RECURSION AND INDUCTION ON POSETS

In this section, we will see that we can generalise the recursion and induction theorems to posets. The notation $m \prec b$ is an abbreviation for $m \preccurlyeq b$ and $m \neq b$, which we call the *strict order* of $\preccurlyeq$.

In a poset A, unlike in a chain, if we have a minimum, it is not always unique. We denote by $M(A)$ the set of minima of the poset:

Definition 1.43 (Minimal set)**.** Let $(A, \preccurlyeq)$ be a poset. The *minimal set* of a subset $B \subseteq A$, denoted $M_{\preccurlyeq}(B)$ or just $M(B)$, is the set defined by

$$M(B) := \{m \in B : \text{there is no } b \in B \text{ such that } b \prec m\}.$$

If A is a chain, then $M(A) = \{\min A\}$ (if $\min A$ exists). Indeed, just as with chains, the set $M(B)$ can sometimes be empty, e.g., if we take the set

$$A = \{\tfrac{1}{n} : n \in \mathbb{N}\},$$

we notice that with the usual order $\leqslant$, we get $M(A) = \varnothing$.

Definition 1.44 (Well-founded order)**.** A partial order $\preccurlyeq$ on A is *well-founded* if $M_{\preccurlyeq}(B) \neq \varnothing$ for any non-empty $B \subseteq A$.

This idea is the poset analogue of a well-order. Indeed, if A is totally ordered by $\preccurlyeq$, and $\preccurlyeq$ is well-founded, then $\preccurlyeq$ is well-ordered. Notice that being well-founded essentially means we do not have any infinitely descending chains like

$$\cdots \prec a_3 \prec a_2 \prec a_1.$$

Indeed, any infinite descending chain in a poset would give $M(\{a_1, a_2, \dots\}) = \varnothing$. (Proof: if $a_k \in M(\{a_1, a_2, \dots\})$, then there is no $a_{k+1} \prec a_k$, so the chain would not be infinite.) Conversely, if a poset A is not well-founded, then there is a non-empty subset $B \subseteq A$ such that $M(B) = \varnothing$. Given any $b \in B$ there must be $m_1 \in B$ such that $m_1 \prec b$ (since $b \notin M(B)$). But also there is $m_2 \prec m_1$ (since $m_1 \notin M(B)$). Similarly there is $m_3 \prec m_2$ (since $m_2 \notin M(B)$), and so on, giving rise to an infinite descending chain $\cdots \prec m_2 \prec m_1 \prec b$.

In any well-founded poset, we can do recursion. We first need to define the notion of an *initial segment*.

Definition 1.45 (Initial segment)**.** Let $(A, \preccurlyeq)$ be a well-founded poset, and pick $a \in A$. The *initial segment* of a, denoted $I[a]$, is the set of all points strictly smaller than a under the partial order, i.e.,

$$I[a] := \{t \in A : t \prec a\}.$$

Notation. A notation which we could have introduced earlier is A^B, which denotes the set of all total functions from B to A. We also have a more general notation for unions. We write $\bigcup_{x \in X} A_x$ for the set

$$\{a : (\exists x \in X)(a \in A_x)\},$$

i.e., the set of all a's such that a is in some A_x for some $x \in X$.

The reason we need these two pieces of notation is because we will be making use of the set $\bigcup_{a \in A} X^{I[a]}$ in the statement of the next theorem. This is the set of all total functions from some initial segment $I[a]$ to a set X.

Theorem 1.46 (Recursion theorem)**.** *Suppose A is a set partially ordered by the well-founded partial order $\preccurlyeq$, let X be any set, and let $g\colon A \times S \to X$ be any total function, where $S := \bigcup_{a \in A} X^{I[a]}$. Then there exists a unique total function $f\colon A \to X$ such that*

$$f(a) = g(a, f \restriction I[a])$$

for all $a \in A$.

In other words, f is allowed to appear in the definition of $f(a)$, as long as we restrict its inputs to points in $I[a]$, i.e., to those points $t \in A$ which are strictly less than a under the partial order.

Just as with theorem 1.41, we will not prove this here, because the proof is technical. The interested reader can find a proof in any standard textbook on set theory.

Before we illustrate this theorem with some examples, we will restate this theorem (less generally) in terms of something called the ranked partition of a poset.

