

# Logic and Set Theory

February 2, 2018

<i>CONTENTS</i>	2
-----------------	---

## Contents

0	Miscellaneous	3
1	Propositional logic	4
2	Syntactic implication	6
3	Well-Orderings and Ordinals	10
4	Ordinals	14

## 0 Miscellaneous

Some introductory speech

## 1 Propositional logic

Let  $P$  denote a set of *primitive proposition*, unless otherwise stated,  $P = \{p_1, p_2, \dots\}$ .

**Definition.** The *language* or *set of propositions*  $L = L(P)$  is defined inductively by:

- (1)  $p \in L \forall p \in P$ ;
- (2)  $\perp \in L$ , where  $\perp$  is read as 'false';
- (3) If  $p, q \in L$ , then  $(p \implies q) \in L$ . For example,  $(p_1 \implies L)$ ,  $((p_1 \implies p_2) \implies (p_1 \implies p_3))$ .

Note that at this point, each proposition is only a finite string of symbols from the alphabet  $(, ), \implies, \perp, p_1, p_2, \dots$  and do not really mean anything (until we define so).

By *inductively define*, we mean more precisely that we set  $L_1 = P \cup \{\perp\}$ , and  $L_{n+1} = L_n \cup \{(p \implies q) : p, q \in L_n\}$ , and then put  $L = L_1 \cup L_2 \cup \dots$

Each proposition is built up *uniquely* from 1) and 2) using 3). For example,  $((p_1 \implies p_2) \implies (p_1 \implies p_3))$  came from  $(p_1 \implies p_2)$  and  $(p_1 \implies p_3)$ . We often omit outer brackets or use different brackets for clarity.

Now we can define some useful things:

- $\neg p$  (not  $p$ ), as an abbreviation for  $p \implies \perp$ ;
- $p \vee q$  ( $p$  or  $q$ ), as an abbreviation for  $(\neg p) \implies q$ ;
- $p \wedge q$  ( $p$  and  $q$ ), as an abbreviation for  $\neg(p \implies (\neg q))$ .

These definitions 'make sense' in the way that we expect them to.

**Definition.** A *valuation* is a function  $v : L \rightarrow \{0, 1\}$  s.t.

- (1)  $v(\perp) = 0$ ; (2)

$$v(p \implies q) = \begin{cases} 0 & v(p) = 1, v(q) = 0 \\ 1 & \text{else} \end{cases} \quad \forall p, q \in L$$

**Remark.** On  $\{0, 1\}$ , we could define a constant  $\perp$  by  $\perp = 0$ , and an operation  $\implies$  by  $a \implies b = 0$  if  $a = 1, b = 0$  and 1 otherwise. Then a valuation is a function  $L \rightarrow \{0, 1\}$  that preserves the structure  $(\perp \text{ and } \implies)$ , i.e. a homomorphism.

**Proposition.** (1) If  $v, v'$  are valuations with  $v(p) = v'(p) \forall p \in P$ , then  $v = v'$  (on  $L$ ).

(2) For any  $w : P \rightarrow \{0, 1\}$ , there exists a valuation  $v$  with  $v(p) = w(p) \forall p \in P$ . In short, a valuation is defined by its value on  $P$ , and any values will do.

*Proof.* (1) We have  $v(p) = v'(p) \forall p \in L_1$ . However, if  $v(p) = v'(p)$  and  $v(q) = v'(q)$  then  $v(p \implies q) = v'(p \implies q)$ , so  $v = v'$  on  $L_2$ . Continue inductively we have  $v = v'$  on  $L_n \forall n$ .

(2) Set  $v(p) = w(p) \forall p \in P$  and  $v(\perp) = 0$ : this defines  $v$  on  $L_1$ . Having defined  $v$  on  $L_n$ , use the rules for valuation to inductively define  $v$  on  $L_{n+1}$  so we can extend  $v$  to  $L$ .  $\square$

**Definition.** We say  $p$  is a *tautology*, written  $\models p$ , if  $v(p) = 1 \forall$  valuations  $v$ .  
Some examples:

(1)  $p \implies (q \implies p)$ : a true statement implies by anything. We can verify this by:

$v(p)$	$v(q)$	$v(q \implies p)$	$v(p \implies (q \implies p))$
1	1	1	1
1	0	1	1
0	1	0	1
0	0	1	1

So we see that this is indeed a tautology;

(2)  $(\neg\neg p) \implies p$ , i.e.  $((p \implies \perp) \implies \perp) \implies p$ , called the "law of excluded middle";

(3)  $[p \implies (q \implies r)] \implies [(p \implies q) \implies (p \implies r)]$ .

