \twoheadrightarrow B$ be a function and assume for all $g, h \colon B \to C$, if $g \circ f = h \circ f$ then $g = h$. We need to show that for all $b \in B$ there exists an $a \in A$ such that $f(a) = b$. Choose $C = [2] = \{ 0, 1 \}$ and define $g = \chi_B$ and $h = \chi_{\overrightarrow{f}(A)}$, where $\overrightarrow{f}(A) \subseteq B$ is also called the range of f, i.e. the set $\{ f(a) \in B \mid a \in A \}$. That is, $g(b) = 1$ for all b, and $h(b) = 1$ for all b in the range of f, and 0 otherwise. Now, for all $a \in A$, $g(f(a)) = h(f(a))$, but by assumption this implies that $g = h$. This is only possible if the range of f is B itself, i.e. f is surjective.

What would be an analogous condition for injections?

Injectivity is equivalent to left-cancellability: $f : B \to C$ is an injection iff for all sets A and functions $g, h \colon A \to B$, if $f \circ g = f \circ h$ then $g = h$.

($\Rightarrow$) Assume $f : B \rightarrowtail C$ is an injection, and suppose that $f \circ g = f \circ h$ for some $g, h \colon A \to B$. We need to show that for all $a \in A$, $g(a) = h(a)$. Injectivity means that for all $b_1, b_2 \in B$, if $f(b_1) = f(b_2)$ then $b_1 = b_2$. Instantiating this for $g(a), h(a) \in B$, and using the assumption $f \circ g = f \circ h$, we deduce that $f(g(a)) = f(h(a))$ implies $g(a) = h(a)$. Since $a \in A$ was arbitrary, we have that $g = h$.

($\Leftarrow$) Assume that for all A and $g, h \colon A \to B$, ① $f \circ g = f \circ h$ implies $g = h$. We need to show that for all $b_1, b_2 \in B$, if $f(b_1) = f(b_2)$ then $b_1 = b_2$. Take $b_1, b_2 \in B$ and assume that ② $f(b_1) = f(b_2)$; for $A = \{ () \}$ the singleton set, define $g, h \colon A \to B$ as $g() = b_1$, and $h() = b_2$. Then, by ②, $f(g()) = f(b_1) = f(b_2) = f(h())$, but then by ① $g = h$ so $b_1 = b_2$.

3. Use the above sufficient condition to show that the identity function is a surjection, and the composition of surjections is a surjection.

Identity. We show that $\mathrm{id}_A : A \to A$ is a surjection. Let X be a set and $g, h : A \to X$ be two functions, and assume $g \circ \mathrm{id}_A = h \circ \mathrm{id}_A$. Since the identity is the unit of composition, we get $g = h$ immediately, so id_A is a surjection by §11.2.2.

Composition. Let $f : A \twoheadrightarrow B$ and $g : B \twoheadrightarrow C$ be surjections. We show that $g \circ f : A \to C$ is a surjection via §11.2.2. Let $h, i : C \to X$ be two functions and assume $h \circ (g \circ f) = i \circ (g \circ f)$; we need to prove that $h = i$. Composition is associative, so $(h \circ g) \circ f = (i \circ g) \circ f$ – and f is a surjection, so we have that $h \circ g = i \circ g : B \to X$. Similarly, g is surjective, so $h = i$.

♪ As is hopefully apparent now, exercises of the form *[object] is [defined set-theoretic concept] iff [it satisfies set-theoretic property]* are powerful and reusable proof principles:

X is the union of A and B	iff $A \subseteq X$, $B \subseteq X$, and for all Y, $(A \subseteq Y \land B \subseteq Y) \Rightarrow X \subseteq Y$
R is an equivalence relation	iff $\mathrm{id}_A \subseteq R$, $R = R^{\mathrm{op}}$ and $R \circ R \subseteq R$
R is a partial function	iff $R \circ R^{\mathrm{op}} \subseteq \mathrm{id}_B$
f is surjective	iff for all $g, h : B \to X$, $g \circ f = h \circ f \Rightarrow g = h$

Thus, to prove that *[object]* is *[defined set-theoretic concept]*, instead of expanding the set-theoretic definition (which is usually given in terms of individual elements), we can reason via the higher-level set-theoretic properties which may result in more abstract and elegant proofs with preorder or equational reasoning. While perhaps not immediately as intuitive and not necessarily shorter, once you get used to the approach, you will be able to recognise and appreciate opportunities for reasoning via sufficient/universal properties rather than reaching for "let $x \in A$ and show $x \in B$" right away and making your proof low-level and often harder to follow. That said, for explicitly defined sets and functions, proving surjectivity (for example) from first principles may be more direct – see some exercises in the next section.

