Cantor's Theorem

Joe Roussos

1 Preliminary ideas

Two sets have the same number of elements (are equinumerous, or have the same cardinality) iff there is a *bijection* between the two sets.

Mappings: A mapping, or function, is a rule that associates elements of one set with elements of another set. We write this $f : X \to Y$, f is called the function/mapping, the set X is called the domain, and Y is called the codomain. We specify what the rule is by writing $f(x) = y$ or $f : x \mapsto y$.

e.g. $X = \{1, 2, 3\}, Y = \{2, 4, 6\}$, the map $f(x) = 2x$ associates each element $x \in X$ with the element in Y that is double it.

A bijection is a mapping that is injective and surjective.[1]

- **Injective (one-to-one)**: A function is injective if it takes each element of the domain onto *at most one* element of the codomain. It never maps *more than one* element in the domain onto *the same* element in the codomain. Formally, if f is a function between set X and set Y, then f is injective iff

$$\forall a, b \in X, f(a) = f(b) \to a = b$$

- **Surjective (onto):** A function is surjective if it maps something onto *every* element of the codomain. It can map more than one thing onto the same element in the codomain, but it needs to hit everything in the codomain. Formally, if f is a function between set X and set Y, then f is surjective iff

$$\forall y \in Y, \exists x \in X, f(x) = y$$

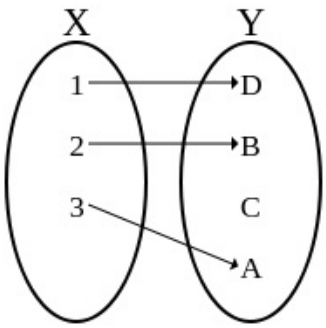

Figure 1: Injective map.

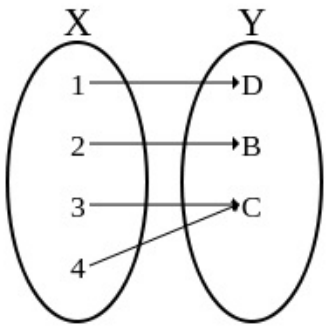

Figure 2: Surjective map.

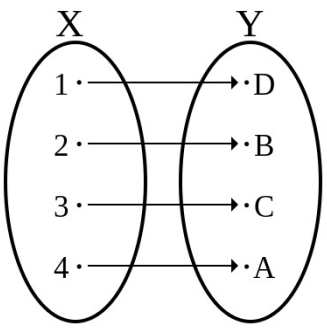

Figure 3: Bijective map.

Final piece of notation: The **image** of X under f is the set of elements in the codomain that f maps onto. It is written as $f(X)$ and is sometimes called the range of f. e.g., In Figure 1 $f(X) = \{A, B, D\}$.

[1] Images from Wikipedia.

2 Cantor's Theorem

For any set A, the cardinality of A is strictly less than the cardinality of A's power set: $|A| < |\mathcal{P}(A)|$

Proof: To prove this, we will show (1) that $|A| \leq |\mathcal{P}(A)|$ and then (2) that $\neg(|A| = |\mathcal{P}(A)|)$. This is equivalent to the *strictly less than* phrasing in the statement of the theorem.

(1) $|A| \leq |\mathcal{P}(A)|$: To show this, we just need to produce a bijection between A and a subset of $\mathcal{P}(A)$. Then we will know A is the same size as that subset, which itself can't be bigger than the size of $\mathcal{P}(A)$.

Consider the set $E = \{\{x\} : x \in A\}$, the set of all single-element subsets of A. Clearly $E \subset \mathcal{P}(A)$, because it is made up of subsets of A. Incidentally, it is a proper subset, since we know it doesn't contain $\varnothing$. The map $g : A \to E$ defined by $g(x) = \{x\}$ is one-to-one and onto. How do we know this? (This is laboured, but useful to be certain that you understand!)

- One-to-one: Suppose we have $x, y \in A$ and $g(x) = g(y)$. Then by the definition of injectiveness above, we want to be sure that this means $x = y$, if g is going to be one-to-one. $g(x) = \{x\}$ and $g(y) = \{y\}$, so $g(x) = g(y)$ means $\{x\} = \{y\}$. These two one-element sets can only be equal if their members are equal, so $x = y$. Therefore g is one-to-one.
- Onto: Is it true that $\forall y \in E, \exists x \in A, g(x) = y$? Yes. We know that $E = \{\{x\} : x \in A\}$ so $\forall y \in E$, $\exists x \in A$ such that $y = \{x\}$. That's because each element of E just is a set with an element from A as its sole member. And since $g(x) = \{x\}$, we have $\forall y \in E, \exists x \in A, g(x) = y$, so g is surjective.

Therefore $|A| = |E| \leq |\mathcal{P}(A)|$.

(2) $\neg\big(|A| = |\mathcal{P}(A)|\big)$: To show this, we need to show that there is no bijection from A to $\mathcal{P}(A)$. Now we already know that there is an injective map from $A \to \mathcal{P}(A)$, it is just g above. (There I defined $g : A \to E$ but since $E \subset \mathcal{P}(A)$ we can also think of g as mapping into $\mathcal{P}(A)$. (Make sure you're happy with this.) So what we need to show is that there is no *surjection* from A to $\mathcal{P}(A)$.

Here's a general way to establish this. Let $f : A \to \mathcal{P}(A)$ be any function. Suppose that f is surjective. We will now derive a contradiction showing that this assumption cannot be true.

Consider the set $B = \{x \in A : \neg(x \in f(x))\}$, the set of all those elements x in A whose image under f does not include x itself. (To give you a feel for this: if f is like the function we used above, $g : x \mapsto \{x\}$, then nothing would be in B because every element x ends up in its image $g(x) = \{x\}$.)

Clearly $B \subset A$, as there is nothing in B that isn't in A. So $B \in \mathcal{P}(A)$, by definition of the power set. If f is surjective, then it must map something onto B, i.e. $\exists a \in A$, such that $f(a) = B$.

Now the question is: is $a \in B$? Suppose that it is. If $a \in B$ then by definition of B we know $\neg(a \in f(a))$. But $f(a) = B$, so a has to be in $f(a)$ if $a \in B$. This is a contradiction.

Suppose that $\neg(a \in B)$. Then a must fail the membership condition for B, which is $\neg(a \in f(a))$. So in this case $a \in f(a)$. But $f(a) = B$, so a has to be in B if $a \in f(a)$. This is a contradiction.

Therefore, our assumption must be wrong. f cannot be a surjection. But f was just any arbitrary function. So there is *no* bijection between A and $\mathcal{P}(A)$.

So, $|A| \neq |\mathcal{P}(A)$. Combined with the result from (1), this concludes the proof.