Indeed, if not then we have some  $v$  with  $v(p \implies (q \implies r)) = 1$ ,  $v((p \implies q) \implies (p \implies r)) = 0$ . So  $v(p \implies q) = 1$ ,  $v(p \implies r) = 0$ . This happens when  $v(p) = 1$ ,  $v(r) = 0$ , so also  $v(q) = 1$ . But then  $v(q \implies r) = 0$ , so  $v(p \implies (q \implies r)) = 0$ .

**Definition.** For  $S \subset L$ ,  $t \in L$ , say  $S$  *entails* or *semantically implies*  $t$ , written  $S \models t$  if  $v(s) = 1 \forall s \in S \implies v(t) = 1$ , for each valuation  $v$ .

("Whenever all of  $S$  is true,  $t$  is true as well.")

For example,  $\{p \implies q, q \implies r\} \models (p \implies r)$ . To prove this, suppose not: so we have  $v$  with  $v(p \implies q) = v(q \implies r) = 1$  but  $v(p \implies r) = 0$ . So  $v(p) = 1$ ,  $v(r) = 0$ , so  $v(q) = 0$ , but then  $v(p \implies q) = 0$ .

If  $v(t) = 1$  we say  $t$  is true in  $v$  or that  $v$  is a model of  $t$ .

For  $S \subset L$ ,  $v$  is a model of  $S$  if  $v(s) = 1 \forall s \in S$ . So  $S \models t$  says that every model of  $S$  is a model of  $t$ . For example, in fact  $\models t$  is the same as  $\emptyset \models t$ .

## 2 Syntactic implication

For a notion of 'proof', we will need axioms and deduction rules. As axioms, we'll take:

1.  $p \implies (q \implies p) \forall p, q \in L$ ;
2.  $[p \implies (q \implies r)] \implies [(p \implies q) \implies (p \implies r)] \forall p, q, r \in L$ ;
3.  $(\neg\neg p) \implies p \forall p \in L$ .

Note: these are all tautologies. Sometimes we say they are 3 axiom-schemes, as all of these are infinite sets of axioms.

As deduction rules, we'll take just *modus ponens*: from  $p$ , and  $p \implies q$ , we can deduce  $q$ .

For  $S \subset L$ ,  $t \in L$ , a *proof* of  $t$  from  $S$  consists of a finite sequence  $t_1, \dots, t_n$  of propositions, with  $t_n = t$ , s.t.  $\forall i$  the proposition  $t_i$  is an axiom, or a member of  $S$ , or there exists  $j, k < i$  with  $t_j = (t_k \implies t_i)$ .

We say  $S$  is the *hypotheses* or *premises* and  $t$  is the *conclusion*.

If there exists a proof of  $t$  from  $S$ , we say  $S$  *proves* or *syntactically implies*  $t$ , written  $S \vdash t$ .

If  $\phi \vdash t$ , we say  $t$  is a *theorem*, written  $\vdash t$ .

**Example.**  $\{p \implies q, q \implies r\} \vdash p \implies r$ .

we deduce by the following:

- (1)  $[p \implies (q \implies r)] \implies [(p \implies q) \implies (p \implies r)]$ ; (axiom 2)
- (2)  $q \implies r$ ; (hypothesis)
- (3)  $(q \implies r) \implies (p \implies (q \implies r))$ ; (axiom 1)
- (4)  $p \implies (q \implies r)$ ; (mp on 2,3)
- (5)  $(p \implies q) \implies (p \implies r)$  (mp on 1,4);
- (6)  $p \implies q$ ; (hypothesis)
- (7)  $p \implies r$ . (mp on 5,6)

**Example.** Let's now try to prove  $\vdash p \implies p$ . Axiom 1 and 3 probably don't help so look at axiom 2; if we make  $(p \implies q)$  and  $p \implies (q \implies r)$  something that's a theorem, and make  $p \implies r$  to be  $p \implies p$  then we are done. So we need to take  $p = p, q = (p \implies p), r = p$ . Now:

- (1)  $[p \implies ((p \implies p) \implies p)] \implies [(p \implies (p \implies p)) \implies (p \implies p)]$ ; (axiom 2)
- (2)  $p \implies ((p \implies p) \implies p)$ ; (axiom 1)
- (3)  $(p \implies (p \implies p)) \implies (p \implies p)$ ; (mp on 1,2)
- (4)  $p \implies (p \implies p)$ ; (axiom 1)
- (5)  $p \implies p$ . (mp on 3,4)

Proofs are made easier by:

**Proposition.** (2, deduction theorem)

Let  $S \subset L$ ,  $p, q \in L$ . Then  $S \vdash (p \implies q)$  if and only if  $(S \cup \{p\}) \vdash q$ .

