A note on a proof of Zorn's Lemma from Axiom of Choice without transfinite induction

Koji Nuida

December 8, 2011

Abstract

In this note, we give a proof of Zorn's Lemma from Axiom of Choice without transfinite induction (whose essential idea is the same as the proof of Theorem 4.19 in the book [H. Rubin and J. E. Rubin, "Equivalents of the Axiom of Choice, II", Second Edition, Studies in Logic and the Foundations of Mathematics vol.116, North-Holland, 1985], as I noticed after writing the note ...). For the sake of comparison, we also include as an appendix a proof of the claim using transfinite induction.

Throughout this note, let $(X, \leq)$ denote an arbitrary non-empty partially ordered set in which every chain (linearly or totally ordered subset) C has an upper bound (with respect to $\leq$), that is, an element $x \in X$ satisfying that $c \leq x$ for every $c \in C$. Then Zorn's Lemma states that such an X always has a maximal element (with respect to $\leq$), that is, an element $x \in X$ satisfying that there exist no elements $y \in X$ with $x < y$. In this note, we give a proof of Zorn's Lemma from Axiom of Choice (in Zermelo–Fraenkel set theory), without transfinite induction which would be used in a "natural" proof of the claim.

Assume, for the contrary, that an X satisfying the hypothesis has no maximal elements. We will derive a contradiction from this assumption.

We summarize some basic definitions and terminologies:

- We say that a chain C in a subset Y of X is *bounded in* Y if there exists an element $x \in Y \setminus C$ satisfying that $c < x$ for every $c \in C$, and *unbounded in* Y otherwise.
- A chain C in X is called *well-ordered* if any non-empty subset of C has a unique minimal element.
- For a chain C and an element $c \in C$, the subset $\{d \in C \mid d < c\}$ is called the *initial segment* of C relative to c and is denoted by $s_C(c)$.

Let $\mathcal{C}$ denote the set of all non-empty chains in X. For each $C \in \mathcal{C}$, define $U_C = \{x \in X \setminus C \mid c < x \text{ for every } c \in C\}$. We notice the following property:

Lemma 1. *We have $U_C \neq \emptyset$ for every $C \in \mathcal{C}$.*

Proof. By the hypothesis of Zorn's Lemma, C has an upper bound $x \in X$. If $y \in U_{\{x\}}$, then we have $c \le x < y$ (hence $c < y$) for every $c \in C$, therefore $y \notin C$ and $y \in U_C$. Hence we have $U_{\{x\}} \subset U_C$, while $U_{\{x\}} \neq \emptyset$ by the assumption that X has no maximal elements, therefore we have $U_C \neq \emptyset$. $\square$

We would like to construct (by using Axiom of Choice) a choice function f of the family $\{U_C\}_{C \in \mathcal{C}}$ which satisfies some more appropriate properties described below.

For an arbitrary pair of $C_1, C_2 \in \mathcal{C}$, we define a symmetric relation $C_1 \sim_{\mathrm{pre}} C_2$ to mean that the chain $C_1 \cap C_2$ is unbounded in both C_1 and C_2. Then we define an equivalence relation $\sim$ on $\mathcal{C}$ to be the transitive closure of $\sim_{\mathrm{pre}}$, that is, we have $C_1 \sim C_2$ if and only if there exists a finite sequence $C_1 = C'_0, C'_1, \dots, C'_n = C_2$ $(n \ge 0)$ of elements of $\mathcal{C}$ with $C'_i \sim_{\mathrm{pre}} C'_{i+1}$ for every $0 \le i \le n-1$. Note that if $C_1 \sim C_2$ and C_1 has a maximum element c, then $c \in C_2$ and c is also a maximum element of C_2. Now we present the following lemma:

Lemma 2. *If $C_1, C_2 \in \mathcal{C}$ and $C_1 \sim C_2$, then $U_{C_1} = U_{C_2}$.*

