Appendix A

Cardinality

Cardinality — that is, Cantor's theory of infinite cardinal numbers — does not play a direct rôle in algorithmic information theory, or more generally in the study of computability, since all the sets that arise there are *enumerable*. However, the proof of Cantor's Theorem below, using Cantor's 'diagonal method', is the predecessor, or model, for our proof of the Halting Theorem and other results in Algorithmic Information Theory. The idea also lies behind Gödel's Unprovability Theorem, that in any non-trivial axiomatic system there are propositions that can neither be proved nor disproved.

A.1 Cardinality

Definition A.1. *Two sets X and Y are said to have the same cardinality, and we write*

$$\#(X) = \#(Y),$$

if there exists a bijection $f : X \to Y$.

When we use the '=' sign for a relation in this way we should verify that the relation is reflexive, symmetric and transitive. In this case that is trivial.

Proposition A.1. *1.* $\#(X) = \#(X)$*;*

2. $\#(X) = \#(Y) \Longrightarrow \#(Y) = \#(X)$*;*

3. $\#(X) = \#(Y)$ & $\#(Y) = \#(Z) \Longrightarrow \#(X) = \#(Z)$.

By convention, we write

$$\#(\mathbb{N}) = \aleph_0.$$

Definition A.2. *We say that the cardinality of X is less than or equal to the cardinality of Y, and we write*

$$\#(X) \leq \#(Y),$$

if there exists a injection $f : X \to Y$.

We say that the cardinality of X is (strictly) less than the cardinality of Y, and we write

$$\#(X) < \#(Y),$$

if $\#(X) \leq \#(Y)$ and $\#(X) \neq \#(Y)$.

Proposition A.2. *1. $\#(X) \leq \#(X)$;*

2. $\#(X) \leq \#(Y)$ & $\#(Y) \leq \#(Z) \Longrightarrow \#(X) \leq \#(Z)$.

Again, these follows at once from the definition of injectivity.

A.1.1 Cardinal arithmetic

Addition and multiplication of cardinal numbers is defined by

$$\#(X) + \#(Y) = \#(X + Y), \quad \#(X) \times \#(Y) = \#(X \times Y),$$

where $X + Y$ is the disjoint union of X and Y (ie if X and Y are not disjoint we take copies that are).

However, these operations are not very useful; for if one (or both) of $\aleph_1, \aleph_2$ are infinite then

$$\aleph_1 + \aleph_2 = \aleph_1 \aleph_2 = \max(\aleph_1, \aleph_1).$$

The power operation is more useful, as we shall see. Recall that 2^X denotes the set of subsets of X. We set

$$2^{\#(X)} = \#(2^X).$$

A.2 The Schröder-Bernstein Theorem

Theorem A.1.

$$\#(X) \leq \#(Y) \ \& \ \#(Y) \leq \#(X) \Longrightarrow \#(X) = \#(Y).$$

Proof ▶. By definition there exist injective maps

$$f : X \to Y,\ g : Y \to X.$$

We have to construct a bijection

$$h : X \to Y.$$

To simplify the discussion, we assume that X and Y are disjoint (taking disjoint copies if necessary).

Given $x_0 \in X$, we construct the sequence

$$y_0 = fx_0 \in Y,\ x_1 = gy_0 \in X,\ y_1 = fx_1 \in Y, \ldots.$$

There are two possibilities:

(i) The sequence continues indefinitely, giving a singly-infinite chain in X:

$$x_0, y_0, x_1, y_1, x_2, \ldots$$

(ii) There is a repetition, say

$$x_r = x_s$$

for some $r < s$. Since f and g are injective, it follows that the first repetition must be

$$x_0 = x_r,$$

so that we have a loop

$$x_0, y_0, x_1, y_1, \ldots, x_r, y_r, x_0.$$

In case (i), we may be able to extend the chain backwards, if $x_0 \in \mathrm{im}(g)$. In that case we set

$$x_0 = gy_{-1},$$

where y_{-1} is unique since g is injective.

