

Introductory Cardinality Theory

Alan Kaylor Cline

Although by name the theory of set cardinality may seem to be an offshoot of combinatorics, the central interest is actually infinite sets. Combinatorics deals with finite sets. As will be seen, the tool employed for the majority of the theory is establishing one-to-one correspondences¹ between sets.

One version of cardinality theory uses “cardinal numbers” - a type of number used to quantify cardinalities. We will not use cardinal numbers here although the results are quite similar.

The initial definition of finiteness says no more than a set is finite if we can “count” its elements. Counting means establishing a one-to-one correspondence with a set of consecutive integers beginning with one. Thus, if f is a function mapping $\{1, 2, \dots, n\}$ one-to-one onto a set A , we “count” A as $f(1), f(2), \dots, f(n)$. The fact that f is one-to-one and onto ensures that each element of A is counted (onto-ness) and that no element is counted more than once (one-to-one-ness).

Definition 1: A set A is *finite* with cardinality n if it is empty or if there exists a one-to-one function mapping $\{1, 2, \dots, n\}$ onto A . A set is *infinite* if it is not finite.

Thus, since the set of lower case Latin characters $\{a, b, c, \dots, z\}$ can be put into one-to-one correspondence with $\{1, 2, \dots, 26\}$, that set is finite. Let’s prove that the set of natural numbers $\mathbb{N} = \{0, 1, 2, \dots\}$ is infinite. Notice to do this we must show that for any n , **no function exists** mapping $\{1, 2, \dots, n\}$ one-to-one onto \mathbb{N} . (It is not sufficient to show that some function doesn’t work – we must establish that it is impossible to have such a function). As one might expect, such arguments use proof-by-contradiction.

Example 1: The set of natural numbers is infinite.

Proof: Suppose there exists an n and a function f mapping $\{1, 2, \dots, n\}$ one-to-one onto \mathbb{N} . The set $\{f(1), f(2), \dots, f(n)\}$ is finite so it has a maximum element. Let $m = 1 + \max\{f(1), f(2), \dots, f(n)\}$. Since $m \in \mathbb{N}$ and f is onto, there must exist some $k \in \{1, 2, \dots, n\}$ such that $f(k) = m$. But then also $f(k) \in \{f(1), f(2), \dots, f(n)\}$ and we would have that $m = 1 + \max\{f(1), f(2), \dots, f(n)\} \geq 1 + f(k) > f(k) = m$, which is a contradiction. We conclude that no such n and f exists, so \mathbb{N} is infinite. \square

That wasn’t hard but it wasn’t much fun either. It might be nice to have a more direct method for proving a set is infinite. Such is provided by this alternative definition.

¹ A “one-to-one correspondence” between sets A and B implies that there is a one-to-one function mapping A onto B (and thus another function, the inverse of the first, mapping B onto A).

Definition 2: A set A is *infinite* if there exists a one-to-one function mapping A onto a proper subset of A . A set is *finite* if it is not infinite.

Whenever we see two definitions for the same concept, we should ask “are they logically identical?”. The answer this time is “not quite”. Let me deal with the easy part of the comparison of the two definitions first and then I’ll discuss the tricky stuff.

It is a simple matter to show that if a set is finite according to Definition 1 then it must be finite under Definition 2. (If a set were finite according to Definition 1 but infinite under Definition 2, we would end up with a one-to-one correspondence between two finite sets of different cardinalities. That can’t happen.) By using the contrapositive to this, we see that if a set is infinite according to Definition 2 then it must be infinite under Definition 1. This is good because, since we will want to use Definition 1 for finite sets and Definition 2 for infinite sets, we know that we are using the “stronger” definitions (i.e., the claims will hold for either definition).

Turning now to the thornier question of implication in the other direction, the answer is that with the standard axioms of set theory², one cannot prove that a set infinite under Definition 1 must also be infinite under Definition 2. In order to establish the implication we add the famous Axiom of Choice (loosely stated as “Given any collection of nonempty sets, we can choose a member from each set in that collection”). The Axiom of Choice allows us to construct the function that Definition 2 requires³.

To see the ease of using Definition 2 for infiniteness, let’s reprove that the set of natural numbers is infinite using Definition 2 instead of 1.

Example 2: The set of natural numbers is infinite.