Definition 1.47 (Ranked poset)**.** A poset $(A, \preccurlyeq)$ is said to be *ranked* if there are no chains of infinite length between two elements, i.e., if $a \prec b$, we can only ever have finitely many a_i's between them:

$$a = a_1 \prec a_2 \prec a_2 \prec \cdots \prec a_n \prec b.$$

Example 1.48*.* Recall example 1.28, where we had the set

$$A = \{1 - \tfrac{1}{n} : n \in \mathbb{N}\} \cup \{1\} = \{0, \tfrac{1}{2}, \tfrac{2}{3}, \tfrac{3}{4}, \tfrac{4}{5}, \ldots, 1\}.$$

This is *not* a ranked poset, since $0 < 1$, and there is a chain of infinite length between them, i.e.,

$$0 < \tfrac{1}{2} < \tfrac{2}{3} < \cdots < 1.$$

This turns out to be the poset analogue of being finitely inducible. Indeed, if $\preccurlyeq$ is a total order which is ranked, then it is finitely inducible. If a poset is ranked and well-founded, we can partition its elements as follows.

Definition 1.49 (Ranked partition)**.** Let A be a set ordered by a ranked well-founded partial order $\preccurlyeq$. Define $A_n : \mathbb{N} \to \wp A$ by recursion as

$$A_n := M\Big(A \setminus \bigcup_{k<n} A_k\Big).$$

Then the *ranked partition of A* is the family of sets $\mathcal{A} = \{A_0, A_1, \ldots\}$.

Example 1.50*.* Consider the poset illustrated in the Hasse diagram in figure 9. As we can see, the ranked partition simply distinguishes between the different "storeys" of the Hasse diagram. Indeed, we have

$$A_0 = M\Big(A \setminus \bigcup_{k<0} A_k\Big) = M(A \setminus \varnothing) = M(A)$$

(since there is no $k < 0$ in $\mathbb{N}$), so the fist set in the partition is simply the minimal set of A. Next,

$$A_1 = M\Big(A \setminus \bigcup_{k<1} A_k\Big) = M(A \setminus A_0),$$

which is the minimal set of A once we remove the bottom layer, i.e., the second "storey". Similarly, $A_2 = M(A \setminus (A_0 \cup A_1))$, and so on.

This actually is a 'partition' of the set A, i.e., a collection of non-overlapping sets which together make up the set A. In other words, we have:

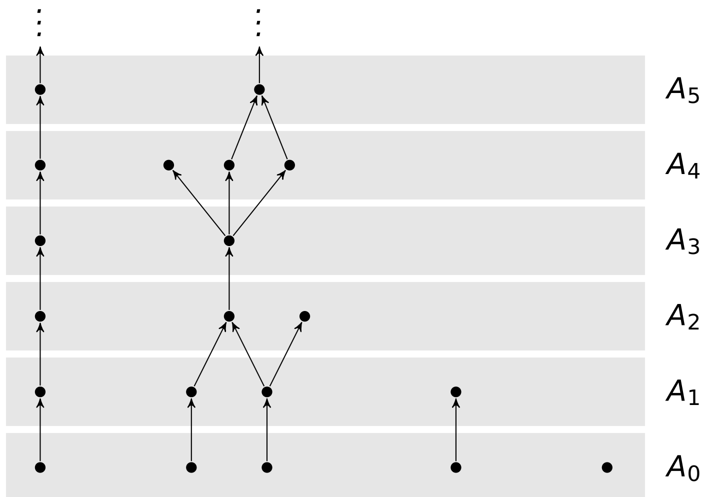

Figure 9: The Hasse diagram of a poset A and its ranked partition

Theorem 1.51. *Let A be a set, let $\preccurlyeq$ be a ranked well-founded order on A, and let $\mathcal{A} = \{A_0, A_1, \dots\}$ be its ranked partition. Then*

(i) $A_i \cap A_j = \varnothing$ *for all* $i \neq j$, *and*

(ii) $A = A_0 \cup A_1 \cup \cdots = \bigcup_{k\in\mathbb{N}} A_k$.

