An Introduction to the Axiom of Choice

- Some history and philosophy
- Some equivalents
- Some consequences

The Axiom of Choice (AC) is the most controversial axiom of Set Theory.

- AC seems intuitively appealing
- AC is everywhere, many results throughout mathematics rely on consequences of it
- AC also yields some consequences that seem counterintuitive

One form that was given by Russell in 1906 and Zermelo in 1908 is as follows:

If $\mathcal{T}$ is a disjoint collection of non-empty sets, there exists a set C which consists of one and only one element from each set in $\mathcal{T}$

So what is the controversy?

The controversy is about the word “exists”.

One group of mathematicians, the *intuitionists* or *constuctivists*, believe that a set exists only if each of its elements can be designated specifically or at least there is a law by which each of its elements can be constructed.

Other mathematicians take a more liberal viewpoint about "existence", that an axiom about the existence of sets may be used if it does not lead to a contradiction. The Axiom of Choice has become almost indispensable in mathematics since a large number of results have been obtained from it in almost all branches of mathematics without leading to contradiction.

Some mathematicians believe the axiom is obvious and tend to use it in proofs whether it is necessary or not.

The question of the “truth” of the axiom of choice is almost in the same position as that of the parallel postulate in geometry.

1922: Fraenkel proved the axiom is independent of the other axioms of set theory.

1940: Gödel proved that the axiom is consistent with the other axioms of set theory.

1963: Paul Cohen used a new technique called “forcing” to prove the independence in set theory of the Axiom of Choice and of the generalized continuum hypothesis.

How did this all start?

Choosing an unspecified element from a single set can be found in Euclid's Elements, if not earlier. This first stage also included the arbitrary choice of an element from each of finitely many sets. By induction, this procedure can be extended to any finite family of sets.

The second stage began when mathematicians made an infinite number of choices by stating a rule.

The third stage is the infinity of choices with the rule left unstated. This most likely began with Cauchy.

Failing to provide a rule allowed the fourth stage to emerge. Implicit use of the axiom (making an infinity of arbitrary choices) occurred in real analysis, algebraic number theory, point-set topology and set theory.

One of the first of these results was the Countable Union Theorem, resulting from Cantor's research in the 1870s:

The union of a countable family of countable sets is countable

A second example is:

Every infinite set has a denumerable subset

This result was demonstrated by Cantor in 1895, Borel in 1898, and Russell in 1902, all with the use of the denumerable form of the axiom of choice.

Cantor did not realize the watershed that he had crossed for the first time. Afterwards, analysts and algebraists increasingly used such arbitrary choices without remarking that a hidden assumption was involved.

In 1878, Cantor put forward his continuum hypothesis, conjecturing that every infinite subset of the continuum is either countable or has the cardinality of the continuum.

The importance of the continuum hypothesis was seen by Hilbert, who made this the first in his famous list of problems at his Paris lecture in 1900. Hilbert saw this as one of the most fundamental questions in mathematics, and he suugested that first one should try to prove another of Cantor's conjectures, that any set can be well-ordered.

Now let us recall some definitions:

The *Cartesian Product* of two sets A and B, denoted $A \times B$, is given by

$$A \times B = \{(a, b) | a \in A \text{ and } b \in B\}.$$

A *relation* on a set A is a subset R of the cartesian product $A \times A$. If R is a relation on A, we use the notation xRy to mean the same thing as $(x, y) \in R$.

A *partial order* in a set A is a relation R on A which has

1. reflexivity: $aRa \ \forall a \in A$
2. anti-symmetry: aRb and bRa imply $a = b$ $\forall a, b \in A$
3. transitivity: aRb and bRc imply $aRc \, \forall a, b, c \in A$

A relation R on a set A is called an *order relation* (or *simple order* or *linear order*) if it has the following properties:

1. comparability: For every a and b in A for which $a \neq b$, either aRb or bRa

2. nonreflexivity: For no a in A does the relation aRa hold

3. transitivity

We often denote an order relation by the symbol $<$.

Suppose A is a set ordered by the relation $<$, and $A_0 \subseteq A$. Then we say that a is the *smallest* element of A_0 if $a \in A_0$ and $a \leq x \ \forall x \in A_0$.

A set A with an order relation $<$ is said to be *well-ordered* if every non-empty subset of A has a smallest element.

The Well-Ordering Theorem: *Every set can be well-ordered*

1883: Cantor proposed the Well-Ordering Principle as a self-evident logical law.

1904: Zermelo used the axiom to prove the Well-Ordering Theorem.