Proof. It suffices to show that $U_{C_1} \subset U_{C_2}$ if $C_1 \sim_{\mathrm{pre}} C_2$. Let $x \in U_{C_1}$. If $x \in C_2$, then the upper bound x $(\in C_2 \setminus (C_1 \cap C_2))$ of C_1 is also an upper bound of $C_1 \cap C_2$, therefore $C_1 \cap C_2$ is bounded in C_2, contradicting the hypothesis $C_1 \sim_{\mathrm{pre}} C_2$. Hence $x \notin C_2$. If $c \in C_2$ and $c \not< x$, then, as $C_1 \cap C_2$ is unbounded in C_2, there exists a $d \in C_1 \cap C_2$ with $c \le d \not< x$. Now we have $d \in C_1$ and $d \not< x$, contradicting the property $x \in U_{C_1}$. Hence we have $c < x$ for every $c \in C_2$, therefore $x \in U_{C_2}$. $\square$

We define a subset $U_{\mathcal{E}}$ of X for each $\sim$-equivalence class $\mathcal{E}$ by $U_{\mathcal{E}} = U_C$, where $C \in \mathcal{E}$. This is well-defined by virtue of Lemma 2. Now we obtain (by using Axiom of Choice) a choice function f of the family $\{U_{\mathcal{E}}\}_{\mathcal{E} \in X/\sim}$. Moreover, for simplicity, we write $f(C) = f([C])$ for any $C \in \mathcal{C}$, where $[C]$ denotes the $\sim$-equivalence class of C. Hence $f(C) \in U_C$ for any $C \in \mathcal{C}$. Moreover, if $C_1, C_2 \in \mathcal{C}$ and $C_1 \sim C_2$, then we have $f(C_1) = f(C_2)$.

From now, we fix an element $x_0 \in X$. We give the following definition:

- We say that a $C \in \mathcal{C}$ is *f-consecutive* if C is well-ordered, x_0 is the minimum element of C, and we have $f(s_C(c)) = c$ for any $c \in C \setminus \{x_0\}$.

We present some properties of f-consecutive chains:

Lemma 3. *Let $C \in \mathcal{C}$ be f-consecutive, and suppose that $\emptyset \neq C' \subset C$ and C' is bounded in C. Then we have $f(C') = d$, where d is the unique minimal element of the set of all upper bounds $x \in C \setminus C'$ of C'.*

Proof. Note that $C' \in \mathcal{C}$. Let d be as specified in the statement, which indeed exists by virtue of the fact that C is well-ordered and C' is bounded in C. We show that $s_C(d) \sim C'$. First, we have $C' \subset s_C(d)$, therefore $s_C(d) \cap C' = C'$ is unbounded in C'. On the other hand, if $x \in s_C(d) \setminus C'$, then x is not an upper bound of C' by the choice of d. Therefore C' is unbounded in $s_C(d)$. Hence we have $s_C(d) \sim C'$. Now we have $f(C') = f(s_C(d))$ by the property of f, while we have $f(s_C(d)) = d$, as C is f-consecutive. This implies that $f(C') = d$. $\square$

Lemma 4. *Let $C_1, C_2 \in \mathcal{C}$ be two f-consecutive chains. Then exactly one of the following three conditions is satisfied:*

1. *$C_1 = C_2$;*
2. *C_1 is an initial segment of C_2;*
3. *C_2 is an initial segment of C_1.*

Proof. It is obvious that two of the three conditions specified in the statement do not occur simultaneously.

First we show that for each $i \in \{1, 2\}$, the conditions $x \in C_1 \cap C_2$ and $y \in s_{C_i}(x)$ imply $y \in C_1 \cap C_2$. Assume, for the contrary, that a counterexample y exists for some x, and let y be the minimal counterexample relative to the fixed x (which exists, as C_i is well-ordered). Then we have $s_{C_i}(y) \neq \emptyset$, as $x_0 \in C_1 \cap C_2$. By the choice of y, we have $s_{C_i}(y) \subset C_{3-i}$ and $y \notin C_{3-i}$. We have $f(s_{C_i}(y)) = y$, as C_i is f-consecutive. On the other hand, as $x \in C_1 \cap C_2$ and $y < x$, the subset $s_{C_i}(y)$ of C_{3-i} is bounded in C_{3-i}, therefore we have $f(s_{C_i}(y)) \in C_{3-i}$ by Lemma 3. But this contradicts the property $y \notin C_{3-i}$. Hence the claim of this paragraph holds. Moreover, as C_1 and C_2 are well-ordered, the above fact implies that $C_1 \cap C_2$ is either equal to or an initial segment of C_i for each $i \in \{1, 2\}$ (if $C_1 \cap C_2 \neq C_i$, consider the minimum element of $C_i \setminus (C_1 \cap C_2)$).