Proof. For (i), suppose $i < j$. Notice that for any $B \subseteq A$, by definition, if $x \in M(B)$, we must have $x \in B$. Now if $x \in A_j$, we must have

$$\begin{aligned} & x \in M\Big(A \setminus \bigcup_{k<j} A_k\Big) \\ \implies & x \in A \setminus \bigcup_{k<j} A_k \\ \implies & x \in A \quad \text{and} \quad x \notin \bigcup_{k<j} A_k. \\ \implies & x \in A \quad \text{and} \quad x \notin A_k \quad \text{for any } k < j, \end{aligned}$$

and in particular we get that $x \notin A_i$. Thus $A_i \cap A_j = \varnothing$.

Now for (ii), we're saying that eventually each $a \in A$ appears in some A_i. Suppose (for contradiction) that $S = \{a \in A : a \notin A_i \text{ for any } i\}$ is non-empty, and let $m \in M(S)$ (by well-foundedness). Then there exists some $t \in A$ such that $t \prec m$, and $t \in A_i$ for some i (otherwise $m \notin M(S)$). Let

$$t = t_i \prec t_{i+1} \prec \cdots \prec t_j = m$$

be a maximal chain joining t to m, i.e., there is no $s \in A$ which can be inserted between t_k and t_{k+1} (i.e., $t_k \prec s \prec t_{k+1}$) for any $i \leqslant k < j$. Such a chain exists, since we can start from $t \prec m$, and insert elements t_k between them until we reach a point where we can no longer do so (we are guaranteed to stop since A is ranked). Now let $T = \{t = t_i, t_{i+1}, \dots, t_j = m\}$, this is a well-ordered, finitely inducible chain under $\preccurlyeq$. Using induction, we can show that $t_k \in A_k$ for each $t_k \in T$, and in particular, we get that $m \in A_j$, which contradicts that $m \in S$. □

We can use the idea of the ranked partition to restate the recursion theorem.

Theorem 1.52 (Recursion theorem for ranked posets)**.** *Let $(A, \preccurlyeq)$ be a ranked well-founded poset with ranked partition $\mathcal{A} = \{A_0, A_1, \dots\}$, let X be any set, let $x\colon A_0 \to X$ be a total function, and let $g\colon A \times S \to X$ be any total function, where $S := \bigcup_{k\in\mathbb{N}} X^{A_k}$. Then there exists a unique total function $f\colon A \to X$ such that*

$$f(a) = \begin{cases} x(a) & \text{if } a \in A_0 \\ g\big(a, f \restriction \bigcup_{k<n} A_k\big) & \text{if } a \in A_n,\ n > 0 \end{cases}$$

for all $a \in A$.

This looks more like theorem 1.41. Indeed, notice that if we let $A = \mathbb{N}$ with the usual order, then $A_0 = \{0\}$, and $\bigcup_{k<n} A_k = \{0\} \cup \{1\} \cup \cdots \cup \{n^-\} = n$. Also, if g is defined by a fixed expression in a for all a, then the only term that can appear in g is $f(a^-)$, since this equality must hold for all a, and in particular, when $a \in A_1 = \{1\}$, we can only have $f(0) = f(a-1)$ appearing. Thus in the case where g is defined by the same expression for all a, we must have $g(a, f \restriction \{a^-\})$.

We can also obtain a version of theorem 1.41 with two base cases. If we order $\mathbb{N}$ such that 0 and 1 are not comparable, i.e., as in

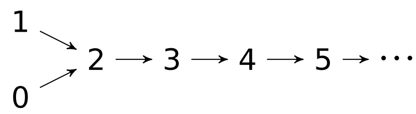

Figure 10: Partial order on $\mathbb{N}$ for two base cases coming from $\{0, 1\} \in A_0$

figure 10, the theorem becomes

$$f(n) = \begin{cases} x(n) & \text{if } n \in A_0 = \{0, 1\} \\ g\big(n, f \restriction \bigcup_{k<r} A_k\big) & \text{if } a \in A_r,\ r > 0, \end{cases}$$

and if we let $x(0) = x_0$, $x(1) = x_1$, and assume that g is defined by some fixed expression in terms of n, then in the case that $n = 2$, only $f(n-1)$ and $f(n-2)$ can make an appearance, which means that we must have

$$f(n) = \begin{cases} x_0 & \text{if } n = 0 \\ x_1 & \text{if } n = 1 \\ g(n, f \restriction \{n-1, n-2\}) & \text{otherwise.} \end{cases}$$