Then we may be able to go further back:

$$y_{-1} = fx_{-1},\ x_{-2} = gy_{-1},\ \ldots.$$

There are three possibilities:

(A) The process continues indefinitely, giving a doubly-infinite chain

$$\ldots, x_{-n}, y_{-n}, x_{-n+1}, y_{-n+1}, \ldots, x_0, y_0, x_1, \ldots.$$

(B) The process ends at an element of X, giving a singly-infinite chain

$$x_{-n}, y_{-n}, x_{-n+1}, \ldots .$$

(C) The process ends at an element of Y, giving a singly-infinite chain

$$y_{-n}, x_{-n+1}, y_{-n+1}, \ldots .$$

It is easy to see that these chains and loops are disjoint, partitioning the union $X + Y$ into disjoint sets. This allows us to define the map h on each chain and loop separately. Thus in the case of a doubly-infinite chain or a chain starting at an element $x_{-n} \in X$, or a loop, we set

$$hx_r = y_r;$$

while in the case of a chain starting at an element $y_{-n} \in Y$ we set

$$hx_r = y_{r-1}.$$

Putting these maps together gives a bijective map

$$h : X \to Y.$$

A.3 Cantor's Theorem

Theorem A.2. *The number of elements of a set is strictly less than the number of subsets of the set:*

$$\#(X) < \#(2^X).$$

Proof ▶. We have to show that $\#(X) \leq \#(2^X)$ but $\#(X) \neq \#(2^X)$.

There is an obvious injection $X \to 2^X$, namely

$$x \mapsto \{x\}.$$

Hence

$$\#(X) \leq \#(2^X).$$

Suppose there is a surjection

$$f : X \to 2^X.$$

Let

$$S = \{x \in X : x \notin f(x)\}.$$

Since f is surjective, there exists an element $s \in X$ such that

$$S = f(s).$$

We ask the question: *Does the element s belong to the subset S, or not?* If $s \in S$, then from the definition of S,

$$s \notin f(s) = S.$$

On the other hand, if $s \notin S$, then again from the definition of S.

$$s \in S.$$

Either way, we encounter a contradiction. Hence our hypothesis is untenable: there is no surjection, and so no isomorphism, $f : X \to 2^X$, ie

$$\#(X) \neq \#(2^X).$$

A.4 Comparability

Our aim in this Section is to prove any 2 sets X, Y are *comparable*, ie

$$\text{either } \#(X) \leq \#(Y) \text{ or } \#(Y) \leq \#(X).$$

To this end we introduce the notion of well-ordering.

Recall that a *partial order* on a set X is a relation $\leq$ such that

1. $x \leq x$ for all x;
2. $x \leq y \ \& \ y \leq z \Longrightarrow x \leq z$;
3. $x \leq y \ \& \ y \leq x \Longrightarrow x = y$.

A partial order is said to be a *total order* if in addition, for all $x, y \in X$,

4. either $x \leq y$ or $y \leq x$.

A total order is said to be a *well-ordering* if

5. every non-empty subset $S \subset X$ has a least element $\mu(S) \in S$.

Examples:

1. The natural numbers $\mathbb{N}$ are well-ordered.

2. The integers $\mathbb{Z}$ are not well-ordered, since $\mathbb{Z}$ itself does not have a least element.

3. The set of positive reals $\mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\}$ is not well-ordered, since the set $S = \{x > 0\}$ does not have a least element *in* S.

4. The set $N \times N$ with the lexicographic ordering

$$(m, n) \leq (m', n') \text{ if } m < m' \text{ or } m = m' \ \& \ n \leq n'$$

is well-ordered. To find the least element (m, n) in a subset $S \subset \mathbb{N} \times \mathbb{N}$ we first find the least m occuring in S; and then among the pairs $(m, n) \in S$ we find the least n.

5. The disjoint sum $\mathbb{N} + \mathbb{N}$, with the ordering under which every element of the first copy of $\mathbb{N}$ is less than every element of the second copy of $\mathbb{N}$, is well-ordered.

It follows at once from the definition that every subset $S \subset X$ of a well-ordered set is well-ordered.

A well-ordered set X has a first (least) element

$$x_0 = \mu(X).$$

Unless this is the only element, X has a second (next least) element

$$x_1 = \mu(X \setminus \{x_0\}).$$

Simitarly, unless these are the only elements, X has a third element

$$x_2 = \mu(X \setminus \{x_0, x_1\}),$$

and so on. Moreover after all these elements $x_0, x_1, x_2, \ldots$ (assuming they have not exhausted X) there is a next element

$$x_\omega = \mu(X \setminus \{x_0, x_1, x_2, \ldots\}).$$

Then comes the element

$$x_{\omega+1} = \mu(X \setminus \{x_0, x_1, x_2, \ldots, x_\omega\}),$$

and after that elements $x_{\omega+2}, x_{\omega+3}, \ldots$.

