5-2014

Axioms of Set Theory and Equivalents of Axiom of Choice

Farighon Abdul Rahim
Boise State University, farrghunabdulrahim@u.boisestate.edu

Axioms of Set Theory and Equivalents of Axiom of Choice

Farighon Abdul Rahim
Advisor: Samuel Coskey
Boise State University

May 2014

1 Introduction

Sets are all around us. A bag of potato chips, for instance, is a set containing certain number of individual chip's that are its elements. University is another example of a set with students as its elements. By elements, we mean members. But sets should not be confused as to what they really are. A daughter of a blacksmith is an element of a set that contains her mother, father, and her siblings. Then this set is an element of a set that contains all the other families that live in the nearby town. So a set itself can be an element of a bigger set.

In mathematics, axiom is defined to be a rule or a statement that is accepted to be true regardless of having to prove it. In a sense, axioms are self evident. In set theory, we deal with sets. Each time we state an axiom, we will do so by considering sets. Example of the set containing the blacksmith family might make it seem as if sets are finite. In truth, they are not! The set containing all the natural numbers $\{1, 2, 3, \cdots\}$ is an infinite set. Our main goal for this paper will be the discussion of Axiom of Choice (AC) and its equivalents.

2 Some Logical and Set Theoretic Symbols

Before we begin our discussion, it is important that we list and define some of the symbols that will be used in this paper.

- $\forall$ and $\exists$ are quantifiers to denote "for all" and "there exists," respectively
- $\wedge$ and $\vee$ abbreviates "and" and "or," respectively
- $\neg$ abbreviates "not"
- $\rightarrow$ abbreviates "if, then"
- $\leftrightarrow$ abbreviates "if and only if"
- $dom(f)$ is the domain of the function f
- $ran(f)$ is the range of the function f

- $\{a, b\}$ is a set containing two elements
- $\{a_1, a_2, \cdots, a_n\}$ is a set with n elements, namely, $a_1, a_2, \cdots, a_n$

These symbols are not the only ones that we encounter as we proceed. New symbols will be defined as they are introduced.

3 Axioms of Set Theory

An important notion in set theory is the concept of belonging. Some x belongs to A would be the same as saying that x is an element that is contained in set A. Using the familiar notation " $\in$ ", we write $x \in A$ to denote belonging. If x does not belong in A, then we write $x \notin A$. But something even more elementary than belonging is the notion of equality, which represents the relationship between objects. To state $x = y$ would imply that x and y is the same object. When dealing with sets, how are we to tell if one set is equal to another set?

Axiom 1 (Axiom of Extension)**.** $\forall x(x \in A \leftrightarrow x \in B) \rightarrow A = B$

Any two given sets are equal if and only if they contain the same elements. It is important to note that Axiom of Extension is essential in describing belonging. So a set can be uniquely determined by its elements.

Definition 1. *If A and B are sets, we write $A \subset B$ if and only if for all x, $x \in A$ implies $x \in B$. If $A \subset B$, we say that A is a subset of B.*

Stated another way, we would say B includes A whenever $A \subset B$. By definition, it's easy to see that $A \subset A$. If $A \subset B$ and $B \subset A$, then $A = B$ by Axiom of Extension since both sets contain the identical elements. The term "inclusion" is also used to mean subset.

Axiom 2 (Axiom of Specification)**.** $\forall A \exists B \forall x(x \in B \leftrightarrow x \in A \wedge S(x))$

Axiom 2 states that for every set A and for every condition $S(x)$, there corresponds a set B whose elements are exactly those elements x of A for which $S(x)$ holds. By Axiom of Extension, set B is determined uniquely. To define this set, we write

$$B = \{x \in A : S(x)\}.$$

Intuitive, but rather informal, explanation of the axiom is that for some given arbitrary set, if we are able to make some intelligent assertion regarding the elements of this set, then we can specify a subset of the given set for which the assertion is true.

Lets consider our earlier example. Let A be the set which contains the children of blacksmith. If x is an arbitrary element of A, then the formula "x is a daughter" is true for some of the elements of A and false for others. After all, it would be illogical to assume the children of blacksmith are all boys. We then generate a subset of A using the formula. So the following

$$\{x \in A : x \text{ is a daughter}\}$$

is a set containing only the daughters of the blacksmith.