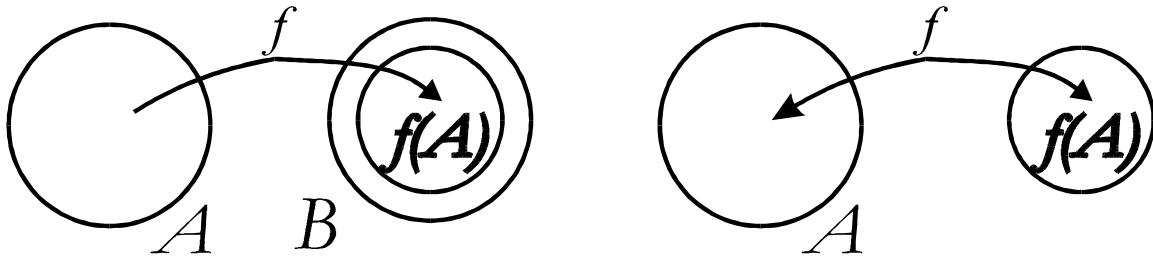
Proof: Let \mathbb{N}^+ denote the set of positive natural numbers and consider the function $f : \mathbb{N} \rightarrow \mathbb{N}^+$ defined by $f(n) = n + 1$. (This is often called the “successor function”.) Since for $n \neq m$, $f(n) = n + 1 \neq m + 1 = f(m)$, and f is one-to-one. But f is also onto since for any $n \in \mathbb{N}^+$, $n - 1 \in \mathbb{N}$ and $f(n - 1) = n$. Because $\mathbb{N}^+ \subseteq \mathbb{N}$ but $0 \in \mathbb{N} \sim \mathbb{N}^+$, \mathbb{N}^+ is a proper subset of \mathbb{N} . \square

Notice that since the definition requires that the set be put into one-to-one correspondence with a proper subset, we must prove that the function is both one-to-one and onto. The following lemma will allow us to cut some of the work. It says that a one-to-one function is also onto - and thus invertible - if its range is restricted to exactly the image of its domain.

² By "standard axioms of set theory", I am referring to Zermelo Fraenkel set theory (see J.M. Henle. *An Outline of Set Theory*. Springer Verlag, 1986).

³ Without getting too deeply into Gödel Consistency Theory, I’ll add that, although the Axiom of Choice is not derivable from the standard set theoretic axioms, if those other axioms are consistent then the Axiom of Choice is consistent with them. Lastly, one should not feel covered in shame because we use the Axiom of Choice. Mathematicians do it all of the time without blinking. It does have some interesting consequences however, such that a line segment 1 inch long can be cut into a finite number of pieces and then glued back together to form a segment one light-year long.

Lemma: If $f:A \xrightarrow{1-1} B$ then f maps A one-to-one onto B $f(A)$ and thus is invertible.



Consider now Definition 2 but omit the necessity for the function to be onto:

Definition 2': A set A is *infinite* if there exists a one-to-one function mapping A into a proper subset of A . A set is *finite* if it is not infinite.

Obviously if a set is infinite under Definition 2 it will be infinite under definition 2'. From the lemma, however, the opposite is also true. To see this, suppose A is infinite under Definition 2' and so a function f mapping A into a proper subset A' of A . exists. Let $\tilde{A} = f(A)$ and notice that $\tilde{A} \subseteq A'$ so \tilde{A} must also be a proper subset of A . We conclude that A is also infinite under Definition 2. By using the contrapositive we may show that Definitions 2 and 2' for finite sets are also equivalent.

Since Definition 2' saves some work, we will use it. More generally we will use Definition 2' for infinite sets and Definition 1 for finite sets.

The set of natural numbers has been proved to be infinite using both Definition 1 and Definition 2. The proof using Definition 2' is the same as that using Definition 2 except that the second to last sentence (showing that the mapping is onto) could be eliminated. We state it as our first theorem and then prove the real interval $[0,1]$ also is infinite..

Theorem 1: The set \mathbf{N} of natural numbers is infinite.

Theorem 2: The real interval $[0,1]$ is infinite.

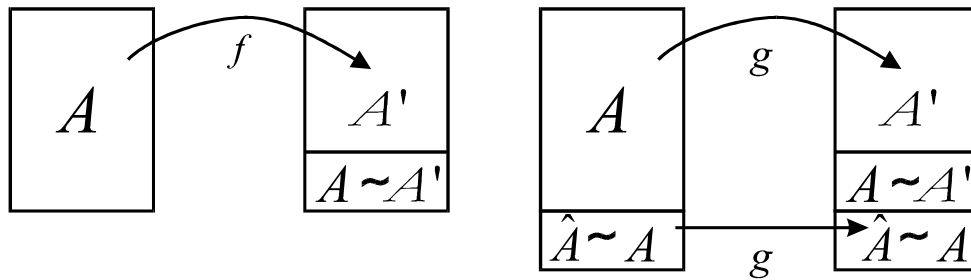
Proof: Consider the function $f(x) = x/2$ defined on $[0,1]$. Clearly f maps the interval into $[0,1/2]$, a proper subset of $[0,1]$. Since for $x \neq y$, $f(x) = x/2 \neq y/2 = f(y)$, f is one-to-one. Thus $[0,1]$ is infinite. \square

The substance of the next theorem may seem obvious: if a set is infinite and has additional elements added, it still is infinite.