*Proof.* Forward: given a proof of  $p \Rightarrow q$  from  $S$ , add the lines  $p$  (hypothesis),  $q$  (mp) to obtain a proof of  $q$  from  $S \cup \{p\}$ .

Backward: if we have proof  $t_1, \dots, t_n = q$  of  $q$  from  $S \cup \{p\}$ . We'll show that  $S \vdash (p \Rightarrow t_i) \forall i$ , so  $p \Rightarrow t_n = q$ .

If  $t_i$  is an axiom, then we have  $\vdash t_i \Rightarrow (p \Rightarrow t_i)$ , so  $\vdash p \Rightarrow t_i$ ;

If  $t_i \in S$ , write down  $t_i, t_i \Rightarrow (p \Rightarrow t_i), p \Rightarrow t_i$  we get a proof of  $p \Rightarrow t_i$  from  $S$ ;

If  $t_i = p$ : we know  $\vdash (p \Rightarrow p)$ , so done;

If  $t_i$  obtained by mp: in that case we have some earlier lines  $t_j$  and  $t_j \Rightarrow t_i$ .

By induction, we may assume  $S \vdash (p \Rightarrow t_j)$  and  $S \vdash (p \Rightarrow (t_j \Rightarrow t_i))$ .

Now we can write down  $[p \Rightarrow (t_j \Rightarrow t_i)] \Rightarrow [(p \Rightarrow t_j) \Rightarrow (t_i)]$  by axiom 2,  $p \Rightarrow (t_j \Rightarrow t_i), p \Rightarrow t_j \Rightarrow (p \Rightarrow t_i)$  (mp),  $p \Rightarrow t_j, p \Rightarrow t_i$  (mp) to obtain  $S \vdash (p \Rightarrow t_i)$ .

These are all of the cases. So  $S \vdash (p \Rightarrow q)$ . □

This is why we chose axiom 2 as we did – to make this proof work.

**Example.** To show  $\{p \Rightarrow q, q \Rightarrow r\} \vdash (p \Rightarrow r)$ , it's enough to show that  $\{p \Rightarrow q, q \Rightarrow r, p\} \vdash r$ , which is trivial by mp.

Now, how are  $\vdash$  and  $\models$  related? We are going to prove the *completeness theorem*:  $S \vdash t \iff S \models t$ .

This ensures that our proofs are sound, in the sense that everything it can prove is not absurd ( $S \vdash t$  then  $S \models t$ ), and are adequate, i.e. our axioms are powerful enough to define every semantic consequence of  $S$ , which is not obvious ( $S \models t$  then  $S \vdash t$ ).

**Proposition.** (3)

Let  $S \subset L, t \in L$ . Then  $S \vdash t \implies S \models t$ .

*Proof.* Given a valuation  $v$  with  $v(s) = 1 \forall s \in S$ , we want  $v(t) = 1$ .

We have  $v(p) = 1 \forall p$  axiom as our axioms are all tautologies (proven earlier);  $v(p) = 1 \forall p \in S$  by definition of  $v$ ; also if  $v(p) = 1$  and  $v(p \Rightarrow q) = 1$ , then also  $v(q) = 1$  (by definition of  $\Rightarrow$ ). So  $v(p) = 1$  for each line  $p$  of our proof of  $t$  from  $S$ . □

We say  $S \subset L$  consistent if  $S \not\vdash \perp$ . One special case of adequacy is:  $S \models \perp \implies S \vdash \perp$ , i.e. if  $S$  has no model then  $S$  inconsistent, i.e. if  $S$  is consistent then  $S$  has a model. This implies adequacy: given  $S \models t$ , we have  $S \cup \{\neg t\} \models \perp$ , so by our special case we have  $S \cup \{\neg t\} \vdash \perp$ , i.e.  $S \vdash ((\neg t) \Rightarrow t)$  by deduction theorem, so  $S \vdash \neg \neg t$ . But  $S \vdash ((\neg \neg t) \Rightarrow t)$  by axiom 3, so  $S \vdash t$  (mp).

**Theorem.** (4)

Let  $S \subset L$  be consistent, then  $S$  has a model.