12. On images

12.1. Basic exercises

1. Let $R_2 = \{ (m, n) \mid m = n^2 \} : \mathbb{N} \nrightarrow \mathbb{Z}$ be the integer square-root relation. What is the direct image of $\mathbb{N}$ under R_2? And what is the inverse image of $\mathbb{N}$?

   By the definition of the direct and inverse relational images, we have:

   $$\overrightarrow{R_2}(\mathbb{N}) = \mathbb{Z} \qquad \overleftarrow{R_2}(\mathbb{N}) = \{ 0 \} \cup \{ n \in \mathbb{N} \mid n \text{ is not square} \}$$

   ♪ This may well be called a "trick question", since the answer could hardly be more counterintuitive – then again, it follows directly from the definition of inverse relation images so there is not much to argue about! R_2 relates every integer (on the right) with its square (on the left), a natural number: $R_2 = \{ (0, 0), (1, -1), (1, 1), (4, 2), (4, -2), (9, -3), (9, 3), \dots \}$. The direct image of the natural numbers is therefore $\mathbb{Z}$ itself, since the square of every integer is in $\mathbb{N}$. It may seem intuitively obvious that the inverse image of $\mathbb{N} \subseteq \mathbb{Z}$ under the square root relation would be the set of square numbers, but this is distinctly *not* the case. Recall the definition of inverse relational images:

   $$\overleftarrow{R}(Y \subseteq B) \triangleq \{ a \in A \mid \forall b \in B. \, a \, R \, b \Rightarrow b \in Y \}$$

For R_2, and $Y = \mathbb{N} \subseteq \mathbb{Z}$, this becomes:

$$\overleftarrow{R_2}(\mathbb{N}) \triangleq \{\, m \in \mathbb{N} \mid \forall n \in \mathbb{Z}.\ m = n^2 \Rightarrow n \in \mathbb{N} \,\}$$

In other words, if there are any integers that square to an element of $\overleftarrow{R_2}(\mathbb{N})$, they all have to be natural numbers. 0 is certainly in the inverse image, since the only number that squares to 0 is 0 itself, and it is in $\mathbb{N}$. The problems start with nonzero square numbers like $1, 4, 9$, etc.: there are exactly two integers that square to the same perfect square number, namely the square root, and the negative of the square root. Only one of these is a natural number, the other violates $m = n^2 \Rightarrow n \in \mathbb{N}$ and therefore cannot be an element of the inverse image. Thus, the inverse image of natural numbers under the square-root relation contains no square numbers other than 0. Even worse is that every natural number which is *not* a perfect square (and therefore isn't related to any integers) vacuously satisfies the condition: for any $n \in \mathbb{Z}$, $2 \neq n^2$ so the hypothesis is never satisfied and the implication holds! As a result, the inverse image contains all the non-square natural numbers and 0.

You may rightly ask: why do we define inverse images in such a way? The answer is simply that this is the most natural way to define it as a dual of the direct image $\overrightarrow{R}(X) \triangleq \{\, b \in B \mid \exists x \in X.\ x\,R\,b \,\}$. Indeed, if we slightly rephrase the condition $\exists x \in X.\ x\,R\,b$ to separate existence and membership of X, and compare it to the inverse image definition, we get:

$$\overrightarrow{R}(X) \triangleq \{\, b \in B \mid \exists x \in A.\ x\,R\,b\ \wedge\ x \in X \,\}$$
$$\overleftarrow{R}(Y) \triangleq \{\, a \in A \mid \forall y \in B.\ a\,R\,y \Rightarrow y \in Y \,\}$$

As is often the case with mathematics, symmetry and simplicity takes precedence over intuition, and trying to define the inverse image to yield the "expected" results would needlessly complicate the definition. In fact, what we intuitively expect the inverse image of $\mathbb{N}$ under R_2 to be (the set of perfect squares) is nothing more than the direct image of $\mathbb{N}$ under the opposite relation R_2^{op}.