Theorem 3: A superset of an infinite set is infinite.

Proof: Let A be an infinite set and assume $A \subseteq \hat{A}$. We seek to show that \hat{A} is infinite as well. By Definition 2', we know there exists an $f:A \xrightarrow{1-1} A'$ for some $A' \subset A$. (We will

use the symbol \subset to indicate proper subsets.). Define a new function g , an "extension" of f to all of \hat{A} , as follows $g(a) = \begin{cases} f(a) & \text{if } a \in A \\ a & \text{if } a \in \hat{A} \sim A \end{cases}$. First we must show that g is one-to-one on \hat{A} . To that end, consider distinct elements $a_1, a_2 \in \hat{A}$. Either $a_1, a_2 \in A$, $a_1, a_2 \in \hat{A} \sim A$, or one element is in each of A and $\hat{A} \sim A$ (and, without loss of generality, we assume $a_1 \in A$ and $a_2 \in \hat{A} \sim A$). If $a_1, a_2 \in A$, then $g(a_1) = f(a_1) \neq f(a_2) = g(a_2)$, since f is one-to-one. If $a_1, a_2 \in \hat{A} \sim A$, then $g(a_1) = a_1 \neq a_2 = g(a_2)$. Finally, if $a_1 \in A$ and $a_2 \in \hat{A} \sim A$, then $g(a_1) = f(a_1) \in A'$, so $g(a_1) \in A$, but $g(a_2) = a_2 \in \hat{A} \sim A$. Since $g(a_1) \in A$ and $g(a_2) \in \hat{A} \sim A$, $g(a_1) \neq g(a_2)$. We have shown that g is one-to-one. Finally, since $A' \subset A$, there exists some element $\bar{a} \in A \sim A'$. We want to show that for no $a \in \hat{A}$ is $g(a) = \bar{a}$. To that end, suppose $g(a) = \bar{a}$. If $a \in A$, then $g(a) = f(a) \in A'$, so $g(a) \neq \bar{a}$. If $a \in \hat{A} \sim A$, then $g(a) = a \in \hat{A} \sim A$, so $g(a) \neq \bar{a}$. We have a contradiction in either case, so we know there exists no $a \in \hat{A}$ so that $g(a) = \bar{a}$. We conclude that g maps \hat{A} into a proper subset of itself and thus \hat{A} is infinite. \square



An easy corollary follows from the fact that if A is a subset of \hat{A} , then \hat{A} is a superset of A . If \hat{A} were finite yet A infinite, we would have a contradiction of the theorem. So, if \hat{A} is finite so must A be finite as well.

Corollary: A subset of a finite set is finite.

We use Theorem three to show that it is easy to establish that a very strange set is infinite by showing that it has an infinite subset.

Example 3: The set of ratios of integers to odd numbers is infinite.

Proof: All natural numbers can be expressed as ratios of themselves to 1. Thus the set of natural numbers is a subset of this set. The set of natural numbers is infinite, therefore this set is infinite. \square

The next theorem emphasizes that infinite cardinality is preserved by one-to-one mappings.

Theorem 4: Let A be infinite and $f: A \xrightarrow{1-1} B$, then B is infinite.