The idea is that we would like to define valuation  $v$  by  $v(p) = 1 \iff p \in S$ , or more sensibly,  $v(p) = 1 \iff S \vdash p$ .

But maybe  $S \not\vdash p_3, S \not\vdash \neg p_3$ , but a valuation maps half of  $L$  to 1, so we want to 'grow'  $S$  to contain one of  $p$  or  $\neg p$  for each  $p \in L$ , while keeping consistency.

*Proof.* Claim: for any consistent  $S \subset L$ ,  $p \in L$ ,  $S \cup \{p\}$  or  $S \cup \{\neg p\}$  consistent.  
*Proof of claim.* If not, then  $S \cup \{p\} \vdash \perp$  and  $S \cup \{\neg p\} \vdash \perp$ , then  $S \vdash (p \implies \perp)$  (deduction theorem), i.e.  $S \vdash \neg p$ , so  $S \vdash \perp$  contradiction.

Now  $L$  is countable as each  $L_n$  is countable, so we can list  $L$  as  $t_1, t_2, \dots$ . Put  $S_0 = S$ ; set  $S_1 = S_0 \cup \{t_1\}$  or  $S_0 \cup \{\neg t_1\}$  so that  $S_1$  is consistent. Then set  $S_2 = S_1 \cup \{t_2\}$  or  $S_1 \cup \{\neg t_2\}$  so that  $S_2$  is consistent, and continue likewise. Set  $\bar{S} = S_0 \cup S_1 \cup S_2 \cup \dots$ . Then  $\bar{S} \supset S$ , and  $\bar{S}$  is consistent (as each  $S_n$  is, and each proof is finite).  $\forall p \in L$ , we have either  $p \in \bar{S}$  or  $(\neg p) \in \bar{S}$ . Also,  $\bar{S}$  is *deductively closed*, meaning that is  $\bar{S} \vdash p$  then  $p \in \bar{S}$ : if  $p \notin \bar{S}$  then  $(\neg p) \in \bar{S}$ , so  $\bar{S} \vdash p$ ,  $\bar{S} \vdash (\neg p)$  so  $\bar{S} \vdash \perp$  contradiction.

Define  $v : L \rightarrow \{0, 1\}$  by  $p \rightarrow 1$  if  $p \in \bar{S}$ , 0 otherwise. Then  $v$  is a valuation:  $v(\perp) = 0$  as  $\perp \notin \bar{S}$ ; for  $v(p \implies q)$ :

If  $v(p) = 1$ ,  $v(q) = 0$ : We have  $p \in \bar{S}$ ,  $q \notin \bar{S}$ , and want  $v(p \implies q) = 0$ , i.e.  $(p \implies q) \notin \bar{S}$ . But if  $(p \implies q) \in \bar{S}$  then  $\bar{S} \vdash q$  contradiction;

If  $v(q) = 1$ : have  $q \in \bar{S}$ , and want  $v(p \implies q) = 1$ , i.e.  $(p \implies q) \in \bar{S}$ . But  $\vdash q \implies (p \implies q)$  so  $\bar{S} \vdash (p \implies q)$ ;

If  $v(p) = 0$ : have  $p \notin \bar{S}$ , i.e.  $(\neg p) \in \bar{S}$  and want  $(p \implies q) \in \bar{S}$ . So we need  $(p \implies \perp) \vdash (p \implies q)$ , i.e.  $p \implies \perp, p \vdash q$  (deduction theorem). Thus it's enough to show that  $\perp \vdash q$ . But  $(\neg \neg q) \implies q$ , and  $\vdash (\perp \implies (\neg \neg q))$  (axiom 3 and 1 – to see the second one, write  $\neg$  explicitly using  $\implies$  and  $\perp$ ), so  $\vdash (\perp \implies q)$ , i.e.  $\perp \vdash q$ .  $\square$

**Remark.** Sometimes this is called 'completeness theorem'. The proof used  $P$  being countable to get  $L$  countable; in fact, result still holds if  $P$  is uncountable (see chapter 3).

By remark before theorem 4, we have

**Corollary.** (5, adequacy)

Let  $S \subset L$ ,  $t \in L$ . Then if  $S \models t$  then  $S \vdash t$ .

And hence,

**Theorem.** (6, completeness theorem)

Let  $S \subset L$ ,  $t \in L$ . Then  $S \vdash t \iff S \models t$ .

