Definition: Let *C* be a collection of nonempty sets ($\varnothing \notin$ C). A choice function, f, is a function such that for all X $\in$ S, f(X) $\in$ X. (Intuitively, we can *choose* a member from each set in that collection.)

Axiom of Choice (AoC): *Every family of nonempty sets has a choice function.*
The AoC was formulated by Zermelo in 1904.

Note: The axiom is non-constructive. It guarantees the existence for a choice function, but gives no indication how to make such a function. Because of the non-constructive nature of the axiom, and some of the fairly non-intuitive results of the axiom, it didn't gain broad acceptance until quite recently. In 1940, Kurt Gödel proved that (as long as the pre-existing axioms were without contradiction) adding the axiom of choice did not lead to a contradiction with the axioms of von Neumann-Bernays-Gödel set theory (a statement is true in ZF iff it is true in NBG). In 1963, Paul Cohen demonstrated that adding the negation of the axiom of choice to ZF also leads to no contradiction (assuming, once again that the ZF axioms were consistent to begin with), thus the AoC is independent of Zermelo-Fraenkel set theory.

The axiom gets its name not because mathematicians prefer it to other axioms.
— A. K. Dewdney

The ZF axioms, with the addition of the AoC, are referred to as *ZFC*.

In general, the AoC is required in order to make arbitrary choices from a family of sets. There are a few specific instances [1] where the AoC is not necessary to accomplish this task:

- If each set in the family is a singleton.
- If there are only finite sets in the family (induction on the number of sets in the family suffices to show that selection can occur for any finite number of sets)
- If each X $\in$ S contains only a finite number of ordered (distinguishable) items, (e.g., f(X) = the least element of a finite set)

Russell phrased it as this: If we have $\aleph_0$ null pairs of shoes, then we can select on shoe from each pair without the axiom of choice (just choose the left shoe for each pair). But, if we had $\aleph_0$ pairs of socks, then we need the axiom of choice to pick one from each set (because socks are not distinguishable from each other).

The AoC has many equivalent statements [2]. A few of the most compelling are:

- For any relation R there is a function F $\subseteq$ R, with domain(F) = domain(R)
- The Cartesian product of a non-empty family of non-empty sets is non-empty.
- For any two sets C and D, C $\preceq$ D or D $\preceq$ C (i.e., the cardinality of any two sets is comparable)

The AoC is also equivalent to a very non-intuitive theorem, known as the Well-ordering theorem. For this theorem, we'll need some additional definitions (that we'll use later, as well)

Definition A binary relation < is a *partial-ordering* of a set S if for all p, q, r in S:

$p \nless p$ (< is irreflexive)

If p < q and q < r then p < r (< is transitive)

Remark: Texts differ as to whether a partial-ordering is defined to be like the real number relation < or $\leqslant$. In this case, we follow the convention of [1], [2], and [4], and use a strict partial ordering (i.e., < is irreflexive rather than reflexive). [3] and [5] use a reflexive partial ordering (based on $\leqslant$). In all cases, the texts quite reasonably use a symbol for the relation that suggests the real relation similar to their convention.

Definition A binary relation < is a *total-ordering* (or *linear-ordering*) of a set S if

(S, <) is a partial-ordering.

for all p, q in S, p < q or p = q or q < p

Example: A good example of a partially ordered set is $\wp(\{1, 2, ..., n\})$ under the relation of $\subset$ ("proper subset"). A set is never a proper subset of itself, and the relation is transitive. In addition, this is not a total-ordering, as $\{1,2\},\{3,4\} \in \wp(\{1,2,...,n\})$, but $\{1,2\} \not\subset \{3,4\}$ and $\{3,4\} \not\subset \{1,2\}$.

Definition An element $a \in X$ is the *least element* of X if $(\forall x \in X)\, a \le x$

Definition A binary relation < is a *well-ordering* of a set S if

(S, <) is a total-ordering.

Every subset of S has a least element

Example: Any subset of the positive integers is a well-ordering under the standard "<" operator (by the Well Ordering Principle of the integers).

Zermelo's Well Ordering Theorem: *Every set can be well ordered.*

Prooflet:

Let A be a set. In order to show that A can be well ordered, we will first show that we can arrange the elements of A into a (possibly transfinite) sequence, $\{a_\alpha : \alpha \in \theta\}$ (where α and θ are ordinal numbers) that enumerates A. Let f be a choice function from the set of non-empty subsets of A (the existence of such a choice function is guaranteed to exist through the AoC).

Now apply transfinite induction:

If A is non-empty, let $\alpha = 0$ (i.e. the first ordinal). $a_0 = f(A)$ (if not, then we are done, as any relation produces a well-ordering on an empty set!). Continue advancing through the ordinals, letting $a_\alpha = f\left(A - \left\{a_\xi : \xi < \alpha\right\}\right)$ until $A - \left\{a_\xi : \xi < \alpha\right\}$ is empty. We are guaranteed to run out of set members prior to running out of ordinals by Hartogs' Theorem [2] (i.e., for every set, there is an ordinal with cardinality larger than the cardinality of our set). Note, by construction of $\{a_\alpha : \alpha \in \theta\}$, $A = \bigcup_{\alpha \in \theta} a_\alpha$ and $(\forall \alpha, \beta \in \theta)\, \alpha \neq \beta \Rightarrow a_\alpha \neq a_\beta$.