2. For a relation $R \colon A \nrightarrow B$, show that:

a) $\overrightarrow{R}(X) = \bigcup_{x \in X} \overrightarrow{R}(\{\, x \,\})$ for all $X \subseteq A$

Let X be a subset of A. We calculate as follows:

$$\begin{aligned}
\overrightarrow{R}(X) &= \{\, b \in B \mid \exists x \in X.\ x\,R\,b \,\} \\
&= \{\, b \in B \mid \exists x \in X.\ \exists y' \in \{\, x \,\}.\ y'\,R\,b \,\} \\
&= \{\, b \in B \mid \exists x \in X.\ b \in \overrightarrow{R}(\{\, x \,\}) \,\} \\
&= \bigcup_{x \in X} \overrightarrow{R}(\{\, x \,\})
\end{aligned}$$

b) $\overleftarrow{R}(Y) = \{\, a \in A \mid \overrightarrow{R}(\{\, a \,\}) \subseteq Y \,\}$ for all $Y \subseteq B$.

Let Y be a subset of B. We calculate as follows:

$$\overleftarrow{R}(Y) = \{\, a \in A \mid \forall y \in B.\ a\,R\,y \Rightarrow y \in Y \,\}$$
$$= \{\, a \in A \mid \forall y \in B.\ \big(\exists a' \in \{a\}.\ a'\,R\,y\big) \Rightarrow y \in Y \,\}$$
$$= \{\, a \in A \mid \forall y \in B.\ y \in \overrightarrow{R}(\{a\}) \Rightarrow y \in Y \,\}$$
$$= \{\, a \in A \mid \overrightarrow{R}(\{a\}) \subseteq Y \,\}$$

♪ This equivalent characterisations of inverse images highlights the requirement that *every* $y \in B$ related to an $a \in A$ has to be in Y, not just at least one.

12.2. Core exercises

1. For $X \subseteq A$, prove that the direct image $\overrightarrow{f}(X) \subseteq B$ under an injective function $f : A \rightarrowtail B$ is in bijection with X; that is, $X \cong \overrightarrow{f}(X)$.

Let $f : A \to B$ be an injective function and let X be a subset of A. We show that the direct image of X under f is isomorphic to X by constructing a bijection $h : X \xrightarrow{\cong} \overrightarrow{f}(X)$. Define h as

$$h(x \in X) = f(x) \in \overrightarrow{f}(X)$$

By construction, h is a function from X to $\overrightarrow{f}(X)$ because every output of f for an input in X ends up in the direct image. We show that h is surjective and injective. Take any element $y \in \overrightarrow{f}(X)$; by definition, there must exist an element $x \in X$ such that $f(x) = h(x) = y$, which is the condition for surjectivity of h. Now, take $x_1, x_2 \in X$ and assume that $h(x_1) = h(x_2)$. Then, $f(x_1) = f(x_2)$, but f is injective, so $x_1 = x_2$ – proving that h is injective too. As a direct corollary, the range of an injection is isomorphic to the domain: $\overrightarrow{f}(A) \cong A$.

♪ This is a situation where proving injectivity and surjectivity is more convenient than trying to precisely formulate an inverse function that maps $y \in \overrightarrow{f}(X)$ to "the element in X that got uniquely mapped to y" and using this to calculate the inverse laws.

2. Prove that for a surjective function $f : A \twoheadrightarrow B$, the direct image function $\overrightarrow{f} : \mathcal{P}(A) \to \mathcal{P}(B)$ is surjective.

