6-14-1999

Review: Set Theory, Functions, and Relations

This is a brief review of material on set theory, functions, and relations. I'll concentrate on things which are often not covered (or given cursory treatment) in courses on mathematical proof.

Definition. Let A and B be sets, and let $f : A \to B$ be a function from A to B.

1. f is **injective** if for all $a_1, a_2 \in A$, $f(a_1) = f(a_2)$ implies $a_1 = a_2$.
2. f is **surjective** if for all $b \in B$, there is an element $a \in A$ such that $f(a) = b$.
3. f is **bijective** if f is injective and surjective.

Definition. Let $f : A \to B$ be a function. Let $B_0 \subset B$.

1. The **image** of f is the set
$$f(A) = \{f(a) \mid a \in A\}.$$
2. The **inverse image** of B_0 is the set
$$f^{-1}(B_0) = \{a \in A \mid f(a) \in B_0\}.$$

Remark. f is surjective if and only if $f(A) = B$.

Example. (a) Let $f : A \to B$ and $g, h : B \to C$ be functions. Suppose that f is surjective and $g \cdot f = h \cdot f$. Prove that $g = h$.

Let $b \in B$. I must show that $g(b) = h(b)$.

Since f is surjective, I may find an element $a \in A$ such that $f(a) = b$. Then $g \cdot f = h \cdot f$ implies that
$$g(f(a)) = h(f(a)), \quad \text{so} \quad g(b) = h(b).$$

Therefore, $g = h$. $\square$

(b) Let $f : A \to B$. Suppose that for all $g, h : B \to C$, if $g \cdot f = h \cdot f$, then $g = h$. Prove that f is surjective.

First, I'll take care of the cases where $A = \emptyset$ or A contains a single element.

If $A = \emptyset$ and $B = \emptyset$, then f is vacuously surjective: Every element of B (there are none) is an image of an element of A.

If $A = \emptyset$ and $B \neq \emptyset$, then B contains at least one element b. Define $g : B \to \{0, 1\}$ by $g(b) = 0$ and $h : B \to \{0, 1\}$ by $h(b) = 1$. It is vacuously true that $g(f(a)) = h(f(a))$ for all $a \in A$ (since there are no a's in A), but $g \neq h$. Therefore, the hypothesis isn't satisfied, and this case is ruled out.

Suppose A consists of a single element a. If $B = \emptyset$, then f is vacuously surjective. If B consists of a single element, then f is obviously surjective.

Finally, suppose that B contains more than one element. $f(a)$ is an element of B; suppose that $b_0 \in B$ and $b_0 \neq f(a)$. Consider the functions $g, h : B \to B$ given by
$$g(b) = b \quad \text{for all} \quad b \in B,$$
$$h(b) = \begin{cases} b & \text{if } b \neq b_0 \\ f(a) & \text{if } b = b_0 \end{cases}.$$

Then $g(f(a)) = f(a) = h(f(a))$, so $g \cdot f = h \cdot f$. However, $g \neq h$. Therefore, the hypothesis isn't satisfied, and this case is ruled out.

This concludes the proof for the trivial cases.

Assume that A contains more than one element.

Suppose, on the contrary, that $f(A) \neq B$. I'll show that the hypothesis is violated.

If $b \in f(A)$, *choose* an element $a_b \in A$ such that $f(a_b) = b$. Since A contains more than one element, I may find elements $a_1, a_2 \in A$ such that $a_1 \neq a_2$. Define $g, h : B \to A$ by

$$g(b) = \begin{cases} a_b & \text{if } b \in f(A) \\ a_1 & \text{if } b \notin f(A) \end{cases} \quad \text{and} \quad g(b) = \begin{cases} a_b & \text{if } b \in f(A) \\ a_2 & \text{if } b \notin f(A) \end{cases}.$$

If $a \in A$, then $f(a) \in f(A)$, so

$$g(f(a)) = a_{f(a)} = h(f(a)).$$

Thus, $g \cdot f = h \cdot f$.

However, since $f(A) \neq B$, there is an element $b_0 \in B - f(A)$. For this element, $g(b_0) = a_1$, but $h(b_0) = a_2$. Therefore, $g \neq h$.

This proves that the hypothesis is violated. This contradiction in turn establishes that $f(A) = B$, so f is surjective. □

Remark. This result is often expressed by saying that surjective functions are **right-cancellable**. The dual result is also true: Injective functions are **left-cancellable**.

Example. Let $f : A \to B$, and suppose that $B_1, B_2 \subset B$. Prove that $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$.

Suppose $a \in f^{-1}(B_1 \cup B_2)$. Then $f(a) \in B_1 \cup B_2$, so $f(a) \in B_1$ or $f(a) \in B_2$.