Given the sequence $\{a_\alpha : \alpha \in \theta\}$, we will define our relation as follows: $(\forall \alpha, \beta \in \theta)\, a_\alpha < a_\beta \Leftrightarrow \alpha < \beta$. This is a well ordering, (through the well ordering of the ordinals).
□

Proposition *Zermelo's Well Ordering Theorem is equivalent to the Axiom of Choice.*
Assume Zermelo's Well Ordering Theorem.
Let S be a family of non-empty sets. All sets are well ordered, so $\bigcup_{A \in S} A$ is well-ordered, and can be enumerated as a (possibly transfinite) sequence. Define our choice function $f(A)$ as the least element of A as defined by our sequence.
□

Note: *Banach-Tarski Paradox*: Using the AoC, it is possible to take the 3-dimensional closed unit ball, and partition it into finitely many pieces, and move those pieces in rigid motions (i.e., rotations and translations, with pieces permitted to move through one another) and reassemble them to form two copies of B. (Note, the pieces of the ball are not Lebesgue measurable.)

> Banach and Tarski had hoped that the physical absurdity of this theorem would encourage mathematicians to discard AoC. They were dismayed when the response of the math community was `Isn't AoC great? How else could we get such counterintuitive results?'

Another statement called Zorn's Lemma is also equivalent to the AoC. We again require some additional definitions:

Definition An element $a \in X$ is a *maximal element* of X if $(\forall x \in X)\, a \not< x$.
Note: This maximal element is not guaranteed to be unique. In the trivial case, the empty set can be used as a relation that makes every element of the set "maximal".

Definition A set X is a *chain* if it is totally ordered.

Definition: An element $a \in X$ is a *upper bound* for P⊂X if $(\forall p \in P)\, a \nless p$.

Zorn's Lemma: If (X, <) is a nonempty partially ordered set such that every chain in X has an upper bound, then X has a maximal element.
Note: In the interest of time, I have chosen a short (but vacuous) proof Zorn's lemma (it amounts to an appeal to the really, really big, and is very similar to the proof for the well ordering theorem). There are other perfectly reasonable proofs of Zorn's lemma, but they are lengthy affairs; see [3] for a possibly more satisfying proof.
Prooflet:
The general construction that we follow is to construct a single chain (we'll represent this as a possibly transfinite sequence) in X that leads to a maximal element using a choice function, f, from ℘(X)-{∅} to X.

Let $a_0 = f(X)$. Now, let $a_1 = f(X - \{a_0\})$ such that $a_0 < a_1$. (This can be thought of as making an arbitrary selection and then discarding those values that do not satisfy this condition; keeping track the members that have been discarded and those that have been kept, removing both sets of them from future selections).

We proceed (i.e., applying transfinite induction) in this way until we run out of set members at some value a_θ (once again, we are guaranteed to run out of set members prior to ordinals, as with the well-ordering theorem).

Note, we have constructed a chain in X (relative to the provided ordering relation!) that has a maximal element a_θ. This element is a maximal element for the set X by construction.
□

Proposition *Zorn's Lemma is equivalent to the Axiom of Choice.*
Proof:
Assume Zorn's Lemma.
Let S be a family of nonempty sets.
Let's consider a choice function on S. Let P={f : f is a choice function on some Z ⊆S}.
Note, (P, ⊂) is a nonempty, partially ordered set under the set, and each chain has an upper bound (each chain could certainly not be "larger" than S by construction), so P has an upper bound, which is S. This upper bound is a choice function on S.
□

Zorn's lemma (and thus the AoC) is used to prove several well known results:
=> Every vector space has a basis
=> Every field has a unique algebraic closure

<=> *Tychonoff's Theorem*: Any product of compact topological spaces is compact (Note: for compact Hausdorff spaces Tychonoff's Theorem is equivalent to the Boolean Prime Ideal Theorem (Rubin and Scott 1954) and hence weaker than the AoC.)

More generally, the AoC is used in ZFC to prove:

=> The countable union of countable sets is countable.

=> The Baire Category Theorem (a weakened version of AoC is required here: The Axiom of Dependent Choice)

=> Every infinite set has a denumerable subset

<= The Generalized continuum hypothesis

The generalized continuum hypothesis (GCH) is not only independent of ZF, but also independent of ZF plus the axiom of choice (ZFC). However, ZF plus GCH implies the AoC, making GCH a strictly stronger claim than AoC.

Q: What's sour, yellow, and equivalent to the axiom of choice? A: Zorn's lemon.

The Axiom of Choice is obviously true, the well-ordering principle obviously false, and who can tell about Zorn's lemma?

- Jerry Bona

References:

[1] Jech, Thomas; Set Theory; Springer 2002.
[2] Enderton, Herbert B; Elements of Set Theory; Academic Press, 1977.
[3] Halmos, Paul; Naive Set Theory; Springer, 1974.
[4] Munkres, James; Topology (2nd ed.); Prentice Hall, 2000.
[5] Introduction to Topology (2nd ed.); Dover Press, 1999.