Some consequences:

**Corollary.** (7, compactness theorem)

Let  $S \subset L$ ,  $t \in L$  with  $S \models t$ . Then  $\exists$  finite  $S' \subset S$  with  $S' \models t$ .

This is trivial if we replace  $\models$  by  $\vdash$  (as proofs are finite).

Special case for  $t = \perp$ : If  $S$  has no model then some finite  $S' \subset S$  has no model. Equivalently,

**Corollary.** (7', compactness theorem, equivalent form)

Let  $S \subset L$ . If every finite subset of  $S$  has a model then  $S$  has a model.

This *isi* equivalent to corollary 7 because  $S \models t \iff S \cup \{\neg t\}$  has no model and  $S' \models t \iff S' \cup \{\neg t\}$  has no model.



**Corollary.** (8, decidability theorem)

There is an algorithm to determine (in finite time) whether or not, for a given finite  $S \subset L$  and  $t \in L$ , we have  $S \vdash t$ .

This is highly non-obvious; however it's trivial to decide if  $S \models t$  just by drawing a truth table, and  $\models \iff \vdash$ .

### 3 Well-Orderings and Ordinals

**Definition.** A *total order* or *linear order* on a set  $X$  is a relation  $<$  on  $X$ , such that

- (1) Irreflexive: Not  $x < x \forall x \in X$ ;
- (2) Transitive:  $x < y, y < z \implies x < z \forall x, y, z \in X$ ;
- (3) Trichotomous:  $x < y$  or  $x = y$  or  $y < x \forall x, y \in X$ .

Note: two of (iii) cannot hold: if  $x < y, y < x$  then  $x < x$  by transitivity.

Write  $x \leq y$  if  $x < y$  or  $x = y$ , and  $y > x$  if  $x < y$ .

We can also define total order in terms of  $\leq$ :

- (1) Reflexive:  $x \leq x \forall x \in X$ ;
- (2) Transitive:  $x \leq y, y \leq z \implies x \leq z \forall x, y, z \in X$ ;
- (3) Antisymmetric:  $x \leq y, y \leq x \implies x = y \forall x, y \in X$ ;
- (4) 'Tri'chotomous (although it's only two):  $x \leq y$  or  $y \leq x \forall x, y \in X$ .

**Example.**  $\mathbb{N}, \mathbb{Q}, \mathbb{R}$  with the usual orders are all total orders.

$\mathbb{N}^+$  the relation 'divides' is not a total order: for example we don't have any of  $2|3, 3|2$  or  $2 = 3$ .

$\mathcal{P}(S)$  for some  $S$  (with  $|S| \geq 2$  to be rigorous), with  $x \leq y$  if  $x \subseteq y$  is not a total order for the same reason.

A total order is a *well-ordering* if every (non-empty) subset has a least element, i.e.  $\forall S \subset X, S \neq \emptyset \implies \exists x \in S, x \leq y \forall y \in S$ .

**Example.** 1.  $\mathbb{N}$  with the usual  $<$  is a well ordering.

2.  $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$  with the usual  $<$  are not well orderings.

3.  $\mathbb{Q}^+ \cup \{0\}$  with the usual  $<$  is not a well ordering (e.g.  $(0, \infty) \subset \mathbb{Q}^+ \cup \{0\}$ ).

4. The set  $\{1 - \frac{1}{n} : n = 2, 3, \dots\}$  as a subset of  $\mathbb{R}$  with the usual ordering is a well ordering.

5. The set  $\{1 - \frac{1}{n} : n = 2, 3, \dots\} \cup \{1\}$  as a subset of  $\mathbb{R}$  with the usual ordering is a well ordering.

6. The set  $\{1 - \frac{1}{n} : n = 2, 3, \dots\} \cup \{2 - \frac{1}{n} : n = 2, 3, \dots\}$  (same assumption) is a well ordering.

**Remark.**  $X$  is well-ordered iff there is no  $x_1 > x_2 > x_3 > \dots$  in  $X$ .

Clearly if there is such a sequence then  $S = \{x_1, x_2, \dots\}$  has no least element.

Conversely, if  $S \subset X$  has no least element, then for each element  $x \in S$  there exists a  $x' \in S$  with  $x' < x$ , so we can just pick  $x, x', \dots$  inductively.

**Definition.** We say total orders  $X, Y$  are *isomorphic* if there exists a bijection  $f : X \rightarrow Y$  that is order-preserving, i.e.  $x < y \iff f(x) < f(y)$ .

For example, 1 and 4 above are isomorphic; 5 and 6 are isomorphic; 4 and 5 are not isomorphic (one has a greatest element, and the other doesn't).