Assume $f : A \twoheadrightarrow B$ is a surjection: for all $b \in B$, there exists an $a \in A$ such that $f(a) = b$. We need to prove that for any element $Y \in \mathcal{P}(B)$ there exists an $X \in \mathcal{P}(B)$ such that $\overrightarrow{f}(X) = Y$. Thus, take a subset $Y \subseteq A$, and let the corresponding subset of A be the inverse image $\overleftarrow{f}(Y) \subseteq A$. We now need to show that $\overrightarrow{f}\big(\overleftarrow{f}(Y)\big) = Y$, for which we calculate:

$$\overrightarrow{f}\big(\overleftarrow{f}(Y)\big) = \{\, b \in B \mid \exists a \in \overleftarrow{f}(Y).\ f(a) = b \,\}$$
$$= \{\, b \in B \mid \exists a \in A.\ f(a) \in Y \ \wedge\ f(a) = b \,\}$$
$$= \{\, b \in Y \mid \exists a \in A.\ f(a) = b \,\}$$

but the last set is precisely Y since f is surjective and therefore the comprehension condition holds for all $b \in Y$. As a direct corollary, the range of a surjection is equal to the

codomain: $\vec{f}(A) = B$. A bijection is both an injection and a surjection, so $A \cong \vec{f}(A) = B$.

3. Show that any function $f : A \to B$ can be decomposed into an injection and a surjection: that is, there exists a set X, a surjection $s : A \twoheadrightarrow X$ and an injection $i : X \rightarrowtail B$ such that $f = i \circ s$.

> Let $f : A \to B$ be a function, not necessarily a surjection or injection. Take X to be the range of f, that is, the direct image of its domain: $X = \vec{f}(A) \subseteq B$. Then, by definition, every element $b \in \vec{f}(A)$ has an associated element $a \in A$ such that $f(a) = b$, so f with its codomain restricted to its range is a surjection – hence, $s(a) = f(a) : A \twoheadrightarrow \vec{f}(A)$. The range of f is a subset of the codomain, so we have the canonical inclusion $i(b) = b : \vec{f}(A) \rightarrowtail B$ which is an injection. For all $a \in A$, $i(s(a)) = i(f(a)) = f(a)$, so the composite $i \circ s$ is indeed equal to f, as required.
>
> ♪ When $A = B$ (and $f : A \to A$ is an endofunction), the construction of course still works. In fact, it gives one half of the idempotent-splitting example §9.2.1, in which an idempotent endofunction $e : A \to A$ is split through its range $\{ e(b) \mid b \in B \} = \vec{e}(B)$ into functions r and s as $s \circ r = e$ which, thanks to the idempotence condition, form a section-retraction pair: $r \circ s = \mathrm{id}_B$. Sections are always injections, and the constructed retraction is a surjection, matching the result shown in this exercise.

4. For a relation $R : A \nrightarrow B$, prove that

   a) $\vec{R}\left(\bigcup \mathcal{F}\right) = \bigcup \left\{ \vec{R}(X) \mid X \in \mathcal{F} \right\}$ for all $\mathcal{F} \subseteq \mathcal{P}(A)$

   > Let $\mathcal{F} \subseteq \mathcal{P}(A)$ be a family of subsets. We have the following calculation:

   $$
   \begin{aligned}
   b \in \vec{R}\left(\bigcup \mathcal{F}\right) &\iff \exists a \in \bigcup \mathcal{F}. \, a R b \\
   &\iff \exists X \in \mathcal{F}. \, \exists a \in X. \, a R b \\
   &\iff \exists X \in \mathcal{F}. \, b \in \vec{R}(X) \\
   &\iff \exists Y \in \left\{ \vec{R}(X) \mid X \in \mathcal{F} \right\}. \, b \in Y \\
   &\iff b \in \bigcup \left\{ \vec{R}(X) \mid X \in \mathcal{F} \right\}
   \end{aligned}
   $$

   b) $\overleftarrow{R}\left(\bigcap \mathcal{G}\right) = \bigcap \left\{ \overleftarrow{R}(Y) \mid Y \in \mathcal{G} \right\}$ for all $\mathcal{G} \subseteq \mathcal{P}(B)$

