Equivalence of the Axiom of Choice, the Well-ordering Theorem, and Zorn’s Lemma

H. Krieger, Mathematics 331, Harvey Mudd College
Fall, 2003

1 Definitions

Definition 1 A set $S$ is partially ordered (by the relation $\leq$) $\iff$

1. $x \leq x$

2. $x \leq y$ and $y \leq x \implies x = y$

3. $x \leq y$ and $y \leq z \implies x \leq z$

Definition 2 Let $S$ be partially ordered and let $B \subset S$.

1. $x$ is an upper bound for $B \iff x \in S$ and $y \leq x$ for all $y \in B$.

2. $x$ is the least upper bound (lub or sup) for $B \iff x$ is an upper bound for $B$ and whenever $z$ is an upper bound for $B$ then $x \leq z$.

3. $x$ is the last element (greatest element) of $B \iff x \in B$ and $x$ is an upper bound for $B$.

4. $x$ is a maximal element of $B \iff x \in B$ and whenever $y \in B$, $y \geq x$, then $y = x$.

There are corresponding definitions for lower bound, greatest lower bound (glb or inf), first element (smallest element), and minimal element.

Definition 3 A set $S$ is linearly (totally) ordered (by the relation $\leq$) $\iff S$ is partially ordered (by the relation $\leq$) and $x \leq y$ or $y \leq x$ for all $x \in S$, $y \in S$ $\iff$

1. $x \leq y$ or $y \leq x$

2. $x \leq y$ and $y \leq x \implies x = y$

3. $x \leq y$ and $y \leq z \implies x \leq z$

A linearly ordered set is also called a chain.
Definition 4 A set $S$ is well ordered (by the relation $\leq$) $\iff$ $S$ is partially ordered (by the relation $\leq$) and every non-empty subset of $S$ has a first element $\iff$ $S$ is linearly ordered (by the relation $\leq$) and every non-empty subset of $S$ has a first element.

Definition 5 A set $S$ is inductively ordered (by the relation $\leq$) $\iff$ $S$ is partially ordered (by the relation $\leq$) and every linearly ordered subset (chain) in $S$ has a least upper bound.

2 Theorem Statements

W.O.T. Every set can be well-ordered.

A.C.1. If $S$ is a non-empty collection of disjoint non-empty sets $S$, then there is a set $R$ which has as its elements exactly one element $x$ from each set $S$ in $S$.

A.C.2. If $S$ is a non-empty collection of non-empty sets $A$, then there is a function $\varphi : S \rightarrow \bigcup\{A : A \in S\}$ such that $\varphi(A) \in A$ for all $A \in S$.

A.C.3. If $I$ is a non-empty set and, for each $i \in I$, $S_i$ is a non-empty set, then the cartesian product

$$\prod_{i \in I} S_i = \{f : I \rightarrow \bigcup\{S_i : i \in I\} : f(i) \in S_i \text{ for all } i \in I\}$$

is non-empty.

Z.L.1. If $S$ is a partially ordered set such that each chain in $S$ has an upper bound, then $S$ has a maximal element.

Z.L.2. If $S$ is inductively ordered, then $S$ has a maximal element.

3 Equivalence of these Statements

A.C.3. $\implies$ A.C.2. Since the cartesian product $\prod\{A : A \in S\}$ is non-empty, there is a function $\varphi : S \rightarrow \bigcup\{A : A \in S\}$ such that $\varphi(A) \in A$ for all $A \in S$.

A.C.2. $\implies$ A.C.1. Since there is a function $\varphi : S \rightarrow \bigcup\{S : S \in S\}$ such that $\varphi(S) \in S$ for all $S \in S$ and the sets in in $S$ are disjoint, the range $R$ of $\varphi$ is a set which has for its elements exactly one element from each set $S$ in $S$.