Here comes the first reason why well orderings are useful:

**Proposition.** (1, Proof by induction)

Let  $X$  be well-ordered, and let  $S \subset X$  be s.t. if  $y \in S \forall y < x$  then  $x \in S$  (each  $x \in X$ ). Then  $S = X$ .

Equivalently, if  $p(x)$  is a property s.t.  $\forall x: \text{if } p(y) \forall y < x \text{ then } p(x)$ , then  $p(x) \forall x$ .

(I think we must assert  $S$  to be non-empty here, but the lecturer didn't agree with me; need to check later.)

*Proof.* If  $S \neq X$  then let  $x$  be the least element of  $X \setminus S$ . Then  $x \notin S$ . But  $y \in S \forall y < x$ , contradiction.  $\square$

A typical use:

**Proposition.** Let  $X, Y$  be isomorphic well-orderings. Then there is a *unique* isomorphism from  $X$  to  $Y$ .

*Proof.* Let  $f, g$  be isomorphisms. We'll show  $f(x) = g(x) \forall x$  by induction. Thus we may assume  $f(y) = g(y) \forall y < x$ , and want  $f(x) = g(x)$ . Let  $a$  be the least element of  $Y \setminus \{f(y) : y < x\}$ . Then we must have  $f(x) = a$ : if  $f(x) > a$ , then some  $x' > x$  has  $f(x') = a$  by surjectivity, contradiction. The same shows  $g(x)$  = least element of  $Y \setminus \{g(y) : y < x\}$ , but this is the same as  $a$ . So  $f(x) = g(x)$ .  $\square$

**Remark.** This is false for total orders in general. One example is, consider from  $\mathbb{Z} \rightarrow \mathbb{Z}$ , we could either take identity, or  $x \rightarrow x - 5$ ; or from  $\mathbb{R}$  to  $\mathbb{R}$  we could take identity or  $x \rightarrow x - 5$  or  $x \rightarrow x^3 \dots$

**Definition.** In a total order  $X$ , an *initial segment*  $I$  is a subset of  $X$  such that  $x \in I, y < x \implies y \in I$ .

**Example.** For any  $x \in X$ , set  $I(x) = \{y \in X : y < x\}$ . Then this is an initial segment.

Obviously, not every initial segment is of this form: for example, in  $\mathbb{R}$  we can take  $\{x : x \leq 3\}$ ; or in  $\mathbb{Q}$ , take  $\{x : x^2 < 2\} \cup \{x < 0\}$  (this cannot be written as above form as  $\sqrt{2} \notin \mathbb{Q}$ ).

Note: in a well-ordering, every proper initial segment *is* of the above form: let  $x$  be the least element of  $X \setminus I$ . Then  $y < x \implies y \in I$ . Conversely, if  $y \in I$ , then we must have  $y < x$ : otherwise  $x \in I$ , contradiction.

Our aim is to show that every subset of a well-ordered  $X$  is isomorphic to an initial segment.

Note: this is very false for total orders: e.g.  $\{1, 5, 9\} \subset \mathbb{Z}$ , or  $\mathbb{Q} \subset \mathbb{R}$ . If we have  $S \subset X$ , we would like to define  $f : S \rightarrow X$  that sends the smallest of  $S$  to the smallest of  $X$ , then remove them from both sets and send the smallest of the remaining to the smallest of the remaining, etc... But to do this we need a theorem.

**Theorem.** (3, definition by recursion)

Let  $X$  be well-ordered,  $Y$  be a set, and  $G : \mathcal{P}(X \times Y) \rightarrow Y$ . Then  $\exists f : X \rightarrow Y$  s.t.  $f(x) = G(f|_{I_x})$  for all  $x \in X$ . Moreover, such  $f$  is unique.

Here we define the restriction as: for  $f : A \rightarrow B$ , and  $C \subset A$ , the restriction of  $f$  to  $C$  is  $f|_C = \{(x, f(x)) : x \in C\}$ . (I think the lecturer is regarding a function as subset of a cartesian product)

In defining  $f(x)$ , make use of  $f|_{I_x}$ , i.e. the values of  $f(y), y < x$ .

*Proof.* Existence: define 'h is an attempt' to mean:  $h : I \rightarrow Y$ , some initial segment  $I$  of  $X$ , and  $\forall x \in I$  we have  $h(x) = G(h|_{I_x})$ . Note that  $h, h'$  are

attempts, both defined at  $x$ , then  $h(x) = h'(x)$  by induction on  $x$ . Since if  $h(y) = h'(y) \forall y < x$  then  $h(x) = h'(x)$ .