Example 1.53 (Fibonacci numbers)*.* The Fibonacci sequence is defined by $f(0) = 0$, $f(1) = 1$, and $f(n) = f(n-1) + f(n-2)$ for $n \geqslant 2$, i.e., each term is the sum of the previous two. This defines a valid function since it agrees with the form above. The first few values of f are given below.

n	0	1	2	3	4	5	6	7	8	9	10
$f(n)$	0	1	1	2	3	5	8	13	21	34	55

Example 1.54 (Binomial coefficients)*.* Another famous example is the number $\binom{n}{k}$ of k-subsets of $n = \{0, \ldots, n-1\}$, i.e.,

$$\binom{n}{k} := \#\{A \in \wp n : \#A = k\}.$$

For example, $\binom{5}{3} = 10$ since there are 10 subsets of $5 = \{0, 1, 2, 3, 4\}$ with size 3, i.e.,

$$\#\left\{\begin{matrix} \{0,1,2\}, \{0,1,3\}, \{0,1,4\}, \{0,2,3\}, \{0,2,4\}, \\ \{0,3,4\}, \{1,2,3\}, \{1,2,4\}, \{1,3,4\}, \{2,3,4\} \end{matrix}\right\} = 10.$$

It can be shown that $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ when $n \geqslant 1$. Indeed, let $C(n,k)$ denote the set $\{A \in \wp n : \#A = k\}$ which defines $\binom{n}{k}$, so that $\binom{n}{k} = \#C(n,k)$. Notice that

$$\begin{aligned} C(n,k) &= \{S \in C(n,k) : k^- \in S\} \cup \{S \in C(n,k) : k^- \notin S\} \\ &= \{S \cup \{k^-\} : S \in C(n-1,k-1)\} \cup C(n-1,k), \end{aligned}$$

and since for disjoint finite sets (i.e., sets with $A \cap B = \varnothing$), we have that $\#(A \cup B) = \#A + \#B$, we get that

$$\begin{aligned} \#C(n,k) &= \#\{S \cup \{k^-\} : S \in C(n-1,k-1)\} + \#C(n-1,k) \\ &= \binom{n-1}{k-1} + \binom{n-1}{k}. \end{aligned}$$

When arranged in a triangular array, we get the famous "Pascal's triangle", illustrated in figures 11 and 12. What the formula

$$\begin{array}{ccccccccccc} &&&&& \binom{0}{0} &&&&& \\ &&&& \binom{1}{0} && \binom{1}{1} &&&& \\ &&& \binom{2}{0} && \binom{2}{1} && \binom{2}{2} &&& \\ && \binom{3}{0} && \binom{3}{1} && \binom{3}{2} && \binom{3}{3} && \\ & \binom{4}{0} && \binom{4}{1} && \binom{4}{2} && \binom{4}{3} && \binom{4}{4} & \\ \binom{5}{0} && \binom{5}{1} && \binom{5}{2} && \binom{5}{3} && \binom{5}{4} && \binom{5}{5} \\ \iddots &&&&& \vdots &&&&& \ddots \end{array}$$

Figure 11: Pascal's triangle, in terms of the numbers $\binom{n}{k}$

$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ tells us is that each entry in this triangular array is the sum of the two above it.

We can choose to define these numbers in a recursive way instead, taking this formula as inspiration. The domain of such a function, let's call it c, should be pairs (n,k) such that $k \leqslant n$, i.e., we let

$$A = \{(n,k) \in \mathbb{N}^2 : k \leqslant n\},$$

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
\iddots & & & & & \vdots & & & & & \ddots
\end{array}
$$