Proof: Let A be an infinite set and $f: A \xrightarrow{1-1} B$. If we set $B' = f(A)$, then from the lemma above f maps A one-to-one onto B' . The idea of the proof is to show that B' is infinite. Since B is a superset of B' , Theorem 3 will guarantee that B is infinite as well. We know that since A is infinite there is some function g mapping A one-to-one into some proper subset A' of itself. Furthermore, since $f: A \xrightarrow{1-1}_{\text{onto}} B'$, we have $f^{-1}: B' \xrightarrow{1-1}_{\text{onto}} A$ so $h = f \circ g \circ f^{-1}: B' \rightarrow B'$ is defined. Since h is the composition of one-to-one functions, it is one-to-one. Now consider some element $a \in A \sim A'$. (We know such an element exists since A' is a proper subset of A .) Let $\bar{b} = f(a)$. If we can show that there is no $b \in B'$ so that $h(b) = \bar{b}$, then h will have been shown to have mapped B' one-to-one into a proper subset of itself - and thus be infinite. To that end, suppose there is such a $b \in B'$ with $h(b) = \bar{b}$. That says $f(g(f^{-1}(b))) = \bar{b}$, so $g(f^{-1}(b)) = f^{-1}(\bar{b})$. But since $\bar{b} = f(a)$, we have $g(f^{-1}(b)) = a$. This is a contradiction because $a \in A \sim A'$ but g maps into A' , so we would have $g(f^{-1}(b))$ both being an element of A' and being outside of A' . We conclude that no such element $b \in B'$ so that $h(b) = \bar{b}$ exists and h maps B' one-to-one into a proper subset of itself. B' is then shown to be infinite and B , a superset of B' , is also infinite. \square

In Example 3, we saw that Theorem 3 simplified showing sets were infinite when we could find infinite subsets. Theorem 4 extends that simplification but no longer must we find infinite sube sets - we may establish infiniteness by finding one-to-one mappings from known infinite sets.

Example 4: The set of points in the plane $\mathcal{W} = \{(x, y) \mid x^2 + (y-3)^4 \leq 6\}$ is infinite.

Proof: Consider the mapping $f: [0,1] \rightarrow \mathcal{W}$ defined by $f(t) = (t, 3)$. Notice that this actually does map into \mathcal{W} since for $0 \leq t \leq 1$, $t^2 + (3-3)^4 = t^2 \leq 1 \leq 6$. To show f is one-to-one, consider distinct $s, t \in [0,1]$, $f(s) = (s, 3) \neq (t, 3) = f(t)$. Thus f is one-to-one and by Theorem 4, \mathcal{W} is infinite. \square

As was stated initially, the theory of cardinality deals with one-to-one correspondences between sets. We will now refine the concept of infinite set by distinguishing those sets that can be put into one-to-one correspondence with the natural numbers from those that cannot.

Definition 3: A set \mathcal{A} is *countably infinite* if there exists a one-to-one function mapping \mathbb{N} onto \mathcal{A} . A set is *countable* if it is finite or countably infinite. A set is *uncountably infinite* if it is not countable.

Let's show that \mathbb{Z} , the set of integers is countable infinite.

Example 5: The set of integers is countably infinite.

Proof: Consider $f : \mathbb{N} \rightarrow \mathbb{Z}$ defined by $f(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ -(n+1)/2 & \text{if } n \text{ is odd} \end{cases}$. We seek to show f maps \mathbb{N} one-to-one onto \mathbb{Z} . First we will show that f is one-to-one. Consider distinct $n, m \in \mathbb{N}$. Either both of n and m are even, both are odd, or one is even and one is odd. If n and m are even, $f(n) = n/2 \neq m/2 = f(m)$. If n and m are odd, $f(n) = -(n+1)/2 \neq -(m+1)/2 = f(m)$. Lastly, if n is even and m is odd then $f(n) \geq 0 > f(m)$. We conclude that f is one-to-one. Lastly we will show that f is onto. Consider any $k \in \mathbb{Z}$. If $k \geq 0$ then $2k \in \mathbb{N}$ and is even so $f(2k) = 2k/2 = k$. If $k < 0$ then $-(1+2k) \in \mathbb{N}$ and is odd so $f(-(1+2k)) = -(-(1+2k)+1)/2 = k$. We conclude that f is both one-to-one and onto and thus \mathbb{Z} is countably infinite. \square

We could show that sets such as the even integers, the odd integers, and the powers of two are all countably infinite. A legitimate question then is "Are there any uncountably infinite sets?". The following theorem shows that there are. The proof uses the classic diagonalization argument of Georg Cantor. It is an proof by contradiction: we assume that the real interval $[0,1]$ is countably infinite, attempt to "count" all of them, find that at least one is missing, and get a contradiction.

Theorem 5: The real interval $[0,1]$ is uncountably infinite.

Proof: Theorem 2 guarantees that the interval is infinite. To prove that it is uncountably infinite, let us assume that it is countably infinite. Thus there exists $g : \mathbb{N} \xrightarrow[\text{onto}]{1-1} [0,1]$. If we can show that there is a number in $[0,1]$ that is not equal to $g(i)$ for any $i \in \mathbb{N}$, then g is not onto and we have a contradiction. We then may conclude that $[0,1]$ is uncountably infinite.