Also,  $\forall x \in X$  there exists an attempt defined at  $x$  by induction on  $x$ : we want attempt defined at  $x$ , given  $\forall y < x$  there exists attempt defined at  $y$ . For each  $y < x$ , we have unique attempt  $h_y$  defined on  $\{z : z \leq y\}$  (unique by what we just showed).

Let  $h = \cup_{y < x} h_y$ : an attempt defined on  $I_x$ . This is single-valued by uniqueness, so is indeed a function.

So  $h' = h \cup \{(x, G(h))\}$  is an attempt defined at  $x$ .

Now set  $f(x) = y$  if  $\exists$  attempt  $h$ , defined at  $x$ , with  $h(x) = y$  (single-valued).

Uniqueness: if  $f, f'$  suitable then  $f(x) = f'(x) \forall x \in X$  (induction on  $X$ ) – since if  $f(y) = f'(y) \forall y < x$  then  $f(x) = f'(x)$ .  $\square$

A typical application:

**Proposition.** (4, subset collapse)

Let  $X$  be well-ordered,  $Y \subset X$ . Then  $Y$  is isomorphic to an initial segment of  $X$ . Moreover, such initial segment is unique.

*Proof.* To have  $f$  an isomorphism from  $Y$  to an initial segment of  $X$ , we need precisely that  $\forall x \in Y : f(x) = \min X \setminus \{f(y) : y < x\}$ . So done (existence and uniqueness) by theorem 3.

Note that  $X \setminus \{f(y) : y < x\} \neq \emptyset$ , e.g. because  $f(y) \leq y \forall y$  (induction), so  $x \notin \{f(y) : y < x\}$ .  $\square$

In particular, a well-ordered  $X$  cannot be isomorphic to a proper initial segment of  $X$  – by uniqueness in subset collapse, as  $X$  is isomorphic to  $X$ .

How do different well-orderings relate to each other?

We say  $X \leq Y$  if  $X$  is isomorphic to an initial segment of  $Y$ . For example,  $\mathbb{N} \leq \{1 - \frac{1}{n} : n = 2, 3, \dots\} \cup \{1\}$ .

**Theorem.** (5)

Let  $X, Y$  be well-orderings. Then  $X \leq Y$  or  $Y \leq X$ .

*Proof.* Suppose  $Y \not\leq X$ . To obtain  $f : X \rightarrow Y$  that is an isomorphism with an initial segment of  $Y$ , need  $\forall x \in X : f(x) = \min Y \setminus \{f(y) : y < x\}$ . So we are done by theorem 3.

Note that we cannot have  $\{f(y) : y < x\} = X$ , as then  $Y$  is isomorphic to  $I_x$ .  $\square$

**Proposition.** (6)

Let  $X, Y$  be well-orderings with  $X \leq Y$  and  $Y \leq X$ . Then  $X$  and  $Y$  are isomorphic.

*Proof.* We have isomorphism  $f$  from  $X$  to an isomorphism of  $Y$ , and  $g$  the other way round. Then  $g \circ f : X \rightarrow X$  is an isomorphism from  $X$  to an initial segment of  $X$  (i.s. of i.s. is i.s.), but that is impossible unless the initial segment is  $X$

itself. So  $g \circ f$  is identity (by uniqueness in subset collapse). Similarly,  $f \circ g$  is identity on  $Y$ .  $\square$

New well-orderings from old:

Write  $X < Y$  if  $X \leq Y$  but  $X$  not isomorphic to  $Y$ . Equivalently,  $X < Y$  iff  $X$  is isomorphic to a proper initial segment of  $Y$ . For example, if  $X = \mathbb{N}$ ,  $Y = \{1 - \frac{1}{n}\} \cup \{1\}$  then  $X < Y$ .

Make a bigger one: given well-ordered  $X$ , choose  $x \notin X$ , and set  $x > y$  for all  $y \in X$ . This is a well-ordering on  $X \cup \{x\}$ : written  $X^+$ . Clearly  $X < X^+$ .

Put some together:

Let  $(X, <_X)$  and  $(Y, <_Y)$  be well-orderings. Say  $Y$  extends  $X$  if  $X \subset Y$ , and  $<_X, <_Y$  agree on  $X$ , and  $X$  an initial segment of  $(Y, <_Y)$ .

