Zorn's Lemma

Theorem [Zorn's Lemma]

Let R be a partial order on a non-empty set A.
If every chain in A has an upper bound in A, then A has a maximal element.

Proof

Notice that the set $\mathcal{C}$ of all chains in A comes with its own natural partial order: it is partially ordered by the inclusion relation "$\subseteq$". Below, we will show that there is a chain C_0 in A, that is maximal with respect to this inclusion relation. Since by assumption every chain in A has an upper bound in A, C_0 must have an upper bound $a_0 \in A$. This element a_0 is a maximal element of A: for if $a \in A$ with $a_0 R a$, then for all $c \in C_0$, cRa_0 and $a_0 R a$. Hence, cRa for all $c \in C_0$, so that $C_0 \cup \{a\}$ is a chain in A. But C_0 is not a subset of any chain in A other than itself, so that $a \in C_0$. Since a_0 is an upper bound for C_0, then aRa_0. Therefore $a = a_0$ (since $a_0 R a$ and aRa_0), proving that a_0 is a maximal element of A. We see that all we need to do in order to establish the theorem is find a maximal chain C_0 in A.

We will find a maximal chain in A in two steps:

(a) We will construct a function g which extends every chain C by exactly one element to a larger chain whenever possible. That is, we will find a function $g : \mathcal{C} \to \mathcal{C}$ such that $g(C) = C \cup \{a\}$ if there is an element $a \in A \setminus C$ such that $C \cup \{a\}$ is chain, or else $g(C) = C$.

(b) We will then show that there is a chain C_0 in A such that $g(C_0) = C_0$, i.e. a chain that cannot be extended; a maximal chain.

(a) Before we define the function g, we define a function f which picks out one element from every non-empty subset S of A. Such a function is called a *choice function* for the set A. Consider the set X of all elements in $\mathcal{P}(\mathcal{P}(A) \times A)$ that are of the form $\{S\} \times S$ with $\emptyset \neq S \subseteq A$. Then the elements of X are all non-empty and pair-wise disjoint: if $(S, s) \in (\{S_1\} \times S_1) \cap (\{S_2\} \times S_2)$, then $S_1 = S = S_2$ so that $\{S_1\} \times S_1 = \{S_2\} \times S_2$. By the *axiom of choice*, there is a set f which contains exactly *one* element (S, s) from every set $\{S\} \times S$ of X. This is the same as saying that f passes the "vertical line test". Therefore f is a function. Denoting $(S, s) \in f$ by $f(S) = s$ as usual, we see that $f(S) \in S$ for all non-empty subsets S of A, as desired.

We now use our choice function f to define g. Given a chain $C \in \mathcal{C}$, consider the set C^* of all elements in $A \setminus C$ by which C could possibly be extended to a larger chain. In symbols, $C^* = \{a \in A \mid a \notin C \wedge C \cup \{a\} \in \mathcal{C}\}$. If C^* is empty, i.e. if C cannot be extended, we define $g(C) = C$. If C^* is not empty, we can pick an element $f(C^*) \in C^*$ using our choice function f and define $g(C) = C \cup \{f(C^*)\}$; an extension of the chain C by *one* element to a larger chain.

(b) Since the set $\mathcal{C}$ of all chains in A is itself partially ordered by inclusion, it is meaningful to speak of a chain of chains: a subset $\mathcal{D} \subseteq \mathcal{C}$ is a chain of chains if for all chains $C_1, C_2 \in \mathcal{D}$ either $C_1 \subseteq C_2$ or $C_2 \subseteq C_1$.

Fix an element $x \in A$. Then $\{x\}$ is a chain in A, albeit a very short one. We will call a set $\mathcal{T}$ of chains in A a *tower*, if it satisfies the following three properties:

(i) $\{x\} \in \mathcal{T}$;

(ii) If $\mathcal{D} \subseteq \mathcal{T}$ is a chain of chains, then $\bigcup \mathcal{D} \in \mathcal{T}$;
(Recall that $\bigcup \mathcal{D}$ denotes the union of all the elements C in $\mathcal{D}$.)

(iii) If $C \in \mathcal{T}$, then $g(C) \in \mathcal{T}$.

An example of a tower is the set $\mathcal{T}_1$ of all chains in A that contain x. It clearly satisfies properties (i) and (iii). It also has property (ii): suppose $\mathcal{D} \subseteq \mathcal{T}_1$ is a chain of chains. If $a, b \in \bigcup \mathcal{D}$, then $a \in C_1$ and $b \in C_2$ for some $C_1, C_2 \in \mathcal{D}$. Since $\mathcal{D}$ is a chain of chains, either $C_1 \subseteq C_2$ or $C_2 \subseteq C_1$. Say, $C_1 \subseteq C_2$. Then a and b are both in the same chain C_2. So, either aRb or bRa. Therefore $\bigcup \mathcal{D}$ is indeed a chain itself. Moreover, since every element of $\mathcal{D}$ contains x so does $\bigcup \mathcal{D}$. Consequently, $\bigcup \mathcal{D} \in \mathcal{T}_1$.