Figure 12: Pascal's triangle, with the numbers evaluated

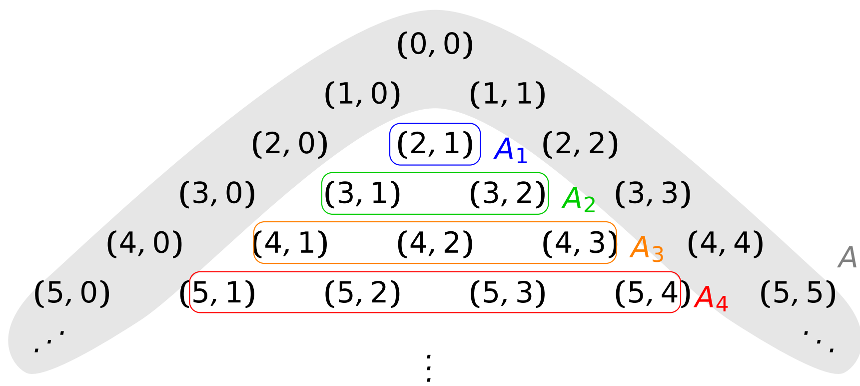

Figure 13: Ranked partition for $c(n, k)$

and we will have $c\colon A \to \mathbb{N}$. A possible base case for c is to take the stream of 1's (i.e. the "sides" of the triangle), so that the inside can be filled in with the "sum of the two above" formula. This corresponds to the definition

$$c(n,k) = \begin{cases} 1 & \text{if } k = 0 \text{ or } n = k \\ c(n-1,k-1) + c(n-1,k) & \text{otherwise.} \end{cases}$$

We must therefore define the partial order on A by the following properties:

- $(n, 0) \preccurlyeq a$ and $(n, n) \preccurlyeq a$ for all $n \in \mathbb{N}$ and $a \in A$
- $(n-1, k) \prec (n, \ell)$ for any $n, k, \ell \in \mathbb{N}$ (where $k < n$ and $\ell \leqslant n$).

This way, our ranked partition is that illustrated in figure 13. Of course, one should verify that this is a well-founded, ranked partial order. To make our definition look more like theorem 1.46, we have

$$c(n,k) = \begin{cases} 1 & \text{if } (n,k) \in A_0 \\ g(n, c \restriction \bigcup_{t<r} A_t) & \text{if } (n,k) \in A_r,\ r > 0, \end{cases}$$

where it is clear that because of the order, $(n-1,k-1),(n-1,k) \in \bigcup_{t<r} A_t$ for all (n,k), so our function definition is valid.

Next we discuss induction on posets. We saw in the section on induction that there are some strong requirements in order for us to prove something by induction; we saw that being a chain was not even enough, we needed a well-order which was finitely inducible. How can we hope to do induction on a poset? Well as we have seen, we have introduced analogues of these ideas for posets:

Chain		Poset
well-ordered	$\longleftrightarrow$	well-founded
$\min A$	$\longleftrightarrow$	$M(A)$
finitely inducible	$\longleftrightarrow$	ranked

Using the ranked partition of a poset, we can create a chain on which we can do induction.

Theorem 1.55. *Let A be a set, let $\preccurlyeq$ be a ranked, well-founded order on A, and let $\mathcal{A} = \{A_0, A_1, \ldots\}$ be its ranked partition. Then $(\mathcal{A}, \ll)$ is a well-ordered finitely inducible chain, where $\ll$ is defined by $A_i \ll A_j \iff i \leqslant j$.*

Proof. This follows immediately by applying the fact that $(\mathbb{N}, \leqslant)$ is well-ordered finitely inducible chain to the subscripts of the A_i. Indeed, this is a chain because we simply need to compare the subscripts of any two A_i, A_j, thus any two are comparable. Similarly, it is well-ordered because any subset $\{A_{i_1}, A_{i_2}, \ldots\}$ of $\mathcal{A}$ has minimal element $A_{\min\{i_1,i_2,\ldots\}}$, and finally it is finitely inducible because each A_i is the successor of A_{i^-} apart from $\min(\mathcal{A}) = A_0$. □

This allows us to state a version of the induction theorem for posets.

Theorem 1.56 (Induction on posets)**.** *Let A be a set, let $\preccurlyeq$ be a a ranked, well-founded order on A, let $\mathcal{A} = \{A_0, A_1, \dots\}$ be its ranked partition, and let $\Phi(a)$ be a statement about some $a \in A$. If*

(i) $\Phi(a)$ *is true for all $a \in A_0$, and*

(ii) for all $A_k \in \mathcal{A}$, if $\Phi(a)$ is true for all $a \in A_k$, then $\Phi(a)$ is true for all $a \in A_{k+1}$,

then $\Phi(a)$ is true for all $a \in A$.