A.C.1. $\implies$ A.C.3. For each $i \in I$, the set $T_i = \{(i, x) : x \in S_i\}$ is non-empty. Thus the collection $S = \{T_i : i \in I\}$ is a non-empty collection of disjoint non-empty sets. Hence there is a set $f$ which has as its elements exactly one element $(i, f(i))$ from each set $T_i$. Then $f : I \rightarrow \bigcup\{S_i : i \in I\}$ with $f(i) \in S_i$ for all $i \in I$. 2
Z.L.1. $\implies$ Z.L.2. If $S$ is inductively ordered, then every chain in $S$ has an upper bound. Thus $S$ has a maximal element.

Z.L.2. $\implies$ Z.L.1. Suppose $S$ is partially ordered and every chain in $S$ has an upper bound. Let $T = \{C : C$ is a chain in $S\}$. Since $T \subset 2^S$ and $2^S$ is partially ordered by inclusion, so is $T$. If $C_0$ is a maximal element of $T$, then there is an upper bound $x_0$ for $C_0$. Then $x_0 \in C_0$ and $x_0$ is a maximal element of $S$. Hence our problem is reduced to showing that $T$ has a maximal element. But if $C$ is a chain in $T$ and $C_0 = \cup\{C : C \in C\}$, then $C_0$ is the least upper bound of $C$. Hence $T$ is inductively ordered which means $T$ does have a maximal element.

W.O.T. $\implies$ A.C.2. Let $S$ be a non-empty collection of non-empty sets $A$.

Let $S = \cup\{A : A \in S\}$ and, then, well-order $S$. Hence each $A$ is then a non-empty subset of the well-ordered set $S$. If $\varphi(A)$ is the first element of $A$, then $\varphi : S \rightarrow \cup\{A : A \in S\}$ such that $\varphi(A) \in A$ for all $A \in S$.

Z.L.2. $\implies$ W.O.T. Let $S$ be a set. For each subset $A$ of $S$ which can be well-ordered, let $W_A$ be the collection of well-orderings of $A$ and then let

$$W = \{(A, \leq_A) : A \subset S, \leq_A \in W_A\}.$$  

Define a partial ordering on $W$ by letting $(A, \leq_A) \leq (B, \leq_B)$

$$\iff A \subset B, \leq_A \leq_B |_A, \text{ and } a \in A, b \in B - A \implies a \leq_B b.$$  

Now if $C$ is a chain in $W$, $C = \{(C, \leq_C)\}$, let $C_0 = \cup\{C : (C, \leq_C) \in C\}$ and let $\leq_{C_0} = \cup\{\leq_C : (C, \leq_C) \in C\}$, i.e. if $x$ and $y$ are in $C_0$, then $x \leq_{C_0} y \iff x \leq_C y$ for some $C$ such that $(C, \leq_C) \in C$. Then $\leq_{C_0}$ is a well-ordering of $C_0$ and $(C_0, \leq_{C_0})$ is the least upper bound of $C$. Hence, $W$ is inductively ordered which means that $W$ has a maximal element. Let $(A, \leq_A)$ be a maximal element of $W$. Then $A = S$, for if $S - A \neq \phi$, we consider $B = A \cup \{x\}$, where $x \in S - A$, and define $\leq_B$ by $y \leq_B x$ for all $y \in A, \leq_B |_A = \leq_A$. However, in this case $(A, \leq_A) \leq (B, \leq_B)$ and $(A, \leq_A) \neq (B, \leq_B)$. Hence, $(S, \leq_S) \in W$ for some well-ordering $\leq_S$, i.e. $S$ can be well-ordered.

**Lemma:** Let $S$ be inductively ordered and let $f : S \rightarrow S$ such that $f(x) \geq x$ for every $x \in S$. Then there is an $x_0 \in S$ such that $f(x_0) = x_0$. In fact, if $a \in S$, there is an $x_0 \geq a$ such that $f(x_0) = x_0$.

**Proof:** Let $a \in S$ and let $B_a$ be the collection of all subsets $B$ of $S$ such that

i) $a \in B$,  
ii) $f(B) \subset B$,  
iii) if $C$ is a chain in $B$, then $\text{lub} \ C \in B$. 