Proposition A.3. *There is at most one order-preserving isomorphism between 2 well-ordered sets.*

Proof ▶. Suppose

$$f, g : X \to Y$$

are 2 isomorphisms between the well-ordered sets X, Y. Let

$$z = \mu\left(\{x \in X : fx \neq gx\}\right).$$

In other words, z is the first point at which the maps f and g diverge.

We may assume without loss of generality that

$$fz < gz.$$

Since g is an isomorphism,

$$fz = gt$$

for some element $t \in X$. But now, since g is order-preserving,

$$gt < gz \Longrightarrow t < z \Longrightarrow gt = ft \Longrightarrow fz = ft \Longrightarrow z = t \Longrightarrow gz = gt = fy,$$

contrary to hypothesis. We conclude that $f = g$, ie $f(x) = g(x)$ for all $x \in X$.

Although we shall make no use of this, we associate an *ordinal number* (or just *ordinal*) to each well-ordered set. By the last Proposition, two well-ordered sets have the same ordinal number if and only if they are order-isomorphic.

A subset $I \subset X$ in a partially-ordered set X is called an *initial segment* if

$$x \in I \ \& \ y \leq x \Longrightarrow y \in I.$$

It is easy to see that the set

$$I(x) = \{y \in X : y < x\}$$

(where $x < y$ means $x \leq y$ but $x \neq y$) is an initial segment in X for each $x \in X$.

In a well-ordered set *every initial subset $I \subset X$, except X itself, is of this form*; for it is easily seen that

$$I = I(x),$$

where

$$x = \mu(X \setminus I).$$

If an element x of a well-ordered set X is not the greatest element of X then it has an immediate successor x', namely

$$x' = \mu(\{y \in X : y > x\}).$$

But not every element $x \in X$ (apart from the minimal element x_0) need be a successor element. We call an element $x \neq x_0$ with no immediate predecessor a *limit element*.

Lemma 6. *If x is a limit element then*

$$I(x) = \cup_{y<x} I(y).$$

Proof ▶. Certainly

$$J = \cup_{y<x} I(y)$$

is an initial segment. Suppose

$$J = I(z),$$

where $z \leq x$.

If $z < x$ then x must be the immediate successor to z. For suppose $z < t < x$. Then

$$z \in I(t) \subset J = I(z),$$

contrary to the definition of the initial segment $I(z)$.

Lemma 7. *Suppose X, Y are well-ordered sets. Then X is order-isomorphic to at most one initial segment I of Y.*

Proof ▶. If X is isomorphic to two different initial segments $I, J \subset Y$ then there are two different order-preserving injective maps

$$f, g : X \to Y.$$

Suppose these maps diverge at $z \in X$:

$$z = \mu(\{x \in X : fx \neq gx\}).$$

We may assume without loss of generality that

$$fz < gz.$$

But then $fz \in J = \mathrm{im}(g)$, and so

$$fz = gt$$

for some $t \in X$. Thus

$$gt < gz \Longrightarrow t < z \Longrightarrow ft = gt = fz \Longrightarrow t = z \Longrightarrow fz = gt = gz,$$

contrary to the definition of z.

Corollary A.1. *If there is such an isomorphism $f : X \to I \subset Y$ then it is unique.*

This follows at once from the Proposition above.

Proposition A.4. *Suppose X, Y are well-ordered sets. Then either there exists an order-preserving injection*

$$f : X \to Y,$$

or there exists an order-preserving injection

$$f : Y \to X.$$

Proof ▶. Suppose there is an order-preserving injection

$$f_x : I(x) \to J$$

onto an initial segment J of Y for every element $x \in X$. If $J = Y$ for some x then we are done. Otherwise

$$J = I(y),$$

where (from the last Lemma) y is completely determined by x, say

$$y = f(x).$$

Then it follows easily that

$$f : X \to Y$$

is an order-preserving injection.