If $f(a) \in B_1$, then $a \in f^{-1}(B_1)$, so $a \in f^{-1}(B_1) \cup f^{-1}(B_2)$. If $f(a) \in B_2$, then $a \in f^{-1}(B_2)$, so $a \in f^{-1}(B_1) \cup f^{-1}(B_2)$.

Therefore, $a \in f^{-1}(B_1) \cup f^{-1}(B_2)$, so $f^{-1}(B_1 \cup B_2) \subset f^{-1}(B_1) \cup f^{-1}(B_2)$.

Conversely, suppose that $a \in f^{-1}(B_1) \cup f^{-1}(B_2)$, so $a \in f^{-1}(B_1)$ or $a \in f^{-1}(B_2)$.

If $a \in f^{-1}(B_1)$, then $f(a) \in B_1$, so $f(a) \in B_1 \cup B_2$, and hence $a \in f^{-1}(B_1 \cup B_2)$. If $a \in f^{-1}(B_2)$, then $f(a) \in B_2$, so $f(a) \in B_1 \cup B_2$, and hence $a \in f^{-1}(B_1 \cup B_2)$.

Therefore, $a \in f^{-1}(B_1 \cup B_2)$, so $f^{-1}(B_1) \cup f^{-1}(B_2) \subset f^{-1}(B_1 \cup B_2)$.

This proves that $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$. □

Definition. Let A be a set. A relation $<$ on A is a **simple order** (or **linear order**) if:

1. (**Comparability**) For all $a, b \in A$, exactly one of the following is true: $a < b$, $b < a$, or $a = b$.

2. (**Nonreflexivity**) There is no element $a \in A$ such that $a < a$.

3. (**Transitivity**) For all $a, b, c \in A$, if $a < b$ and $b < c$,then $a < c$.

Example. Let $<$ be a simple order on a set A. The **dictionary order** (or **lexicographic order**) on $A \times A$ is given by: $(a, b) < (a', b')$ if and only if

1. $a < a'$; or

2. $a = a'$, and $b < b'$.

The words in a dictionary are ordered by their first letter. Among words with the same first letter, they are ordered by their second letter, and so on. This explains the name of this order relation.

To give a specific example, in $\mathbb{R}^2$ under the dictionary order, $(1, 4) < (2, 4)$ and $(1, 6) < (1, 8)$.

Definition. If A is an ordered set, $a, b \in A$, and $a < b$, then the **open interval** from a to b is the set

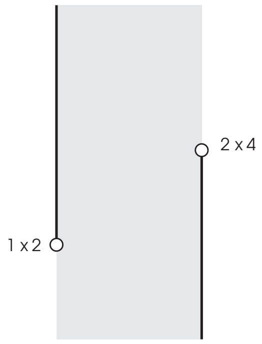
$$(a, b) = \{x \mid a < x < b\}.$$

It may happen that this set is empty. In that case, since there are no elements "between" a and b, it's natural to say that a is the **predecessor** of b and b is the successor of a.

Unfortunately, "(a, b)" can mean a *point* in a Cartesian product or an *open interval* in a simple order. To avoid confusion, I'll sometimes write $a \times b$ for the point when both concepts occur in the same discussion.

Example. Consider the dictionary order on $\mathbb{R}^2$. The open interval from the point 1×2 to the point 2×4 is the union of:

1. The open ray $\{1 \times y \mid y > 2\}$.
2. The lines $\{x \times y \mid 1 < x < 2\}$.
3. The open ray $\{2 \times y \mid y < 4\}$.



$\square$

Definition. Suppose A and B are sets with simple orders $<_A$ and$<_B$, respectively. Let $f : A \to B$ be a function. f is **order-preserving** if for all $a_1, a_2 \in A$,

$$a_1 <_A a_2 \quad \text{implies} \quad f(a_1) <_B f(a_2).$$

Example. With the usual order relation on $\mathbb{R}$, define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^3$. Then f is order-preserving.

In fact, a function $\mathbb{R} \to \mathbb{R}$ is order-preserving if it is an **increasing** function. $\square$

Definition. Let A and B be ordered sets. A and B have the **same order type** if there's an order-preserving bijection from A to B.

Definition. A set is **well-ordered** if every nonempty subset has a smallest element.

The **Well-Ordering Axiom** says that every set can be well-ordered. This is an axiom of set theory; it's equivalent to other results which can be taken as axioms, such as the **Axiom of Choice** and **Zorn's lemma**.

Well-ordering means that for a given set, there is *some* order relation on the set relative to which the set is well-ordered. That well-ordering will usually have little to do with familiar order relations.

Definition. Let A be an ordered set, and let $B \subset A$.

1. $c \in B$ is the **largest element** of B if $b < c$ for all $b \in B$, $b \neq c$.

   Likewise, $c \in B$ is the **smallest element** of B if $c < b$ for all $b \in B$, $b \neq c$.

2. An element $a \in A$ is an **upper bound** for B if $b < a$ for all $b \in B$, $b \neq a$. If B has an upper bound, B is **bounded above**.