As another example, consider

$$\{x \in \mathbb{R} : x^2 = 5\} = \{-\sqrt{5}, \sqrt{5}\}$$

which would be a set consisting of two members.

For both of the examples, we used the term "formula". In the first example, the formula was "x is a daughter" and in the second example, the formula was "$x^2 = 5$." Lets define what is meant by "formula." There are two underlying type of formulas:

1. assertion of belonging, $x \in A$, and
2. assertion of equality, $A = B$.

Together, the two formulas are referred to as atomic formulas. Using the two atomic formulas, we append the logical operators $\wedge, \vee, \neg, \rightarrow, \leftrightarrow, \exists, \forall$ to formulate compound formulas. For example, appending the operator $\wedge$ would give "If A and B are formulas, then $A \wedge B$ is a formula." Lets note that "formula" is another term for "condition."

Axiom 3 (Axiom of Empty Set)**.** $\exists x \forall y \neg(y \in x)$

By Axiom of Empty Set, there exists a set with no elements. Such set is unique by Axiom of Extension and we denote the empty set by $\emptyset$. Lets note that for any set A, $\emptyset \subset A$. So, the empty set is a subset of every set.

Axiom 4 (Axiom of Pairing)**.** $\forall x \forall y \exists A \forall v (v \in A \leftrightarrow v = x \vee v = y)$

To state the axiom plainly, if A and B are sets, then there exists a set X such that $A, B \in X$.

If we're given two sets, a natural desire could be to combine the elements of the two into one set. So for A and B where $a \in A$ and $b \in B$, might it be possible to define some whole new set W such that $W = \{a, b\}$? Yes! We are able to do so by the following axiom.

Axiom 5 (Axiom of Unions)**.** $\forall X \exists Y \forall z (z \in Y \leftrightarrow \exists W (z \in W \wedge W \in X)$

Stated another way, for every collection C, there exists a set U such that if $x \in X$ for some X in C, then $x \in U$. Here, U is the union of the two sets and we write $\cup C$ where $\cup$ is used to denote the union of the collection. If $a \in A$ and $b \in B$, then $\cup \{a, b\}$ would simply translate to $a \cup b$. A simple example would be

$$\cup\{A\} = A.$$

An equally simple, yet more concrete, example would be to consider the union of the sets $\{a, b, c\}$ and $\{c, d, e\}$, where

$$\{a, b, c\} \cup \{c, d, e\} = \{a, b, c, d, e\}.$$

Axiom 6 (Axiom of Powers)**.** $\forall A \exists P \forall x [x \in P \leftrightarrow \forall y \in x (y \in A)]$

To put more simply, Axiom of Powers states that if X is a set, then $P(X)$ is a set, where $P(X)$ denotes the power set. Intuitively, we can think that for each set, there exists a collection of sets that contain among its elements all the subsets of the given set. If we assume for a moment that there exists some set E for which all the sets are subsets of E, then we write

$$P = \{X : X \subset E\}.$$

As a consequence, it is perfectly possible that P could contain other elements than the ones that are in X. An easy fix would be to apply the Axiom of Extension which would imply that the set is unique. An example would be the power set of the empty set, which gives us

$$P(\emptyset) = \{\emptyset\}.$$

So the power set of the empty set is the singleton set which contains only one element. The power set of a set containing only one element is

$$P(\{a\}) = \{\emptyset, \{a\}\}.$$

The power set of a set containing two elements is

$$P(\{a, b\}) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}.$$

Power set of a set containing three elements would give a set with eight elements and a power set of a set containing four elements would give a set with sixteen elements, and so on. In general, the power set of a set with n elements would give a resulting set that contains 2^n elements.

Definition 2. *For every set x, successor x^+ of x is $x^+ = x \cup \{x\}$.*

As an example, we define $0 = \emptyset$ since there are no elements in the empty set. Then by definition, we define 1 and 2 as

$$1 = 0^+ (= \{0\})$$

$$2 = 1^+ (= \{0, 1\})$$

Continuing in relatively the same manner, we define $3, 4, 5, 6, \cdots$ as we did above. Having defined definition 2, we can state the following axiom.