Proof. This is very straightforward, simply define

$$\Psi(k) = \text{“}\Phi(a) \text{ is true for all } a \in A_k\text{”},$$

for all $k \in \mathbb{N}$, and apply theorem 1.35. □

Similarly, the "strong" analogue (theorem 1.40) is

Theorem 1.57 (Strong induction on posets)**.** *Let A be a set, let $\preccurlyeq$ be a ranked, well-founded order on A, let $\mathcal{A} = \{A_0, A_1, \dots\}$ be its ranked partition, and let $\Phi(a)$ be a statement about some $a \in A$. If*

(i) $\Phi(a)$ *is true for all $a \in A_0$, and*

(ii) for all $A_k \in \mathcal{A}$, if $\Phi(a)$ is true for all $a \in \bigcup_{t<k} A_t$, then $\Phi(a)$ is true for all $a \in A_{k+1}$,

then $\Phi(a)$ is true for all $a \in A$.

The proof is identical to that of theorem 1.56, but it simply invokes theorem 1.40 rather than theorem 1.34 with the same $\Psi(k)$.

Remark 1.58*.* The essential idea behind what was discussed in this section is the following. Any ranked well-founded poset is equivalent to a well-ordered finitely inducible chain if we consider elements on the "same storey" of the Hasse diagram equivalent by grouping them together as one element (i.e., elements in the same minimal set).

Indeed, as we have seen, any well-founded poset gives us a well-ordered finitely inducible chain, but conversely, given any partition $\{A_1, A_2, \dots\}$ of a set A which is a well-ordered finitely inducible chain (by $\ll$, say), we can simply convert it to a ranked well-founded poset by defining the partial order

$$a \preccurlyeq b \qquad \iff \qquad a \in A_i, \quad b \in A_j \quad \text{and} \quad i \leqslant j.$$

Thus there is a very clear back-and-forth between the two.

LANGUAGES AND STRINGS

Now we define the objects which are the main focus of this set of notes: languages and strings.

Definition 1.59 (Alphabet)**.** An *alphabet* is a non-empty finite set, whose elements we call *symbols*.

By 'symbols' here we mean that the members of an alphabet are typically thought of as representing letters, characters or digits; essentially things which we are used to juxtaposing. For example, a common alphabet is the binary alphabet $\{\texttt{0}, \texttt{1}\}$, or the English alphabet $\{\texttt{a}, \texttt{b}, \texttt{c}, \dots, \texttt{z}\}$. We usually use the letter Σ to denote an alphabet.

Definition 1.60 (String)**.** A *string* drawn from an alphabet Σ is a tuple whose entries are members of Σ.

For example, $(\texttt{1}, \texttt{0}, \texttt{1}, \texttt{1}, \texttt{0}, \texttt{1})$ and $(\texttt{1}, \texttt{1}, \texttt{1})$ are strings drawn from $\{\texttt{0}, \texttt{1}\}$, and $(\texttt{s}, \texttt{t}, \texttt{e}, \texttt{f}, \texttt{a}, \texttt{n}, \texttt{i}, \texttt{a})$ is a string drawn from the alphabet $\{\texttt{a}, \texttt{b}, \texttt{c}, \dots, \texttt{z}\}$. For convenience, we will stop using tuple notation and simply juxtapose the symbols. So instead of $(\texttt{1}, \texttt{0}, \texttt{1}, \texttt{1}, \texttt{0}, \texttt{1})$ we write $\texttt{101101}$, and similarly we write $\texttt{stefania}$ instead of $(\texttt{s}, \texttt{t}, \texttt{e}, \texttt{f}, \texttt{a}, \texttt{n}, \texttt{i}, \texttt{a})$. We will denote strings using bold letters such as $\boldsymbol{s}$ or $\boldsymbol{\sigma}$ to distinguish them from members of the alphabet.

Definition 1.61 (The Empty String)**.** The *empty string* is the tuple of length zero, and is denoted by ϵ.

This definition might cause some discomfort from a set theoretic point of view, but remember how we encoded tuples as functions; ϵ here is a bijection equal to the empty set.

Definition 1.62 (Kleene-closure)**.** Let Σ be an alphabet. The *Kleene-closure* of Σ, denoted Σ^*, is the set of strings given by

$$\Sigma^* := \bigcup_{n \in \mathbb{N}} \Sigma^n$$

where $\Sigma^0 = \{\epsilon\}$.