Now assume, for the contrary, that $C_1 \cap C_2 \neq C_1$ and $C_1 \cap C_2 \neq C_2$. Then $C_1 \cap C_2$ is an initial segment of both C_1 and C_2, therefore we have $f(C_1 \cap C_2) \in C_1 \cap C_2$, as C_1 and C_2 are f-consecutive. But this contradicts the definition of f. Hence we have $C_1 \cap C_2 = C_i$ for some $i \in \{1, 2\}$, and now the above result implies that $C_i = C_{3-i}$ or C_i is an initial segment of C_{3-i}. $\square$

Let C_0 denote the union of all f-consecutive chains (note that at least one f-consecutive chain, say $\{x_0\}$, exists). Then Lemma 4 implies that $C_0 \in \mathcal{C}$ (for two elements $x, y \in C_0$, there exist f-consecutive chains C containing x and C' containing y, and $x \in C \subset C'$ or $y \in C' \subset C$ by the lemma). Now we present the following lemma:

Lemma 5. *Let $c \in C_0$, and let C be an f-consecutive chain containing c. Then $s_{C_0}(c) = s_C(c)$.*

Proof. First note that $s_C(c) \subset s_{C_0}(c)$, as $C \subset C_0$ by the definition of C_0. We show that $d \in C$ for an arbitrary element $d \in s_{C_0}(c)$. By the definition of C_0, there exists an f-consecutive chain C' such that $d \in C'$. Now we have $d \in C' \subset C$ if $C' = C$ or C' is an initial segment of C. On the other hand, if C is an initial segment of C', then the properties $c \in C$, $d \in C'$ and $d < c$ imply $d \in C$. Hence we have $d \in C$ by Lemma 4. $\square$

From now, we will show that C_0 itself is f-consecutive. It is obvious that x_0 is the minimum element of C_0.

Lemma 6. *C_0 is well-ordered.*

Proof. Let A be an arbitrary non-empty subset of C_0. Choose an element $a \in A$, and let C be an f-consecutive chain with $a \in C$ (which exists by the definition of C_0). As C is well-ordered, $A \cap C$ has the minimum element, say a_0. Now if $a' \in A$ and $a' < a_0$, then we have $a' \in s_{C_0}(a_0) = s_C(a_0)$ by Lemma 5, therefore $a' \in A \cap C$ and $a' < a_0$, contradicting the choice of a_0. Hence we have $a' \geq a_0$ for every $a' \in A$, therefore A itself has the minimum element a_0. $\square$

Lemma 7. *We have $f(s_{C_0}(c)) = c$ for any $c \in C_0 \setminus \{x_0\}$.*

Proof. Let C be an f-consecutive chain with $c \in C$ (which exists by the definition of C_0). Then we have $s_{C_0}(c) = s_C(c)$ by Lemma 5. As C is f-consecutive, now we have $f(s_{C_0}(c)) = f(s_C(c)) = c$. $\square$

By Lemma 6 and Lemma 7, it follows that C_0 is f-consecutive. Now it is straightforward to verify that $C_0 \cup \{f(C_0)\}$ is also an f-consecutive chain, but it is not a subset of C_0, contradicting the definition of C_0. Hence X should have a maximal element, concluding the proof of Zorn's Lemma.

Appendix: A proof using transfinite induction

In this appendix, for the sake of comparison, we describe a proof of Zorn's Lemma from Axiom of Choice using transfinite induction. First we clarify the statement of the principle for "definition by transfinite recursion" (see e.g., Theorem 9.3 in Chapter I of [K. Kunen, "SET THEORY, An Introduction to Independence Proofs", Elsevier, 1980]):

Theorem 1. *Let $\varphi(x, y)$ be a formula (in Zermelo–Fraenkel set theory) with free variables x, y satisfying $\forall x \exists ! y \varphi(x, y)$. Then there exists a formula $\Phi(x, y)$ with free variables x, y satisfying the following two conditions;*