To this end, consider the decimal expansions of $g(0), g(1), \dots$. For $i \in \mathbb{N}$, let $g(i)$ be expressed as $.d_1^i d_2^i \dots d_{i+1}^i \dots$. Some real numbers may have two different decimal expansions: one terminating in zeros, the other in nines. If there is the option, we will choose the expansion terminating in nines. Notice that 0 itself will then be the only number that terminates in zeros and that every number in $[0,1]$ has a unique such decimal expansion.

Now consider constructing the number $e = .e_1 e_2 \dots e_{i+1} \dots$ defined by $e_{i+1} = \begin{cases} 1 & \text{if } d_{i+1}^i \neq 1 \\ 2 & \text{if } d_{i+1}^i = 1 \end{cases}$ for $i = 0, 1, 2, \dots$. First recognize that e is a real number and $e \in [0,1]$ (in fact $e \in [1/9, 2/9]$), and yet, e cannot equal $g(i)$ for any $i \in \mathbb{N}$. Suppose $e = g(i)$, then the decimal expansions of the two must agree in every position, but in fact they differ in the $i+1^{\text{st}}$ decimal digit. Since for no $i \in \mathbb{N}$ is $e = g(i)$, g is not onto. We conclude that no such g exists and that $[0,1]$ is uncountably infinite. \square

The name "diagonalization" is suggested in the construction of the number e in the proof. If one were to make a column of the decimal expansions of $g(0), g(1), \dots$, then e is created by altering the entries on the diagonal of this table.

To come are several theorems that can be summarized as "the union of a countable collection of countable sets is countable." For that purpose the following theorem is very helpful.

Theorem 6: If there exists a function $f : \mathbf{N} \xrightarrow{\text{onto}} A$ then A is countable.

Proof: If A is finite we are done. Assume then that A is infinite. Consider an array of the elements of A induced by the function f

$$a_0 a_1 \cdots a_k \cdots$$

where $a_k = f(k)$. We will define a new function $g : \mathbf{N} \rightarrow A$ and show that this function is both one-to-one and onto. To that end, define $g(0) = a_0$ and then remove all copies of elements of the array equal to a_0 . Define next $g(1)$ equal to the leading element on the remaining array – and then remove all copies of it. In general, define $g(i)$ as the leading element of the array after all copies of $g(0), g(1), \dots, g(i-1)$ have been removed. Since A is infinite will not end. Thus for every $i \in \mathbf{N}$, $g(i)$ is defined. We need to show this mapping is both one-to-one and onto. To that end consider to properties of g :

- For any element a_j in the array (and therefore in A), there will be some value of $i \in \mathbf{N}$, so that $g(i) = a_j$. In fact, the value of $i \leq j$.
- By construction, the value of $g(i)$ must be distinct from $g(0), g(1), \dots, g(i-1)$, since all copies of those elements were removed prior to the definition of $g(i)$.

The first property guarantees that g is onto and the second property guarantees that it is one-to-one. We have established that either the set A is finite or it is countably infinite. \square

This Theorem has an immediate consequence.

Theorem 7: A subset of a countable set is countable.

Proof: Assume set \hat{A} is countable and $A \subseteq \hat{A}$. If \hat{A} is empty then it is obviously finite. Henceforth, we assume \hat{A} is nonempty. If \hat{A} is finite then so will A be because of Theorem 3. If \hat{A} is countably infinite then there exists a function $f : \mathbf{N} \xrightarrow{\text{onto}} \hat{A}$. Choose any fixed element $a' \in A$ and define a new function $g : \mathbf{N} \rightarrow A$ as follows:

$$g(n) = \begin{cases} f(n) & \text{if } f(n) \in A \\ a' & \text{if } f(n) \notin A \end{cases}$$

Since f is onto, for any element $a \in A$ there exists an n such that $f(n) = a$. But then $g(n)$ is also equal to a since $g(n) = f(n)$ if $f(n) \in A$. We conclude that g is onto and from Theorem 6, A is countable. \square

Corollary: A superset of an uncountably infinite set is uncountably infinite.

Proof: Assume set A is uncountably infinite and $A \subseteq \hat{A}$. If \hat{A} were countable then so would A . by Theorem 7. That is a contradiction so \hat{A} must be uncountably infinite. \square

Theorem 8: The union of a finite collection of finite sets is finite.