Let $\mathcal{T}_0$ be the intersection of all towers. Then $\mathcal{T}_0$ is again a tower. In fact, $\mathcal{T}_0$ is the *smallest* tower, since it is a subset of every tower. In particular, every element C of $\mathcal{T}_0$ is an element of $\mathcal{T}_1$ and therefore contains x. It turns out that the tower $\mathcal{T}_0$ of chains is actually a *chain* of chains. We will verify this claim below. For the moment, let us assume this as true.

Put $C_0 = \bigcup \mathcal{T}_0$. Since $\mathcal{T}_0$ is itself a chain of chains, then $C_0 = \bigcup \mathcal{T}_0 \in \mathcal{T}_0$ by Property (ii). Consequently, $g(C_0) \in \mathcal{T}_0$ by Property (iii). So, we have $g(C_0) \subseteq C_0$, since C_0 is the union of *all* the elements of $\mathcal{T}_0$, and we have $C_0 \subseteq g(C_0)$ by definition of the function g. This shows that $g(C_0) = C_0$ and completes Step (b).

All that is left to prove is that $\mathcal{T}_0$ is a chain of chains. That is, we need to verify that for all $C_1, C_2 \in \mathcal{T}_0$ either $C_1 \subseteq C_2$ or $C_2 \subseteq C_1$. To this end, we define $\mathcal{E} = \{C \in \mathcal{T}_0 \mid \forall D \in \mathcal{T}_0 \, [C \subseteq D \vee D \subseteq C]\}$. Then our goal is to show that $\mathcal{E} = \mathcal{T}_0$. We do this by showing that $\mathcal{E}$ is a tower. Since $\mathcal{E} \subseteq \mathcal{T}_0$ and $\mathcal{T}_0$ is the smallest tower, this will imply that $\mathcal{E} = \mathcal{T}_0$. Property (i) is not an issue, because every element D of $\mathcal{T}_0$ contains x, i.e. $\{x\} \subseteq D$ for all $D \in \mathcal{T}_0$, so that $\{x\} \in \mathcal{E}$. To show Property (ii), let $\mathcal{D} \subseteq \mathcal{E}$ be a chain of chains. Then $\bigcup \mathcal{D} \in \mathcal{T}_0$, since $\mathcal{T}_0$ is a tower. Now let $D \in \mathcal{T}_0$. Since every element of $\mathcal{D}$ is an element of $\mathcal{E}$, we know that for each $C \in \mathcal{D}$ either $C \subseteq D$ or $D \subseteq C$. If $C \subseteq D$ for *all* $C \in \mathcal{D}$, then $\bigcup \mathcal{D} \subseteq D$. If, on the other hand, $D \subseteq C$ for at least *one* $C \in \mathcal{D}$, then $D \subseteq C \subseteq \bigcup \mathcal{D}$. Hence, $\bigcup \mathcal{D} \in \mathcal{E}$. To show Property (iii), let $C \in \mathcal{E}$. Consider the set $\mathcal{F} = \{D \in \mathcal{T}_0 \mid g(C) \subseteq D \vee D \subseteq C\}$. The final paragraph of our proof will show that $\mathcal{F}$ is a tower. Consequently, $\mathcal{F} = \mathcal{T}_0$, because $\mathcal{F} \subseteq \mathcal{T}_0$ and $\mathcal{T}_0$ is the smallest tower. So, if $D \in \mathcal{T}_0 = \mathcal{F}$, then either $g(C) \subseteq D$ or $D \subseteq C \subseteq g(C)$. This places $g(C)$ into $\mathcal{E}$ and proves that $\mathcal{E}$ is a tower.

The remainder of the proof consists of checking that for a fixed $C \in \mathcal{E}$ the set $\mathcal{F} = \{D \in \mathcal{T}_0 \mid g(C) \subseteq D \vee D \subseteq C\}$ is a tower. Properties (i) and (ii) are verified as for the set $\mathcal{E}$. To verify Property (iii) we let $D \in \mathcal{F}$. Then either $g(C) \subseteq D$ or $D \subseteq C$. Moreover, since $C \in \mathcal{E}$, either $C \subseteq g(D)$ or $g(D) \subseteq C$. Consequently, either $g(C) \subseteq g(D)$ or $g(D) \subseteq C$, so that $g(D) \in \mathcal{F}$. (The only case that needs to be looked at is where $D \subseteq C \subseteq g(D)$. However, $g(D)$ has at most one more element than D. This makes either the first or the second inclusion an equality.)