Axiom 7 (Axiom of Infinity)**.** $\exists x(\emptyset \in x \wedge \forall y \in x \exists z \in x(y \in z))$

By axiom of infinity, there exists a set containing 0 and containing the successor of each of its elements. Set A is a successor set if $0 \in A$ and if $x^+ \in A$ whenever $x \in A$. Axiom of infinity implies that there is at least one infinite set. The set of all the natural numbers, ω[1], for example, is an infinite set. Few other infinite sets include the set of real numbers, $\mathbb{R}$, the set of rational numbers, $\mathbb{Q}$, and the set of integer numbers, $\mathbb{Z}$. Aside from asserting that infinte sets exists, Axiom of Infinity is also giving the description of the elements of the sets as

$$\emptyset, \emptyset \cup \{\emptyset\}, \emptyset \cup \{\emptyset\} \cup \{\emptyset \cup \{\emptyset\}\}, \cdots.$$

1

[1] In set theory, ω denotes the set of natural numbers, $\mathbb{N}$

4 Axiom of Choice

Formulated by Ernst Zermelo in 1904, Axiom of Choice states that when we are given a collection of sets, say C, there is some function that chooses an element from each set. Such a statement begs for questions like "What is the function?" or "Does such a function even exist?" to be asked. Before we delve into the discussion of the axiom, we need some definitions to work with.

Definition 3. *A function $f : A \to B$ is injective (or one-to-one) if and only if for all $a, b \in A$, $f(a) = f(b)$ implies $a = b$.*

Definition 4. *A function $f : A \to B$ is surjective (or onto), if for any $b \in B$, there exists an $a \in A$ for which $f(a) = b$.*

Definition 5. *A function f is bijective if and only if it is both injective and sujective.*

Definition 6. *The cardinality of a finite set A is the number of elements that the set contains and is denoted $|A|$.*

For example, the set $A = \{0, 1, 2, 3\}$ contains 4 elements and thus has a cardinality $|A| = 4$. When there is no source of confusion, "size" may also be used to refer to the cardinality of a set. At first, it may seem as if cardinality of a set can only be considered when dealing with finite sets. Although it might be simpler to deal with finite sets, it is perfectly possible to consider the cardinality of infinte sets.

Definition 7. *The ordered pair of a and b is denoted by the set (a, b) and defined as*

$$(a, b) = \{\{a\}, \{a, b\}\}.$$

Definition 8. *If $\{X_i\}$ is a family of sets $(i \in I)$, the Cartesian Product of the family is the set of all families $\{x_i\}$ with $x_i \in X_i$ for each $i \in I$.*

Given two sets A and B, the Cartesian Product $A \times B$ is the unique set of all ordered pairs (a, b) where $a \in A$ and $b \in B$. If $A = \{0, 1\}$ and $B = \{2, 3\}$, then $A \times B = \{(0, 2), (0, 3), (1, 2), (1, 3)\}$.

Having defined Cartesian Product of a family of sets, we now formally state Axiom of Choice.

Axiom 8 (Axiom of Choice)**.** *The Cartesian Product of a nonempty family of nonempty sets is nonempty.*

In other words, if $\{X_i\}$ is a family of nonempty sets indexed by nonempty set I, then there exists a family $\{x_i\}$, $i \in I$, such that $x_i \in X_i$ for each $i \in I$.

Let C be the collection of nonempty sets. For the choice function f with domain C, if $A \in C$, then $f(A) \in A$.

Example 4.1. We can define a choice function f on a collection C of all nonempty subsets of $\{1, 2, 3, ...\}$ by letting $f(A)$ be defined to be the smallest member of the set A.

Consider $A = \{1, 2\}$. Then

$$f(A) = f(\{1, 2\}) = min(1, 2) = 1$$

is a choice function on C.

Example 4.2. Given infinitely many pair of socks, we choose one sock from each pair where our choice is arbitrary using AC. The choice function selects an element of each family from infinite family of sets of size 2.

Example 4.2 is fairly interesting. We must choose one sock from each pair at once. So the function f can be described as a simultaneous choice of a sock from each of the infinitely many pair of socks. We are allowed to do this by AC although the choice function is not clear as to how this might be done. If we instead assume finitely many pair of socks, then AC would not be needed.