   > Let $\mathcal{F} \subseteq \mathcal{P}(A)$ be a family of subsets. We have the following calculation:

   $$
   \begin{aligned}
   a \in \overleftarrow{R}\left(\bigcap \mathcal{G}\right) &\iff \forall b \in B. \, a R b \Rightarrow a \in \bigcap \mathcal{G} \\
   &\iff \forall b \in B. \, a R b \Rightarrow \forall Y \in \mathcal{G}. \, a \in Y \\
   &\iff \forall Y \in \mathcal{G}. \, \forall b \in B. \, a R b \Rightarrow a \in Y \\
   &\iff \forall Y \in \mathcal{G}. \, a \in \overleftarrow{R}(Y) \\
   &\iff \forall X \in \left\{ \overleftarrow{R}(Y) \mid Y \in \mathcal{G} \right\}. \, a \in X \\
   &\iff a \in \bigcap \left\{ \overleftarrow{R}(Y) \mid Y \in \mathcal{G} \right\}
   \end{aligned}
   $$

5. Show that, by the inverse image, every map $A \to B$ induces a *Boolean algebra map* $\mathcal{P}(B) \to \mathcal{P}(A)$. That is, for every function $f : A \to B$, its inverse image preserves set operations:

- $\overleftarrow{f}(\emptyset) = \emptyset$

$$a \in \overleftarrow{f}(\emptyset) \iff f(a) \in \emptyset \iff \text{false} \iff a \in \emptyset$$

- $\overleftarrow{f}(B) = A$

$$a \in \overleftarrow{f}(B) \iff f\,a \in B \iff \text{true} \iff a \in A$$

- $\overleftarrow{f}(X \cup Y) = \overleftarrow{f}(X) \cup \overleftarrow{f}(Y)$

$$a \in \overleftarrow{f}(X \cup Y) \iff f(a) \in (X \cup Y) \iff f(a) \in X \lor f(a) \in Y$$
$$\iff a \in \overleftarrow{f}(X) \lor a \in \overleftarrow{f}(X) \iff a \in \overleftarrow{f}(X) \cup \overleftarrow{f}(Y)$$

- $\overleftarrow{f}(X \cap Y) = \overleftarrow{f}(X) \cap \overleftarrow{f}(Y)$

$$a \in \overleftarrow{f}(X \cap Y) \iff f(a) \in (X \cap Y) \iff f(a) \in X \land f(a) \in Y$$
$$\iff a \in \overleftarrow{f}(X) \land a \in \overleftarrow{f}(X) \iff a \in \overleftarrow{f}(X) \cap \overleftarrow{f}(Y)$$

- $\overleftarrow{f}(X^c) = \left(\overleftarrow{f}(X)\right)^c$

$$a \in \overleftarrow{f}(X^c) \iff f(a) \in X^c \iff \neg(f(a) \in X) \iff \neg(a \in \overleftarrow{f}(X)) \iff a \in \left(\overleftarrow{f}(X)\right)^c$$

13. On countability

13.1. Basic exercises

1. Prove that every finite set is countable.

   If the set is empty, it is countable by definition. Otherwise, if A is finite, it has at most $\#A = n > 0$ elements. Thus, an enumeration $\mathbb{N} \twoheadrightarrow A$ can be constructed by mapping the first n natural numbers to distinct elements of A (e.g. by putting them in some order and assigning $k : [0..n-1]$ to the kth element), and the rest of the naturals to a single element $a_0 \in A$. The mapping is surjective by construction (the kth element of A is listed at k) so it is an enumeration.

2. Demonstrate that $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ are countable sets.

   $\mathbb{N}$ is enumerated by the identity function, which is in particular a surjection.

   $\mathbb{Z}$ is enumerated by alternating between positive and negative numbers: $0, 1, -1, 2, -2, \ldots$. Explicitly, $e : \mathbb{N} \twoheadrightarrow \mathbb{Z}$ is the enumeration

   $$e(n) \triangleq \begin{cases} \frac{n+1}{2} & \text{if } n \text{ is odd} \\ -\frac{n}{2} & \text{if } n \text{ is even} \end{cases}$$

   $\mathbb{Q}$ is enumerable using the traversal of the coordinate plane demonstrated on .