Proof: For $n \geq 1$, let $\{A_1, A_2, \dots, A_n\}$ be a collection of finite sets. We seek to prove a stronger version of the theorem – that the cardinality of the union is less than or equal to the sum of the cardinalities of the sets A_i . To this end, for $1 \leq i \leq n$ let $n_i = \#(A_i)$. We proceed by induction. For $n = 1$, $\#(A_1) = \#(A_1)$. Assume now that all unions of n sets: $\#(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n \#(A_i)$. We seek to show that $\#(\bigcup_{i=1}^{n+1} A_i) \leq \sum_{i=1}^{n+1} \#(A_i)$. But then $\#(\bigcup_{i=1}^{n+1} A_i) = \#(\bigcup_{i=1}^n A_i \cup A_{n+1}) \leq \#(\bigcup_{i=1}^n A_i) + \#(A_{n+1}) \leq \sum_{i=1}^n \#(A_i) + \#(A_{n+1}) = \sum_{i=1}^{n+1} \#(A_i)$. \square

Theorem 9: The union of a countably infinite collection of finite sets is countable.

Proof: Let the finite sets be A_0, A_1, \dots , define n_k to be the cardinality of A_k , for $k \in \mathbb{N}$, and let $A = \bigcup_{i=0}^{\infty} A_i$. Since for $k \in \mathbb{N}$, each A_k is finite, the elements can be ordered in some form $a_1^k a_2^k \cdots a_{n_k}^k$, where n_k is the cardinality of A_k . Consider an array of all of the elements of A stretched out as $\overbrace{a_1^0 a_2^0 \cdots a_{n_0}^0}^{A_0} \overbrace{a_1^1 a_2^1 \cdots a_{n_1}^1}^{A_1} \cdots \overbrace{a_1^k a_2^k \cdots a_{n_k}^k}^{A_k} \cdots$. We will define a function $f: \mathbb{N} \rightarrow A$ and then prove that this function is onto. To define the function we use the array: $f(i) = a_l^k$ if a_l^k is the i th element of the array (starting the counting from $i = 0$.) Since every element of the array (hence every element of A) is the image under f for some i , f is onto. From Theorem 4, A is countable. \square

Example 5: The set of rational numbers is countably infinite.

Proof: For $k \in \mathbb{N}$, define $A_k = \{p/q \mid p, q \in \mathbb{Z} \wedge -k \leq p \leq k \wedge -k \leq q \leq k \wedge q \neq 0\}$. Notice that each A_k is finite (in fact, having exactly $(2k+1) \cdot 2k$ elements) and given any rational number, this will be contained in p/q is A_k for $k = \max\{|p|, |q|\}$. Thus the set of rationals equals $\bigcup_{k \in \mathbb{N}} A_k$ and this set is countable by Theorem 9. Since it contains the natural numbers as a subset, it is infinite. We conclude that it is countably infinite. \square

The last of these theorems allows us to union a countably infinite collection of countably infinite sets and still the result is countably infinite.

Theorem 10: The union of a countably infinite collection of countably infinite sets is countably infinite.

Proof: Let the countably infinite sets be A_0, A_1, \dots , and let $A = \bigcup_{i=0}^{\infty} A_i$. Since for $k \in \mathbb{N}$, each A_k is countably infinite, there is a function $f_k: \mathbb{N} \xrightarrow{\text{onto}} A_k$. Thus, the elements can

be ordered in the form $a_0^k a_1^k \cdots a_i^k \cdots$, where $a_i^k = f_k(i)$. Form now a new collection of sets $\{B_0, B_1, \dots\}$ defined as follows: for $i \in \mathbb{N}$ $B_i = \{a_i^0, a_{i-1}^1, a_{i-2}^2, \dots, a_1^{i-1}, a_0^i\}$. From Theorem 9, A is countable. Since the infinite set $A_0 \subseteq A$, A is infinite, and therefore it is countably infinite. \square

Theorem 11: Let A be uncountably infinite and $f : A \xrightarrow{1-1} B$, then B is uncountably infinite.

Proof: From the lemma above, $f : A \xrightarrow[\text{onto}]{1-1} B'$, where $B' = f(A)$. By Theorem 4, B' is infinite. Suppose B' is countably infinite. Then there exists a function $g : \mathbb{N} \xrightarrow[\text{onto}]{1-1} B'$. Consider the function $h : \mathbb{N} \rightarrow A$ defined as $h = f^{-1} \circ g$. This function is both one-to-one and onto since f^{-1} and g are. We would then have that A is countably infinite but that would be a contradiction. We conclude that B is uncountably infinite. \square