Two other classical forms of the Well-Ordering Theorem:

Every set is equivalent to an ordinal number

Every set is equivalent to a subset of an ordinal number

Here are 4 common forms of the axiom of choice:

1. If $\mathcal{S}$ is a set of non-empty sets, there is a function f such that for every $X \in \mathcal{S}$, $f(X) \in X$
2. If $\mathcal{T}$ is a disjoint collection of non-empty sets, there exists a set C which consists of one and only one element from each set in $\mathcal{T}$
3. For every relation R there is a function f such that the domain of f equal the domain of R and $f \subseteq R$.
4. The Cartesian product of a set of non-empty sets is non-empty

The Law of the Trichotomy:

For all sets X *and* Y*, either* X *is equivalent to a subset of* Y*,* X *is equivalent to* Y*, or* Y *is equivalent to a subset of* X*.*

The equivalence of the axiom of choice and the trichotomy was given by Hartogs in 1915. It had been considered self-evident and was used without hesitation before 1915.

Maximal Principles

As mathematics developed further, there also developed a need for another non-constructive proposition. Kuratowski, Hausdorff, Zorn and others, used a principle to replace transfinite induction and the well-ordering theorem. It seems unrelated to the Axiom of Choice, but it is actually equivalent. It is often referred to as **Zorn's Lemma**.

If R *is a transitive relation on a non-empty set* X *and if every subset of* X *which is well-ordered (linearly ordered) by* R *has an* $R-$*upper bound, then there is an* $R-$*maximal element in* X.

Another form of Zorn's Lemma:

Let X be a non-empty partially-ordered set in which every totally ordered subset has an upper bound in X. Then X contains at least one maximal element.

An application of Zorn's Lemma

Theorem: *Let V be a vector space; then V has a basis.*

Another example using the Axiom of Choice:

If $x, y \in \mathbb{R}^n$, we say that x is **rationally equivalent** to y, written $x \sim y$, if $x - y \in \mathbb{Q}^n$.

Rational equivalence is an equivalence relation on the set $\mathbb{R}^n$, and so it divides it up into a collection of disjoint equivalence classes. Using the Axiom of Choice, we form a set $\mathcal{V}$ by choosing one representative from each equivalence class. Any such set $\mathcal{V}$ is referred to as Vitali's set.

Lemma: Let $q_1, q_2, ...$ be an enumeration of the countable set $\mathbb{Q}^n$. Then $\mathbb{R}^n$ can be represented by the disjoint union

$$\mathbb{R}^n = \bigcup_{i=1}^{\infty} \mathcal{V}_{q_i} = \bigcup_{i=1}^{\infty} (q_i \oplus \mathcal{V})$$

Thus every element $x \in \mathbb{R}^n$ has a unique representation in the form $x = q_i + v$ for some q_i and some $v \in \mathcal{V}$.

Theorem: *Vitali's set is NOT Lebesgue measurable*

Proof: Suppose that $\mathcal{V}$ is measurable. Then either (i) $m(\mathcal{V}) > 0$ or (ii) $m(\mathcal{V}) = 0$.

Some amazing consequences of the Axiom of Choice:

1. The Bolzano-Weierstrass Theorem: For metric spaces (or first countable topological spaces), compactness is equivalent to sequential compactness

2. The Hahn-Banach Extension Theorem: Every bounded linear functional on a subspace of a normed space has an extension to the whole space of the same norm

3. Tychonoff's Theorem: The product of a collection of compact topological spaces (with the product topology) is compact (this is equivalent to AC in ZF)

4. Every proper ideal of a ring with identity is contained in a maximal ideal (this is equivalent to AC in ZF)

5. The Weak Hilbert Nullstellenstatz: If $\mathbb{F}$ is an algebraically closed field, then an ideal $I \subseteq \mathbb{F}[x_1, ..., x_n]$ is proper if and only if its sero set $Z(I) \subseteq \mathbb{F}^n$ is non-empty

6. Multiplication of Cardinals: If X is an infinite set, Y is a nonempty set, and $|X| \geq |Y|$, then $|X \times Y| = |X|$

7. Another Cardinality Result: If X is an infinite set, then the cardinality of X is equal to the cardinality of $\bigcup_{n=1}^{\infty} X^N$ (finite sequences in X)

A few pure mathematicians and many applied mathematicians are not exactly comfortable with the Axiom of Choice. Although it simplifies some parts of mathematics, it also yields some results that are unrelated to, and even contrary to “ordinary” experience. Perhaps the most bizarre example is the **Banach-Tarski Paradoxical Decomposition**.

Banach and Tarski used the axiom to prove that it is possible to take a 3D closed unit ball

$$B = \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 \leq 1\}$$

and partition it into finitely many pieces, and then move these pieces in rigid motions (rotations and translations, with pieces permitted to move through one another) and reassemble them to form 2 copies of B.