13.2. Core exercises

1. Let A be an infinite subset of $\mathbb{N}$. Show that $A \cong \mathbb{N}$. *Hint*: Adapt the argument shown in the proof of Proposition 144, showing that the map $\mathbb{N} \to A$ is both injective and surjective.

> Let A be an infinite subset of $\mathbb{N}$. We construct a bijection $\mathbb{N} \xrightarrow{\cong} A$ to show that they are isomorphic. To this end, define the function $\mu \colon \mathbb{N} \to A$ as follows:
>
> $$\mu(0) \triangleq \min(A) \qquad \mu(n+1) \triangleq \min\{\, k \in A \mid \mu(n) < k \,\} = \min(A \setminus \{\, \mu(k) \mid k \leq n \,\})$$
>
> We will denote the set $A \setminus \{\, \mu(k) \mid k \leq n \,\}$ as A_n, so that $\mu(n+1) = \min(A_n)$.
>
> To show that μ is an injection, we need to prove that if $\mu(m) = \mu(n)$ then $m = n$. We equivalently prove the contrapositive: if $m \neq n$, then $\mu(m) \neq \mu(n)$. Without loss of generality, assume that $m < n$; then, $\mu(m) \in \{\, \mu(k) \mid k \leq n-1 \,\}$, so $\mu(m) \notin \{\, \mu(k) \mid k \leq n-1 \,\}^{\mathsf{c}} = A \setminus \{\, \mu(k) \mid k \leq n-1 \,\}$. On the other hand, $\mu(n)$ is an element of the latter set (its minimum), which means that $\mu(m)$ cannot equal $\mu(n)$.
>
> To show that μ is a surjection, we let a be an arbitrary element of A and show that there is an $i \in \mathbb{N}$ such that $\mu(i) = a$. Consider the set $\{\, k \in \mathbb{N} \mid \mu(k) < a \,\}$ of numbers which get mapped to an element below a in A, and let N be the size of this set (which, by the Pigeonhole Principle, must be at most a). Now, A_N is the subset of A obtained by removing its N least elements, and by construction, its least element is a. But if $a = \min(A_N)$, then it is equal to $\mu(N+1)$ by the definition of μ, so we indeed have the natural number $i = N+1$ such that $\mu(i) = a$.

2. For an infinite set A, prove that the following are equivalent:

   a) There is a bijection $\mathbb{N} \xrightarrow{\cong} A$.
   b) There is a surjection $\mathbb{N} \twoheadrightarrow A$.
   c) There is an injection $A \rightarrowtail \mathbb{N}$.

   > **(a) $\Rightarrow$ (b), (c)** Every bijection $f \colon \mathbb{N} \xrightarrow{\cong} A$ has an inverse $f^{-1} \colon A \xrightarrow{\cong} \mathbb{N}$ which is itself a bijection. Every bijection is a surjection (giving $\mathbb{N} \twoheadrightarrow A$ from f) and an injection (giving $A \rightarrowtail \mathbb{N}$ from f^{-1}).
   >
   > **(b) $\Rightarrow$ (c)** Let $s \colon \mathbb{N} \twoheadrightarrow A$ be a surjection. We need to construct an injection $i \colon A \rightarrowtail \mathbb{N}$, assigning a unique numeric code to every element of A. As s is a surjection, the inverse images of the singleton subsets of A (also called the set of *fibres* of elements of A) are all non-empty, and as they are subsets of the natural numbers, they must have a minimal element. Thus, define $i \colon A \to \mathbb{N}$ as a function that maps an $x \in A$ to the smallest natural number that maps to x:
   >
   > $$i(x) = \min\big(\overleftarrow{s}(\{x\})\big) = \min\{\, n \in \mathbb{N} \mid s(n) = x \,\}$$
   >
   > This encodes an element $x \in A$ by the position of its first occurrence in the enumeration given by s. We can see that i is injective as s acts as its retraction: for any $x \in S$, $s(i(x)) = s(n)$ where n is the smallest natural number such that $s(n) = x$ so clearly

$s(i(x)) = s(n) = x$ and $s \circ i = \text{id}_A$, as required.

