Zorn's Lemma and Transfinite Induction

Yijia Chen

1. Zorn's Lemma and Well-Ordering

Let M be a set and nonempty $\mathcal{U} \subseteq \mathscr{P}ow(M) = \{T \mid T \subseteq M\}$. We say that a *nonempty* subset $C \subseteq \mathcal{U}$ is a *chain* in $\mathcal{U}$ if for every $T_1, T_2 \in C$ either $T_1 \subseteq T_2$ or $T_2 \subseteq T_1$.

Lemma 1.1 (Zorn's Lemma)**.** *Assume that for every chain* C *in* $\mathcal{U}$ *we have*

$$\bigcup C := \{a \mid a \in T \textit{ for some } T \in C\} \in \mathcal{U}.$$

Then $\mathcal{U}$ *has a* maximal element T*, i.e., there is no* $T' \in \mathcal{U}$ *with* $T \subsetneq T'$. ⊣

Definition 1.2. A *total order* on a set A is a binary relation $\leqslant \subseteq A \times A$ with the following properties. Let $a, b, c \in A$ be arbitrary.

(i) $a \leqslant a$ (i.e., $\leqslant$ is reflexive).

(ii) If $a \leqslant b$ and $b \leqslant a$, then $a = b$ (i.e., $\leqslant$ is anti-symmetric).

(iii) If $a \leqslant b$ and $b \leqslant c$, then $a \leqslant c$ (i.e., $\leqslant$ is transitive).

(iv) $a \leqslant b$ or $b \leqslant a$ (i.e., $\leqslant$ is total).

If furthermore

(v) every nonempty $A' \subseteq A$ has a *minimum* element a, i.e., $a \in A'$ and $a \leqslant a'$ for any $a' \in A'$,

then $\leqslant$ is a *well order*. ⊣

Assume that for every set A there is a well order $\leqslant \subseteq A \times A$. Towards a proof of Zorn's Lemma, let $\mathcal{U} \subseteq \mathscr{P}ow(M)$ be nonempty such that every chain C in $\mathcal{U}$ we have

$$\bigcup C \in \mathcal{U}.$$

We need to show $\mathcal{U}$ has a maximal element.

By assumption, there is a well order $\leqslant \subseteq \mathcal{U} \times \mathcal{U}$. As usual $u < u'$ means $u \leqslant u'$ and $u \neq u'$. Furthermore, let u_0 be the minimum element of $\mathcal{U}$ with respect to $\leqslant$.

As the core of the proof, we define a function $f : \mathcal{U} \to \mathcal{U}$ by the so-called *transfinite induction*.

(F1) $f(u_0) := u_0$.

(F2) Let $u \in \mathcal{U}$ with $u_0 < u$. We assume that $f(v)$ has been already defined for all $v < u$. Furthermore, for every $v_1 < v_2 < u$ we have

$$f(v_1) \subseteq f(v_2).$$

As a consequence,

$$\{f(v) \mid v < u\}$$

is a chain in $\mathcal{U}$. Then set

$$f(u) := \begin{cases} u & \text{if } f(v) \subseteq u \text{ for every } v < u, \\ \bigcup_{v<u} f(v) & \text{otherwise.} \end{cases} \tag{1}$$

Note that $\bigcup_{v<u} f(v)$ is defined, since we have assumed that $\{f(v) \mid v < u\}$ is a chain.

We argue that f is defined for every $u \in \mathcal{U}$. Otherwise, choose the minimum $u \in \mathcal{U}$ such that $f(u)$ is not defined. By (F1), we must have $u_0 < u$, otherwise $f(u) = f(u_0) = u_0$. Let $v_1 < v_2 < u$. Since u is the minimum element on which f is not defined, f is defined for every $v < u$ including v_1 and v_2. By (1)

$$f(v_2) := \begin{cases} v_2 & \text{if } f(v) \subseteq v_2 \text{ for every } v < v_2. \\ \bigcup_{v<v_2} f(v) & \text{otherwise.} \end{cases}$$

It is clear then that $f(v_1) \subseteq f(v_2)$. Therefore

$$\{f(v) \mid v < u\}$$

is a chain. Consequently (F2) implies that $f(u)$ is defined as well, which contradicts our assumption that f is not defined on u.

Next, similarly as argued above,

$$\{f(u) \mid u \in \mathcal{U}\}$$

is a chain. So we can set

$$w := \bigcup_{u \in \mathcal{U}} f(u).$$

We claim that w is maximal in $\mathcal{U}$. Otherwise, let $u \in \mathcal{U}$ such that

$$w \subsetneq u.$$

This implies that for any $u' < u$

$$f(u') \subseteq \bigcup_{v \in \mathcal{U}} f(v) = w \subsetneq u.$$

So by (1) we have $f(u) = u$, and

$$u = f(u) \subseteq \bigcup_{v \in \mathcal{U}} f(v) = w \subsetneq u,$$

which is a contradiction.

But what do we mean by defining the function f by transfinite induction as exhibited by (F1) and (F2).