3
Note that $B_a$ is not empty, since \( \{x : x \geq a\} \in B_a \). Let \( A = \cap \{B : B \in B_a\} \). Then it is easy to see that \( A \in B_a \). Our objective is to show that \( A \) is a chain, for then if \( x_0 = \operatorname{lub} A \), it follows that \( x_0 \in A \implies f(x_0) \leq x_0 \leq f(x_0) \), i.e. \( f(x_0) = x_0 \). The intuitive idea is that \( A \) consists only of \( a \) and its successive images under \( f \). Hence, we let
\[
P = \{p \in A : y \in A \text{ and } y < p \implies f(y) \leq p\}.
\]
Note that \( P \subset A \) but we would like to have equality. As a start, we show that if \( p \in P \) and \( z \in A \), then either \( z \leq p \) or \( z \geq f(p) \) (i.e. elements of \( P \) and \( A \) are comparable and there are no elements of \( A \) between \( p \) and \( f(p) \) if \( p \in P \)). To see this, let \( p \in P \) and let
\[
B(p) = \{z \in A : \text{ either } z \leq p \text{ or } z \geq f(p)\}.
\]
Then,
i) \( a \leq p \implies a \in B(p) \).

ii) If \( z \in A \) and \( z < p \), then \( f(z) \leq p \); if \( z \in A \) and \( z = p \), then \( f(z) = f(p) \geq f(p) \); if \( z \in A \) and \( z \geq f(p) \), then \( f(z) \geq z \geq f(p) \). Hence, \( z \in B(p) \implies f(z) \in B(p) \), i.e., \( \{B(p)\} \subset B(p) \).

iii) Let \( C \) be a chain in \( B(p) \) and let \( c_0 = \operatorname{lub} C \). Then \( c_0 \in A \). If \( c \leq p \) for all \( c \in C \), then \( c_0 \leq p \implies c_0 \in B(p) \). Otherwise, there is some \( c \in C \) such that \( c \geq f(p) \implies c_0 \geq c \geq f(p) \implies c_0 \in B(p) \). Thus \( c_0 = \operatorname{lub} C \in B(p) \). Consequently, \( B(p) \in B_a \) and, since \( B(p) \subset A \), we have \( B(p) = A \).

Now we can show that \( P = A \).

i) \( a \in P \) since there is no \( y \in A \) with \( y < a \).

ii) Let \( p \in P \). Suppose that \( y \in A \) and \( y < f(p) \). Note that \( f(p) \in A \).

Then we must have \( y \leq p \). If \( y < p \) then \( f(y) \leq p \), and if \( y = p \) then \( f(y) = f(p) \leq f(p) \). Hence \( f(p) \in P \), i.e., \( \{P\} \subset P \).

iii) Let \( C \) be a chain in \( P \). Then \( c_0 = \operatorname{lub} C \in A \). Suppose that \( y \in A \) and \( y < c_0 \). If \( y \in P \), then \( c_0 \geq f(y) \). Otherwise, for each \( c \in C \), either \( y < c \) or \( y \geq f(c) \geq c \). But we can’t have \( y \geq c \) for all \( c \in C \), since \( c_0 = \operatorname{lub} C \). Hence, for some \( c \in C \), \( y < c \implies f(y) \leq c \leq c_0 \).

Therefore, \( c_0 \in P \).

Consequently, \( P \in B_a \) and, since \( P \subset A \), we have \( P = A \). Finally, if \( x \in A \) and \( y \in A \), then either \( x \leq y \) or \( x \geq f(y) \geq y \) and hence \( A \) is a chain.

A.C.2. \( \implies \) Z.L.2. Let \( S \) be inductively ordered. Let \( S \) be the collection of all non-empty subsets of \( S \) and let \( \varphi : S \to \cup\{A : A \in S\} \) be a mapping with \( \varphi(A) \in A \) for all \( A \in S \). Define \( f : S \to S \) by letting \( f(x) = x \) if \( x \) is a maximal element and letting \( f(x) = \varphi(\{y : y > x\}) \) if \( x \) is not a maximal element. Then \( f(x) \geq x \) for all \( x \in S \). Hence by the lemma, there is an \( x_0 \in S \) such that \( f(x_0) = x_0 \). Then, \( x_0 \) must be a maximal element of \( S \).