(c) $\Rightarrow$ **(a)** Let $i : A \rightarrowtail \mathbb{N}$ be an injection. We need to construct a bijection $A \xrightarrow{\cong} \mathbb{N}$, or equivalently, show that A and $\mathbb{N}$ are isomorphic. By §12.2.1, the direct image of the domain under the injection i (i.e. the range of i) is isomorphic to the domain: $\overrightarrow{i}(A) \cong A$. By assumption, A is infinite, so $\overrightarrow{i}(A) \subseteq \mathbb{N}$ is infinite as well. But then it is an infinite subset of the natural numbers, and by §13.2.1, it is isomorphic to $\mathbb{N}$. Hence we have the chain $A \cong \overrightarrow{i}(A) \cong \mathbb{N}$, establishing the bijection $\mathbb{N} \xrightarrow{\cong} A$.

♪ If you look at other resources on countability, you will face several competing, but equivalent (but sometimes not *quite* equivalent) definitions which make translating between various statements and proofs a bit of a chore – especially since the same terms are used by different authors for different purposes. This course uses *enumerable* for sets which have a surjection $\mathbb{N} \twoheadrightarrow A$, and *countable* for sets which are enumerable or empty (since one can't have a function into the empty set so this needs to be handled as a special case). Other literature (e.g. Wikipedia) calls sets which are isomorphic to $\mathbb{N}$ *countably infinite*, and sets which are either finite or countably infinite are called *countable*. The countability condition can be equivalently stated as the set being isomorphic to some subset of the natural numbers, i.e. coming with an injection $A \rightarrowtail \mathbb{N}$. Yet other terms used for the above notions are *at most countable, enumerable, denumerable, equinumerous, listable*, etc.

As this exercise shows, the bijection/surjection/injection notions are equivalent when the set A is infinite, and appropriate connections can be made when the sets are empty or finite. This gives us two equivalent ways of showing that a set is enumerable: either by constructing an enumeration $\mathbb{N} \twoheadrightarrow A$, or by defining an encoding function $A \rightarrowtail \mathbb{N}$ that maps every element of A to a unique natural number. This is related to a concept called *Gödel encoding* which will be covered in more detail in the IB Computation Theory course.

3. Prove that:

   a) Every subset of a countable set is countable.

   > Assume $S \subseteq A$ for some sets A. If A is finite, so is S and it is countable. If A is infinite and S is finite, S is countable. If S is also infinite, we can show that it is enumerable by providing an injection $S \rightarrowtail \mathbb{N}$. But by assumption we have an injection $f : A \rightarrowtail \mathbb{N}$ and subsets come with a canonical injective inclusion function $\iota : S \rightarrowtail A$, so the composite $S \rightarrowtail A \rightarrowtail \mathbb{N}$ is an injection.

   b) The product and disjoint union of countable sets is countable.

   > Assume A and B are countable sets.

   > If either is empty, the Cartesian product will be empty too and therefore countable. In the general case, assume they are enumerable and there are injections $f : A \rightarrowtail \mathbb{N}$ and $g : B \rightarrowtail \mathbb{N}$ that uniquely encode the elements of the sets. We define a function

$p : A \times B \to \mathbb{N}$ as follows:

$$p(a, b) = 2^{f(a)} \cdot 3^{g(a)}$$

By the Fundamental Theorem of Arithmetic, the mapping is injective: the output of this mapping will have a unique prime decomposition, and the number of 2 and 3 factors will give the unique code of elements of A and B, respectively. By §13.2.2, the injection p will imply that $A \times B$ is enumerable.

If both sets are empty, their disjoint union will be empty and therefore countable. If either is empty, the disjoint union will be isomorphic to the other set, which is countable by assumption. In the general case, assume A and B are enumerable and come with injections $f : A \rightarrowtail \mathbb{N}$ and $g : B \rightarrowtail \mathbb{N}$. We define the function $u : A \uplus B \to \mathbb{N}$ as follows:

$$u(0, a) = 2^{f(a)} \qquad u(1, b) = 3^{g(b)}$$

Again, by the Fundamental Theorem of Arithmetic, the prime decomposition of the output of u will uniquely determine the output of $f(a)$ or $g(b)$ which in turn uniquely determine the input a and b by assumption. By §13.2.2, the injection u will imply that $A \uplus B$ is enumerable.