To further investigate this idea, lets assume that we are given infinitely many pair of shoes. Unlike socks, we can choose one shoe from each pair because shoes can be ordered for the left and right foot. So infinitely many pair of shoes makes it possible to define the choice function explicitely. We may just pick left shoe from each pair.

Axiom of Choice is different from the other axioms. From intuitive perspective, AC is true. However, it is not possible to prove AC from other axioms of set theory. For that reason, we adopt it as an axiom. When first formulated, AC was subject to controversy. Much of the controversy was whether AC should be accepted as an axiom or not. Today, AC is accepted as a valid principle by most mathematicians and used to prove other theorems.

5 Equivalents of Axiom of Choice

Now that we have some understanding of AC, we next look at two equivalents and prove that they are consistent with AC. In particular, we consider Zorn's Lemma and the Well-Ordering (WO) Theorem. To begin with, we start by defining partial ordering.

Let S be a set. A relation $\leq$ on S is said to be a partial order if the following are satisfied:

1. $\forall x \in S$, $x \leq x$ (reflexive)
2. $\forall x, y, z \in S$, if $x \leq y$ and $y \leq z$, then $x \leq z$ (transitive)
3. $\forall x, y \in S$, if $x \leq y$ and $y \leq x$, then $x = y$ (antisymmetric)

Definition 9. *x is an upper bound of A if $a \leq x$ for each $a \in A$.*

Definition 10. *$a \in A$ is a maximal element if $a < x$ for no $x \in A$.*

Lemma 1 (Zorn's Lemma)**.** *If X is a partially ordered set such that every chain in X has an upper bound, then X contains a maximal elements.*

A chain is a totally ordered set. By totally ordered, we mean that $\forall x, y \in S$, either $x \leq y$ or $y \leq x$ is true. If A is a chain in X, then upper bound for A in X is guaranteed by the hypothesis of Zorn's lemma.

Proof of Zorn's Lemma. Let $\bar{s}$ be a function from X to $P(X)$ and define $\bar{s}(x) = \{y \in X : y \leq x\}$ where x is also an element of X and $S = ran(\bar{s})$ be partially ordered by inclusion. Then the function $\bar{s}$ is one-to-one. Necessary and sufficient condition for $\bar{s}(x) \subset \bar{s}(y)$ is that $x \leq y$. With this, the task of finding a maximal element in X is the same as the task of finding a maximal element in S. The hypothesis about chains in X implies the corresponding statement about chains in S.

Let $\mathcal{X}$ be the set of all chains in X. Then every member of $\mathcal{X}$ is included in $\bar{s}(x)$ for some $x \in X$. The collection of $\mathcal{X}$ is a non-empty collection of sets, partially ordered by inclusion, and such that if C is a chain in $\mathcal{X}$, then $\cup A$ belongs to $\mathcal{X}$ for $A \in C$. Since each set in $\mathcal{X}$ is dominated by some set in S, the passage from S to $\mathcal{X}$ cannot introduce any new maximal elements. One advantage of the collection $\mathcal{X}$ is the slightly more specific form that the chain hypothesis assumes. Instead of saying that each chain in C has some upper bound in S, we can say explicitly that the union of the sets of C, which is clearly an upper bound of C, is an element of the collection $\mathcal{X}$. Another technical advantage of $\mathcal{X}$ is that it contains all the subsets of each of its sets. This makes it possible to enlarge non-maximal sets in $\mathcal{X}$ slowly, one element at a time.

Now we can forget about the given partial order in X. In what follows we consider a non-empty collection $\mathcal{X}$ of subsets of a non-empty set X, which is subject to two conditions:

1. every subset of each set in $\mathcal{X}$ is in $\mathcal{X}$ and,
2. the union of each chain of sets in $\mathcal{X}$ is in $\mathcal{X}$.

The first condition implies that $\emptyset \in \mathcal{X}$. So now, our task is to prove that there exists a maximal set in $\mathcal{X}$.