Transfinite induction. In the following we show how to "define" the function f using usual set constructions. We begin with some definitions.

A binary relation $R \subseteq \mathcal{U} \times \mathcal{U}$ is *functional* if for every $u \in \mathcal{U}$ there is *at most one* $w \in \mathcal{U}$ with

$$(u, w) \in R.$$

Basically, it says that R can be viewed as a "partial function" which is defined on u with value w.

Furthermore, R is *good* if the following conditions are all satisfied.

(R1) R is functional.

(R2) $(u_0, u_0) \in R$.

(R3) If $(u, w) \in R$ and $(u', w') \in R$ with $u < u'$, then $w \leqslant w'$. That is, R is "monotone."

(R4) If $(u, w) \in R$, then for every $u' < u$ there is a $w' \in \mathcal{U}$ such that $(u', w') \in R$. That is, if R is "defined" on u, then it is "defined" on any $u' < u$ as well.

Together with (R3), it implies that

$$\{w' \mid (u', w') \in R \text{ for some } u' \leqslant u\}$$

is a chain in $\mathcal{U}$.

(R5) Assume $(u, w) \in R$ with $u_0 < u$. If for every $(u', w') \in R$ with $u' < u$ we have $w' \subseteq u$, then $w = u$.

Otherwise, i.e., for some $(u', w') \in R$ with $u' < u$ we have $w' \not\subseteq u$. Then

$$w = \bigcup \underbrace{\{w' \mid (u', w') \in R \text{ for some } u' < u\}}_{\text{a chain}}.$$

Claim 1. $\{(u_0, u_0)\}$ is good.

Proof of the claim. Trivial. ⊣

Claim 2. Let R_1 and R_2 be good. Furthermore, $(u, w_1) \in R_1$ and $(u, w_2) \in R_2$. Then $w_1 = w_2$. That is, for any $u \in \mathcal{U}$, any good R either is not "defined" on u, or maps u to a unique value independent of the choice of R.

Proof of the claim. Assume otherwise. We choose the minimum $u \in \mathcal{U}$ such that $(u, w_1) \in R_1$ and $(u, w_2) \in R_2$ with $w_1 \neq w_2$. Observe that $u \neq u_0$ by (R1) and (R2). And by (R1), (R4), and the minimality of u, for any $v < u$ we have a *unique* w_v with

$$(v, w_v) \in R_1 \quad \text{and} \quad (v, w_v) \in R_2 \tag{2}$$

We distinguish two cases. First, if for all $(v, w_v) \in R_1$ with $v < u$, i.e., $(v, w_v) \in R_2$ by (2), we have $w_v \subseteq u$, then (R4) implies that $w_1 = w_2 = u$. Otherwise, again by (R4),

$$\begin{aligned} w_1 &= \bigcup \{w_v \mid (v, w_v) \in R_1 \text{ for some } v < u\} \\ &= \bigcup \{w_v \mid (v, w_v) \in R_2 \text{ for some } v < u\} = w_2. \end{aligned}$$

This finishes the proof of the claim. ⊣

Let

$$R_f := \bigcup \{R \mid R \text{ is good}\}. \tag{3}$$

Claim 3. R_f is good.

Proof of the claim. By Claim 2, we see that R_f is functional. The other properties are also easy to verify. $\dashv$

Claim 4. For every $u \in \mathcal{U}$ there is a $w \in \mathcal{U}$ with $(u, w) \in R_f$.

Proof of the claim. Otherwise, let u be the minimum element such that there is no $w \in \mathcal{U}$ with $(u, w) \in R_f$. Since R_f is functional, for every $v < u$ there is a unique $w_v \in \mathcal{U}$ with

$$(v, w_v) \in R_f.$$

We define

$$C := \{w_v \mid w < u\}$$

and show that C is a chain. Let $v_1 < v_2 < u$. Thus $(v_1, w_{v_1}), (v_2, w_{v_2}) \in R_f$. By (3) there is a good R such that

$$(v_2, w_{v_2}) \in R.$$

By (R3) and (R4) there is a $w' \leqslant w_{v_2}$ with $(v_2, w') \in R \subseteq R_f$. Since R_f is functional, we conclude $w_{v_2} = w' \leqslant w_{v_1}$.

Now if $w_v \subseteq u$ for every $v < u$, we let

$$w_u := u.$$

Otherwise,

$$w_u := \bigcup C,$$

which is well defined, since C is a chain. It is then not hard to see that

$$R_f \cup \{(u, w_u)\}$$

is good too. This implies

$$R_f \cup \{u, w_u)\} \subseteq R_f$$

by our definition (3) of R_f, which is a contradiction. Therefore, R is "defined" on every $u \in \mathcal{U}$. $\dashv$

Now we are ready to define our function $f : \mathcal{U} \to \mathcal{U}$. For every $u \in \mathcal{U}$ we let $f(u)$ be the unique $w_u \in \mathcal{U}$ with $(u, w_u) \in R_f$.