   Likewise, an element $a \in A$ is a **lower bound** for B if $a < b$ for all $b \in B$, $b \neq a$. If B has a lower bound, B is **bounded below**.

3. If the set of upper bounds of B is nonempty, then the smallest element of that set is the **least upper bound** for B. It is denoted $\operatorname{lub} B$ or $\sup B$.

   Likewise, if the set of lower bounds of B is nonempty, the largest element of that set is the **greatest lower bound** for B. It is denoted $\operatorname{glb} B$ or $\inf B$.

4. A set has the **least upper bound property** if every nonempty subset that is bounded above has a least upper bound.

   A set has the **greatest lower bound property** if every nonempty subset that is bounded below has a greatest lower bound.

Example. The real numbers with the usual ordering satisfy the least upper bound property. In fact, it is one of the axioms for the real numbers. □

Mathematics is based on a collection of axioms for set theory. They are called the **Zermelo-Fraenkel Axioms with the Axiom of Choice** (often abbreviated to **ZFC**).

While the Axiom of Choice is accepted and used by most mathematicians, the situation was different earlier in this century. Some mathematicians felt that the Choice Axiom was problematic and refused to use it; others took care to note when they used the Choice Axiom in a proof. Many of the objections stemmed from the often startling results which can be proved using the axiom.

Axiom of Choice. Let $\{X_a\}$ be a collection of disjoint nonempty sets. There exists a set X containing exactly one element from each X_a.

The Axiom of Choice is equivalent to the Well-Ordering Axiom, and also to an important result known as **Zorn's lemma**. The statement of Zorn's lemma requires some preliminaries.

Definition. A relation $\leq$ on a set X is a **partial order** if:

1. (**Reflexivity**) $x \leq x$ for all $x \in X$.

2. (**Antisymmetry**) If $x \leq y$ and $y \leq x$, then $x = y$.

3. (**Transitivity**) If $x \leq y$ and $y \leq z$, then $x \leq z$.

Example. The set of subsets of a set is partially ordered by inclusion. □

Example. The usual "less than or equal to" relation $\leq$ on $\mathbb{R}$ is a partial order. □

Definition. Let X be a partially ordered set.

1. A subset $Y \subset X$ is **totally ordered** (or **linearly ordered**) if for all $x, y \in Y$, either $x \le y$ or $y \le x$.

   A totally ordered subset is also said to be a **chain**.

2. If $Y \subset X$, then Y has an **upper bound** if there is an element $x \in X$ such that $y \le x$ for all $y \in Y$.

3. An element $x \in X$ is **maximal** if $y \in X$ and $x \le y$ implies $x = y$.

Zorn's Lemma. If X is a nonempty partially ordered set in which every chain has an upper bound, then X has maximal elements.

Zorn's lemma is an immensely useful form of the Axiom of Choice. For example, you may have seen a proof that every *finite-dimensional* vector space has a basis. Zorn's lemma can be used to show that *every* vector space has a basis.

Example. Let G be a group and let $g \in G$, $g \ne 1$. There is a subgroup $H < G$ such that $g \notin H$ and H is maximal among subgroups with this property.

Let $\mathcal{S}$ be the collection of subgroups of G which do not contain g. $\mathcal{S} \ne \emptyset$, since $\{1\} \in \mathcal{S}$. $\mathcal{S}$ is partially ordered by inclusion.

Let $\mathcal{T}$ be a chain in $\mathcal{S}$. Let

$$H_0 = \bigcup_{H \in \mathcal{T}} H.$$

Thus, H_0 is just the union of all the elements of the chain.

I claim that H_0 is a subgroup of G which does not contain g.

First, $1 \in H$ for all $H \in \mathcal{T}$, so certainly $1 \in H_0$.

If $h, h' \in H_0$, then $h \in H$ and $h' \in H'$ for some $H, H' \in \mathcal{T}$. But $\mathcal{T}$ is a chain, so either $H \subset H'$ or $H' \subset H$.

Suppose $H \subset H'$. Then $h, h' \in H'$, so (since H' is a subgroup) $hh'^{-1} \in H'$. Therefore, $hh'^{-1} \in H_0$, and hence H_0 is a subgroup.

If $g \in H_0$, then $g \in H$ for some $H \in \mathcal{T}$. But $H \in \mathcal{S}$, so $g \notin H$ by definition of $\mathcal{S}$. Therefore, $g \notin H_0$.

I've shown that H_0 is a subgroup of G which does not contain g, so $H_0 \in \mathcal{S}$. Since $H \subset H_0$ for all $H \in \mathcal{T}$, it follows that H_0 is an upper bound for $\mathcal{T}$ in $\mathcal{S}$.

Since every chain has an upper bound, Zorn's lemma implies that $\mathcal{S}$ has a maximal element H. Then H is a subgroup of G which is maximal among subgroups which do not contain g. $\square$