Let f be a choice function for X. In other words, let f be a function from the collection of all nonempty subsets of X to X such that $f(A) \in A$ for all A in the domain of f. For each set $A \in \mathcal{X}$, let $\hat{A} = \{x \in X : A \cup \{x\} \in \mathcal{X}\}$. In other words, we let $\hat{A}$ be the set of all elements x of X whose adjunction to A produces a set in $\mathcal{X}$. Now, define a function $g : \mathcal{X} \to \mathcal{X}$ as follows: if $\hat{A} - A \neq \emptyset$, then $g(A) = A \cup \{f(\hat{A} - A)\}$. So if $\hat{A} - A = \emptyset$, then $g(A) = A$. It follows that from the definition of $\hat{A}$ that $\hat{A} - A = \emptyset$ if and only if A is maximal. In these terms, therefore, what we must prove is that there exists a set A in $\mathcal{X}$ such that $g(A) = A$. It turns out that the crucial property of g is the fact that $g(A)$ contains at most one more element than A. Let's note that $g(A)$ always includes A.

Now to facilitate the exposition, we introduce a temporary definition. We shall say that a subcollection $\mathcal{J}$ of $\mathcal{X}$ is a tower if

1. $\emptyset \in \mathcal{J}$,
2. if $A \in \mathcal{J}$, then $g(A) \in \mathcal{J}$ and,
3. if C is a chain in $\mathcal{J}$, then $\cup A \in \mathcal{J}$ for $A \in C$.

Towers surely exist. The whole collection $\mathcal{X}$ is one. Since the intersection of a collection of towers is again a tower, it follows, in particular, that if $\mathcal{J}_0$ is the intersection of all towers, then $\mathcal{J}_0$ is the smallest tower. So we must prove that the tower $\mathcal{J}_0$ is a chain.

We say that a set C in $\mathcal{J}_0$ is comparable if it is comparable with every set in $\mathcal{J}_0$. By this, we mean that if $A \in \mathcal{J}_0$, then either $A \subset C$ or $C \subset A$. By saying that $\mathcal{J}_0$ is a chain, we mean that all the sets in $\mathcal{J}_0$ are comparable. Comparable sets do exists. $\emptyset$ is one of them.

Now, let C be arbitrary (but temporarily) fixed comparable set. Suppose that $A \in \mathcal{J}_0$ and A is a proper subset of C. We assert that $g(A) \subset C$ because since C is comparable, then we have that either $g(A) \subset C$ or C is a proper subset of $g(A)$. If C is a proper subset of $g(A)$, then A us a proper subset of a proper subset of $g(A)$. But this contradicts the fact that $g(A) - A$ cannot be more than a singleton.

Next, let's consider the collection $\mathcal{U}$ of all sets A in $\mathcal{J}_0$ for which either $A \subset C$ or $g(C) \subset A$. The collection $\mathcal{U}$ is somewhat smaller than the collection of sets in $\mathcal{J}_0$ comparable with $g(C)$. If $A \in \mathcal{U}$, then, since $C \subset g(C)$, either $A \subset g(C)$ or $g(C) \subset A$. Now we assert that $\mathcal{U}$ is a tower. Since $\emptyset \subset C$, the first condition on towers is satisfied. Proving the second condition, that is, if $A \in \mathcal{U}$, then $g(A) \in \mathcal{U}$, involves us to split the discussion into three cases.

Case 1: A is a proper subset of C. Then $g(A) \subset C$ by the preceding paragraph, and therefore $g(A) \in \mathcal{U}$.

Case 2: $A = C$. So, we then $g(A) = g(C)$, so that $g(C) \subset g(A)$, and therefore $g(A) \in \mathcal{U}$.

Case 3: $g(C) \subset A$. Then $g(C) \subset g(A)$, and therefore $g(A) \in \mathcal{U}$. The third condition on towers, which states that the union of chains in $\mathcal{U}$ belongs to $\mathcal{U}$, is immediate from the definition of $\mathcal{U}$. With this, we can conclude that $\mathcal{U}$ is a tower included in $\mathcal{J}_0$, and therefore, since $\mathcal{J}_0$ is the smallest tower, we have $\mathcal{U} = \mathcal{J}_0$.

Our preceding considerations imply that for each comparable set C, the set $g(C)$ is also comparable. The is because for given C, form $\mathcal{U}$ as above. The fact that $\mathcal{U} = \mathcal{J}_0$, which means that if $A \in \mathcal{J}_0$, then either $A \subset C$ (in which case $A \subset g(C)$) or $g(C) \subset A$.