Suppose there is no such map f_x for some $x \in X$. Let $z \in X$ be the smallest such element. If $u < v < z$ then the corollary above shows that f_u is the restriction of f_v to $I(u)$. It follows that the maps f_u for $u < z$ 'fit together' to give an order-preserving injection

$$f : I(z) \to Y.$$

More precisely, if $x < z$ then (from the definition of z) there is an order-preserving isomorphism

$$f_x : I(x) \to I(y),$$

where $y \in Y$ is well-defined. We set $fx = y$ to define an order-preserving injection

$$f : I(z) \to Y$$

onto an initial segment of Y, contrary to the definition of z.

If X, Y are two well-ordered sets, we say that the ordinal of X is $\leq$ the ordinal of Y if there exists an order-preserving injection $f : X \to Y$. Ordinals are even further from our main theme than cardinals; but nevertheless, every young mathematician should be at least vaguely familiary with them.

We denote the ordinal of $\mathbb{N}$ with the usual ordering (which we have observed is a well-ordering) by ω:

$$\omega = \{0, 1, 2, \dots\}.$$

If X, Y are well-ordered sets then so is the disjoint union $X + Y$, taking the elements of X before those of Y. This allows us to add ordinals. For example

$$\omega + 1 = \{0, 1, 2, \dots, \omega\},$$

where we have added another element after the natural numbers. It is easy to see that

$$\omega + 1 \neq \omega :$$

the two ordered sets are not order-isomorphic, although both are enumerable. So different ordinals may correspond to the same cardinal. Of course this is not true for finite numbers; there is a one-one correspondence between finite cardinals and finite ordinals.

Note that addition of ordinals is not commutative, eg

$$1 + \omega = \omega,$$

since adding an extra element at the beginning of $\mathbb{N}$ does not alter its ordinality.

A.4.1 The Well Ordering Theorem

The Axiom of Choice states that for every set X we can find a map

$$c : 2^X \to X$$

such that

$$c(S) \in S$$

for every non-empty subset $S \subset X$.

We call such a map c a *choice function* for X.

The Well Ordering Theorem states that *every set X can be well-ordered.*

Proposition A.5. *The Axiom of Choice and the Well Ordering Theorem are equivalent.*

Proof ▶. If X is a well-ordered set then there is an obvious choice function, namely

$$c(S) = \mu(S).$$

The converse is more difficult.

Suppose c is a choice function for X. Let us say that a well-ordering of a subset $S \subset X$ has property $\mathcal{P}$ if

$$\mu(S \setminus I) = c(X \setminus I)$$

for every initial segment $I \subset S$ (except S itself).

We note that such a subset S must start with the elements $x_0, x_1, x_2, \ldots, x_\omega, \ldots$ unless S is exhausted earlier.

Lemma 8. ?? *A subset $S \subset X$ has at most one well-ordering with property $\mathcal{P}$.*

Proof ▶. Suppose there are two such well-orderings on S. Let us denote them by $<$ and $\subset$, respectively. If $x \in S$ let us write

$$I_<(x) \equiv I_\subset(x)$$

to mean that not only are these initial subsets the same but they also carry the same orderings.

If this is true of every $x \in S$ then the two orderings are the same, since

$$u < v \Longrightarrow u \in I_<(v) = I_\subset(v) \Longrightarrow u \subset v.$$

If however this is not true of all x, let z be the least such *according to the first ordering*:

$$z = \mu_< \left(\{ x \in X : I < (x) \not\equiv I_\subset(x) \} \right).$$

If $u, v \in I_<(z)$ then

$$u < v \Longrightarrow u \in I_<(v) = I_\subset(v) \Longrightarrow u \subset v.$$

It follows that $I_<(z)$ is also an initial segment in the second ordering, say

$$I_<(z) = I_\subset(t).$$

Hence

$$z = \mu_<(X \setminus I_<(z)) = c(X \setminus I_<(z)) = c(X \setminus I_\subset(z)) = \mu_< \left(X \setminus I_\subset(t) \right) = t.$$

Thus

$$I_<(z) = I_\subset(z).$$

Since, as we saw, the two orderings coincide, it follows that

$$I_<(z) \equiv I_\subset(z),$$

contrary to hypothesis. So there is no such z, and therefore the two orderings on S coincide.