1. $\forall x((x \in \mathbf{ON} \to \exists ! y \Phi(x, y)) \wedge (\neg x \in \mathbf{ON} \to \neg \exists y \varphi(x, y)))$;
2. $\forall x(x \in \mathbf{ON} \to \forall y, z(y = \Phi \restriction_x \wedge \varphi(y, z) \to \Phi(x, z)))$,

where "$x \in \mathbf{ON}$" is an abbreviation of "x is an ordinal number" and "$\Phi \restriction_x$" is an abbreviation of the set $\{\langle a, b \rangle \mid a \in x \wedge \Phi(a, b)\}$ (with $\langle a, b \rangle$ denoting the ordered pair of a and b).

Intuitively, the theorem means that, if we would like to define a "function" Φ with domain consisting of all ordinal numbers (the whole of which is never a set) in such a way that the value of Φ at each ordinal number α is determined by a given rule from the values of Φ at ordinal numbers less than α, then there indeed exists such a "function" Φ. Note that this is a theorem of ZF set theory and does not depend on Axiom of Choice.

Now we give a proof of Zorn's Lemma from Axiom of Choice using Theorem 1 (as well as transfinite induction). Let $X \neq 0$ be a partially ordered set appeared in the statement of Zorn's Lemma. Assume, for the contrary, that X has no maximal elements. Then, for each non-empty subset C of X which is isomorphic to an ordinal number (hence a chain), it follows from Axiom of Choice that there exists a distinguished upper bound b_C of C with $b_C \in X \setminus C$.

To apply Theorem 1, first we define a formula $\varphi(x, y)$ in the following manner, where we fix an element $a \in X$ throughout the proof:

- If $x = 0$ $(= \emptyset)$, then let $\varphi(x, y)$ mean that $y = a$.
- If x is a function from an ordinal number $\alpha > 0$ to X which is an isomorphism (between partially ordered sets) onto the image $\mathrm{Im}(x)$ of x, then let $\varphi(x, y)$ mean that $y = b_{\mathrm{Im}(x)}$ (note that $\mathrm{Im}(x)$ is isomorphic to the non-empty ordinal number α, therefore $b_{\mathrm{Im}(x)}$ is indeed defined).
- Otherwise, let $\varphi(x, y)$ mean that $y = 0$.

This formula $\varphi(x, y)$ satisfies the hypothesis of Theorem 1, therefore a formula $\Phi(x, y)$ as in the theorem exists. Now we have the following lemma:

Lemma 8. *Let x be an ordinal number, and let x' be the unique element satisfying $\Phi(x, x')$.*

1. *We have $x' \in X$.*
2. *If $y < x$ and $\Phi(y, y')$, then $y' < x'$ in X.*

Proof. We prove the claim by transfinite induction on x. First, if $x = 0$, then it follows from the definition of the formula φ that $x' = a$, therefore the specified conditions are satisfied. Secondly, suppose that $x > 0$. Then, by the hypothesis of the transfinite induction, the set $\Phi \restriction_x$ in the statement of Theorem 1 is an isomorphism from x to a subset of X, say, C (note that x is totally ordered). Now by the definitions of Φ and φ, it follows that $x' = b_C$, therefore the specified conditions are satisfied for x (the second condition follows from the property that $b_C \in X \setminus C$ is an upper bound of C). Hence the claim holds. $\square$

By the second property shown in Lemma 8, for each element $v \in X$, there exists at most one ordinal number x satisfying $\Phi(x, v)$. Let X' denote the subset of X defined in such a way that $v \in X'$ if and only if $v \in X$ and $\Phi(x, v)$ for some (or equivalently, a unique) ordinal number x. By the Axiom Schema of Replacement applied to the set X' and the formula $\Phi'(x, y) := \Phi(y, x)$, there exists a set Y for which we have $y \in Y$ if y is an ordinal number and the unique element y' satisfying $\Phi(y, y')$ belongs to X'. Now by the first property shown in Lemma 8, the set Y contains every ordinal number. However, this contradicts Burali–Forti Paradox (which states that there exist no sets containing all ordinal numbers). Hence X should have a maximal element, concluding the proof of Zorn's Lemma.