Now we know that $\emptyset$ is comparable and g maps comparable sets onto comparable sets. Since the union of a chain of comparable sets is comparable, it follows that the comparable sets (in $\mathcal{J}_0$) constitute a tower, and hence they exhaust $\mathcal{J}_0$. This is what we set out to prove about $\mathcal{J}_0$.

Since $\mathcal{J}_0$ is a chain, the union, syay A, of all the sets in $\mathcal{J}_0$ is itself a set in J_0. Since the union includes all the sets in $\mathcal{J}_0$, it follows that we have $g(A) \subset A$. Since always $A \subset g(A)$, it follows that $A = g(A)$. With this, we have proved Zorn's lemma [2]. □

Having proved Zorn's lemma, we next consider an application of the lemma. In different fields of mathematics, Zorn's lemma is sometimes the most crucial part of a proof that involves using the lemma. For instance, in Functional Analysis, Zorn's lemma is used to prove the Hahn–Banach Theorem. The lemma is also used to prove Tychonoff's Theorem in Topology. Another simple application includes proving that every vector space has a basis in Linear Algebra, which we will show.

Theorem 1. *Every vector space V has a basis.*

Proof. Let V be our vector space. Assume that V is nonzero. Ordered by inclusion, let K be the set whose elements are the linealy independent subsets of V. So K is a nonempty partially ordered set. Basis of V is the maximal element of K. If $T \subset K$ is a chain, then we have that $\cup T$ where T is a linearly independent subset of V, and hence is an upper bound for T. So by Zorn's Lemma, we conclude that every vector space V has a basis. □

Definition 11. *A partially ordered set is called well ordered if every nonempty subset of it has a smallest element.*

One consequence of definition 11 is that every well ordered set is totally ordered.

Definition 12. *An ordinal number is a well ordered set α such that $s(\beta) = \beta$ for all β in α where $s(\beta)$ is the initial segment $\{n \in \alpha : n < \beta\}$.*

Definition 13. *We say that a well ordered set A is a continuation of a well ordered set B, if, in the first place, B is a subset of A, if, in fact, B is an initial segment of A, and if, finally, the ordering of the elements in B is the same as their ordering in A.*

Theorem 2. *(Well-Ordering Theorem) Every set can be well ordered.*

Note: To prove WO, we invoke Zorn's lemma.

Proof. Let's assume that we are given a set X. Then consider the collection W of all well-ordered subsets of X. In other words, an element of W is a subset A of X with a well ordering of A. Let's partially order W by continuation.

The collection W is not empty, that is, $\emptyset \in W$. If $X \neq \emptyset$, then the less annoying elements of W can be exhibited like $\{(x, x)\}$ for any particular x of X. If C is a chain in W, then the union of the sets in C has a unique well ordering that makes the union of the sets larger than or equal to each set in C. So we have verified the principle hypothesis of Zorn's Lemma. In conclusion, there exists a maximal well-ordered set B in W and this set must be equal to the entire set X. With this, we have the proof of the Well-Ordering Theorem [2]. □

Now that we have stated and proved two variants of AC, we next prove that AC implies WO, WO implies Zorn's Lemma, and Wo implies AC.

AC $\rightarrow$ WO

Proof. Let A be a set. Finding the well-ordering of A means for us to find an ordinal number a and a a-sequence such that

$$r_0, r_1, \cdots, r_b$$

enumerates A where $b < a$. Then let C be a choice function on the family of all nonempty subsets of A. Then for the above sequence, we construct it by transfinite recursion as follows:

$$r_0 = C(A),$$

$$r_b = C(A - \{r_n : n < b\}).$$

We continue until we have all the elements of A [5]. □

WO $\rightarrow$ Zorn's lemma

Proof. Let $(P,<)$ be a nonempty partially ordered set and assume that every chain in P has an upper bound. We find a maximal element of P. By assumption, the set P can be well-ordered. By this, we mean that there is an enumeration

$$P = \{p_0, p_1, \cdots, p_b, \cdots\}$$

where $b < a$ and a is some ordinal number. So then by transfinite recursion, we let $c_0 = p_0$ and $c_b = p_y$ where y is the least ordinal such that p_y is an upper bound of the chain $Ch = \{ch_n : n < b\}$ and $p_y \notin Ch$. Let's note that $\{ch_n : n < b\}$ is always a chain and that p_y exists unless ch_{b-1} is a maximal element of P. As we keep going recursively, we will eventually get the maximal element of P [5]. $\square$