Lemma 9. *Suppose the subsets $S, T \subset X$ both carry well-orderings with property $\mathcal{P}$. Then*

$$\text{either } S \subset T \text{ or } T \subset S.$$

Moreover, in the first case S is an initial segment of T, and in the second case T is an initial segment of S.

Proof ▶. As before, we denote the well-orderings on S and T by $<$ and $\subset$, respectively.

Consider the elements $x \in S$ such that the initial segment $I = I_<(x)$ in S is also an initial segment in T. By the last Lemma, the two orderings on I coincide.

If

$$I_<(x) = T$$

for some x we are done: T is an initial segment of S. Otherwise

$$I = I_<(x) = I_\subset(x),$$

with

$$x = c(X \setminus I).$$

Suppose this is true for all $x \in S$.

If S has a largest element s then

$$S = I_<(s) \cup \{s\} = I_\subset(s) \cup \{s\}.$$

Thus $S \subset T$, and either $S = T$,

$$S = I_\subset(s'),$$

where s' is the successor to s in T. In either case S is an initial segment of T.

If S does not have a largest element then

$$S = \cup_{x \in S} I_<(x) = \cup_{x \in S} I_\subset(x).$$

Thus S is again an initial segment of T; for

$$u \in S,\ v \in T,\ v \subset u \Longrightarrow v \in I_\subset(u) = I_<(u) \Longrightarrow v \in S.$$

Now suppose that $I_<(x)$ is *not* an initial segment of T for some $x \in S$. Let z be the smallest such element in S.

Since $I_<(z)$ is not an initial segment of T there is an element $tinT$ such that

$$t < z \ \& \ z \subset t.$$

We are now in a position to well-order X. Let us denote by $\mathcal{S}$ the set of subsets $S \subset X$ which can be well-ordered with property $\mathcal{P}$; and let

$$U = \cup_{S \in \mathcal{S}} S.$$

We shall show that U is well-ordered with property $\mathcal{P}$.

Firstly we define a total ordering on U. Suppose $u, v \in U$. There exists a set $S \in \mathcal{S}$ containing u, v; for if $u \in S_1, \ v \in S_2$, where $S_1, S_2 \in \mathcal{S}$ then by Lemma 9 either $S_1 \subset S_2$, in which case $u, v \in S_2$ or $S_2 \subset S_1$, in which case $u, v \in S_1$.

Also if $u, v \in S$ and T then by the same Lemma the two orderings are the same. Thus we have defined the order in U unambiguously.

To see that this is a well-ordering, suppose A is a non-empty subset of U; and suppose $a \in A$. Then $a \in S$ for some $S \in \mathcal{S}$. Let

$$z = \mu(A \cap S).$$

We claim that z is the smallest element of A. For suppose $t < z, \ t \in A$. Then $t \in I(z) \subset S$, and so $t \in A \cap S$, contradicting the minimality of z.

Finally, to see that this well-ordering of U has property $\mathcal{P}$, suppose I is an initial segment of U, $I \neq U$. Let z be the smallest element in $U \setminus I$; and suppose $z \in S$, where $S \in \mathcal{S}$. Then $I = I(z)$ is an initial segment in S, with

$$z = \mu_S(S \setminus I) = c(X \setminus I),$$

since S has property $\mathcal{P}$.

Thus

$$U \in \mathcal{S}.$$

If $U \neq X$ let

$$u = c(X \setminus U),$$

and set

$$V = U \cup \{u\}.$$

We extend the order on U to V by making u the greatest element of V. It is a straightforward matter to very that V is well-connected, and has property $\mathcal{P}$. It follows that

$$V \in \mathcal{S} \Longrightarrow V \subset U,$$

which is absurd.

Hence $U = X$, and so X is well-connected.

Now we can prove the Comparability Theorem.

Theorem A.3. *Any 2 sets X, Y are comparable, ie*

$$\textit{either } \#(X) \leq \#(Y) \textit{ or } \#(Y) \leq \#(X).$$

Proof ▶. Let us well-order X and Y. Then by Proposition A.4 either there exists an injection

$$j : X \to Y,$$

or there exists an injection

$$j : Y \to X.$$

In other words,

$$\#(X) \leq \#(Y) \text{ or } \#(Y) \leq \#(X).$$