🎵 Constructing unique encodings using products of primes is a useful alternative to the visually descriptive "diagonal traversal" enumeration which is often quite difficult to define explicitly. The specific choice of encoding can of course vary (e.g. we could have encoded disjoint unions via even and odd numbers) but there is no reason to look for the most "efficient" solution since all we care about is whether the enumeration/encoding is possible or not.

4. For a set A, prove that there is no injection $\mathcal{P}(A) \rightarrowtail A$.

We suppose there is an injection $f : \mathcal{P}(A) \rightarrowtail A$ and derive a contradiction. By §11.2.1, the injection f has a retraction $r : A \to \mathcal{P}(A)$ which must be a surjection since it undoes the application of f on any element of A. But then r would be a surjection from a set to its powerset, which is impossible due to Cantor's Theorem.

13.3. Optional advanced exercise

1. Prove that if A and B are countable sets then so are A^*, $\mathcal{P}_{\text{fin}}(A)$ and $\text{PFun}_{\text{fin}}(A, B)$.

All the results follow from Proposition 154: an enumerable indexed disjoint union of enumerable sets is enumerable. An enumeration-style proof is presented in the notes, but an encoding-style argument is straightforward too: we can encode elements of $\biguplus_{i \in I} A_i$ as

$$d(i \in I, a \in A_i) = p(c(i), c_i(a)) = 2^{c(i)} \cdot 3^{c_i(a)}$$

where $c : I \rightarrowtail \mathbb{N}$ is the encoding of the index set, and $c_i : A_i \rightarrowtail \mathbb{N}$ is an encoding for every element of the indexed family.

$A^* = \biguplus_{n \in \mathbb{N}} A^n$ is the set of finite sequences on A. If A is empty, the only finite sequence with elements from A is the empty sequence, so $\{\,()\,\}$ is finite and countable. In the general case, we know that $\mathbb{N}$ is enumerable, and A^n is the iterated binary Cartesian product of enumerable sets and hence is an enumerable set for every $n \in \mathbb{N}$ (quick inline induction proof: $A^0 = \emptyset$ is countable; $A^{k+1} = A^k \times A$ is the Cartesian product of countable A^k by IH, and countable A by assumption). By the above proposition, the $\mathbb{N}$-indexed disjoint union of countable sets is countable.

$\mathcal{P}_{\text{fin}}(A) = \{\, S \subseteq A \mid S \text{ is finite} \,\}$ is the set of finite subsets of A. If A is empty, $\mathcal{P}_{\text{fin}}(A) = \{\emptyset\}$ which is finite so countable. Otherwise, the set A has an encoding $c : A \rightarrowtail \mathbb{N}$ which imposes an ordering on the elements on A based on the ordering of their code: for $a, b \in A$, $a \sqsubseteq b$ if $c(a) \leq c(b)$. This ordering restricts to every subset of A, so in particular, finite subsets of A can be mapped to finite sequences of elements of A according to the ordering $\sqsubseteq$. Then, the set of finite subsets of A is isomorphic to the set of finite sequences on A, which is countable for a countable A.

$\text{PFun}_{\text{fin}}(A, B) = \biguplus_{S \in \mathcal{P}_{\text{fin}}(A)} S \Rightarrow B$ is the set of partial functions with a finite domain of definition from A to B. If A or B are empty, the totally undefined function is the only element of the set so it is countable. Otherwise, the disjoint union is indexed by $\mathcal{P}_{\text{fin}}(A)$ which is enumerable by the result above. The function space $S \Rightarrow B$ has a finite domain S, so a single function $f : S \to B$ can be captured as a finite sequence of elements of B as $(f(s_1), f(s_2), f(s_3), \ldots, f(s_n))$ where $n = \#S$ and s_i is the ith element of S in some ordering (which is always possible to define for a finite S). Thus, $S \Rightarrow B \cong B^{\#S}$ which is countable for any countable B. By Proposition 154, the set $\text{PFun}_{\text{fin}}(A, B)$ is a countable indexed disjoint union of countable sets and is therefore itself countable.