WO $\rightarrow$ AC

Proof. Let C be a set containing nonempty sets. Then $A = \cup_{S \in C} S$ is a set. By WO, A is well-ordered under some relation $<$. Since each S is a nonempty subset of A, then each S has a least member m with respect to the relation $<$. We define $f : C \rightarrow A$ by $f(S) = m$ where f is our choice function. With this, we can conclude that AC holds. $\square$

Since we have shown that AC implies WO, WO implies AC, and WO implies Zorn's lemma, we assume that Zorn's lemma implies AC is true.

6 Final Remark

Sets are quite interesting. Although a set can be thought of as being a basket, the number of objects that this basket can hold might not be obvious. For instance, if we consider the set of all the natural numbers $\mathbb{N} = \{0, 1, 2, \cdots\}$ and the set of all the even numbers $E = \{2, 4, 6, \cdots\}$, then $|\mathbb{N}| = |E|$ since there is a bijection from $\mathbb{N}$ to E for the function $f(n) = 2n$. One might easily think $|\mathbb{N}| > |E|$ since $\mathbb{N}$ includes both the even and odd numbers. To consider another example, let X be a set. Then $|X| < |P(X)|$ where $P(X)$ is the power set consisting of all the possible subsets of X. Of course, this is completely obvious. But still, it's quite mind blowing!

Using sets, we stated and proved two variants of Axiom of Choice, namely Zorn's Lemma and the Well-Ordering Theorem. There is a large collection of statements that are equivalent to AC. We only considered Zorn's Lemma and WO because the two are most widely used in proving other theorems. There is one other variant of AC. Although not as popular as Zorn's lemma or WO, it is one of the few variants that are used by mathematicians. It's also quite fasinating. The statement is known as Tukey's lemma.

Definition 14. *Set A is said to have finite character if $A \neq \emptyset$, and for each set X, X is a member of A if and only if every finite subset of X is a member of A.*

Lemma 2 (Tukey's Lemma)**.** *Let $\mathcal{F}$ be a nonempty family of sets. If $\mathcal{F}$ has finite character, then $\mathcal{F}$ has a maximal element (maximal with respect to inclusion).*

Using Zorn's lemma, we can easily prove Tukey's lemma. Conversely, we now show Tukey's lemma implies Axiom of Choice.

Tukey's Lemma $\rightarrow$ AC

Proof. Let $\mathcal{F}$ be a set containing family of nonempty sets. Our objective is to find a choice function $\mathcal{F}$. So then, we consider the family

$$\mathcal{G} = \{f : f \text{ is a choice function on some } \mathcal{E} \subset \mathcal{F}\}.$$

A subset of a choice function is a choice function and this it is easy to see that $\mathcal{G}$ has finite character. By assumption, $\mathcal{G}$ has a maximal element F. A simple appeal to the maximality of F shows that the domain of F is the whole family $\mathcal{F}$ [5]. $\square$

7 References

[1] Devlin, K. J., and Devlin, K. J. (1993). The joy of sets: Fundamentals of contemporary set theory. New York: Springer-Verlag.

[2] Halmos, P. R. (1960). Naive set theory. Princeton, N.J: Van Nostrand.

[3] Hamilton, A. G. (1982). Numbers, sets, and axioms: The apparatus of mathematics. Cambridge [Cambridgeshire: Cambridge University Press.]

[4] Hayden, S., Zermelo, E., Fraenkel, A. A., and Kennison, J. F. (1968). Zermelo-Fraenkel set theory. Columbus, Ohio: C. E. Merrill.

[5] Jech, T. J. (1973). The axiom of choice. Amsterdam: North-Holland Pub. Co.

[6] Nanzetta, P., and Strecker, G. E. (1971). Set theory and topology. Tarrytown-on-Hudson, N.Y: Bogden and Quigley.