Well-orderings  $(X_i : i \in I)$  are nested if  $\forall i, j \in I : X_i$  extends  $X_j$  or  $X_j$  extends  $X_i$ .

**Proposition.** (7)

Let  $(X_i : i \in I)$  be a nested family of well-orderings. Then there exist well-ordering  $X$  with  $X \geq X_i \forall i$ .

*Proof.* Let  $X = \cup_{i \in I} X_i$ , with  $x < y$  if  $\exists i$  with  $x, y \in X_i$  and  $x <_i y$ . Then  $<$  is a well-defined total order on  $X$ . given  $S \subset X$ ,  $S \neq \emptyset$ , choose  $i$  with  $S \cap X_i \neq \emptyset$ . Then  $S \cap X_i$  has a minimal element (as  $X_i$  is well-ordered), which must also be a minimal element of  $S$  (as  $X_i$  an i.s. of  $X$ ). Also,  $X \geq X_i \forall i$ .  $\square$

## 4 Ordinals

Are the well-orderings themselves well-ordered?

An ordinal is a well-ordered set, with two well-ordered sets regarded as the same if they are isomorphic. (Just as a rational is an expression  $\frac{M}{N}$ , with  $\frac{M}{N}$ ,  $\frac{M'}{N'}$  regarded as the same if  $MN' = M'N$ . But, unlike for  $\mathbb{Q}$ , we cannot formalise by equivalence classes – see later).

If  $X$  is a well-ordering corresponding to ordinal  $\alpha$ , say  $X$  has order-type  $\alpha$ .

**Example.** For each  $k \in \mathbb{N}$ , write  $k$  for the order-type of the (unique) well-ordering of a set of size  $k$ , and write  $\omega$  for order-type of  $\mathbb{N}$ . So, in  $\mathbb{R}$ ,  $\{1, 3, 7\}$  has order-type 3.  $\{1 - \frac{1}{n} : n = 2, 3, \dots\}$  has order-type  $\omega$ . For  $X$  of o-t  $\alpha$  and  $Y$  of o-t  $\beta$ , write  $\alpha \leq \beta$  if  $X \leq Y$  (this is independent of choice of  $X, Y$ ). Similarly for  $\alpha < \beta$  etc.

We know:  $\forall \alpha, \beta, \alpha \leq \beta$  or  $\beta \leq \alpha$ , and if  $\alpha \leq \beta, \beta \leq \alpha$  then  $\alpha = \beta$ .

**Theorem.** Let  $\alpha$  be an ordinal. Then the ordinals  $< \alpha$  form a well-ordered set of order-type  $\alpha$ . e.g. the ordinals  $< \omega$  are  $0, 1, 2, 3, \dots$

*Proof.* Let  $X$  have o-t  $\alpha$ . the well-orderings  $< X$  are precisely (up to isomorphism) the proper initial segments of  $X$ , i.e. the  $I_x, x \in X$ .

But these are isomorphic to  $X$  itself, via  $x \rightarrow I_x$ . □

We often write  $I_\alpha$  to be the set of ordinals less than  $\alpha$ .

**Proposition.** (9)

Let  $S$  be a non-empty set of ordinals. Then  $S$  has a least element.

*Proof.* Choose  $\alpha \in S$ . If  $\alpha$  minimal in  $S$  then done. If not, then  $S \cap I_\alpha \neq \emptyset$ , so have a minimal element of  $S \cap I_\alpha$ , which is therefore minimal in  $S$ . □

**Theorem.** (10, Burali-Forti paradox):

The ordinals do not form a set.

*Proof.* Suppose not, let  $X$  be set of all ordinals. Then  $X$  is a well-ordering, say order-type  $\alpha$ . So  $X$  is isomorphic to  $I_\alpha$ . But  $I_\alpha$  is a proper i.s. of  $X$ . □

Given  $\alpha$ , we have  $\alpha^+ > \alpha$ . Also, if  $\{\alpha_i : i \in I\}$  is a set of ordinals, then there exists  $\alpha$  with  $\alpha \geq \alpha_i \forall i$  (by applying prop 7 to the nested family of  $I_{\alpha_i}; i \in I$ ).

In fact, there is therefore a least upper bound for  $\{\alpha_i : i \in I\}$  by applying prop 9 to the set  $\{\beta \leq \alpha : \beta \text{ an upper bound for the } \alpha_i\}$ . This is written  $\sup\{\alpha_i : i \in I\}$ , e.g.  $\sup\{2, 4, 6, 8, \dots\} = \omega$ .