Final Exam Solutions

Math 114S, Winter 2022

Instructions. Please carefully read the instructions for each section. Also, for each problem on this exam, unless explicitly stated otherwise, you may use all axioms of ZFC. There are 100 points in total.

Short Answer Questions.

For each question in this section, you do not need to give any justification for your answers.

Question 1 (16 points)

Mark each of the following as True or False.

(a) Without making any other changes to our development of mathematics within ZFC, we could have defined an ordered pair as a function with domain $\{\varnothing, \{\varnothing\}\}$.

> **Solution:** $\boxed{\text{False.}}$ Since our definition of "function" depended on the definition of "ordered pair," this definition would be circular.

(b) For every set x, $\text{rank}(\text{tc}(x)) = \text{rank}(x)$, where $\text{tc}(x)$ denotes the transitive closure of x.

> **Solution:** $\boxed{\text{True.}}$ First note that since $x \subset \text{tc}(x)$ we have
>
> $$\text{rank}(x) = \sup\{\text{rank}(y) + 1 \mid y \in x\} \leq \sup\{\text{rank}(y) + 1 \mid y \in \text{tc}(x)\} = \text{rank}(\text{tc}(x)).$$
>
> Now we need to show that $\text{rank}(\text{tc}(x)) \leq \text{rank}(x)$. There are two ways to do prove this, the "top down" approach and the "bottom up" approach.
>
> The "top down" approach is to recall that the transitive closure of x is the smallest transitive set of which x is a subset. Now recall that if $\text{rank}(x) = \alpha$ then $x \subseteq V_\alpha$. Since V_α is a transitive superset of x, we must have $\text{tc}(x) \subset V_\alpha$. Hence $\text{tc}(x)$ has rank at most α.
>
> The "bottom up" approach is to recall our construction of $\text{tc}(x)$ as $x \cup \bigcup x \cup \bigcup\bigcup x \cup \ldots$. Note that if $y \in \bigcup x$ then there must be some $z \in x$ such that $y \in z$. Hence $\text{rank}(y) < \text{rank}(z) < \text{rank}(x)$. Similarly, one can show by induction on ω that every element of $\text{tc}(x)$ must have rank less than x. Since the rank of $\text{tc}(x)$ is the supremum of the successors of these ranks, $\text{rank}(\text{tc}(x)) \leq \text{rank}(x)$.

(c) There is an ordinal $\alpha \geq \omega$ such that $\alpha^\alpha = \alpha$ (where α^α here is referring to ordinal exponentiation, not cardinal exponentiation).

> **Solution:** $\boxed{\text{False.}}$ Since ordinal addition, multiplication and exponentiation are all order preserving, for any $\alpha \geq \omega$ we have
>
> $$\alpha < \alpha + 1 \leq \alpha + \alpha = \alpha \times 2 \leq \alpha \times \alpha = \alpha^2 \leq \alpha^\alpha.$$
>
> Actually, it's true more generally that for any $\alpha > 1$ and any $\beta < \beta'$, $\alpha^\beta < \alpha^{\beta'}$ (this can be proved using transfinite induction). Applying this with $\beta = 1$ and $\beta' = \alpha$ also gives the result.

(d) If M, A, and B are sets such that $A, B \in M$ and $M \vDash$ "$|A| = |B|$" then $|A| = |B|$.

Solution: $\boxed{\text{False.}}$ Suppose B is a much larger set than A but their intersections with M have the same size. Then M could "think" that A and B are in bijection (because M ignores all elements of B that are not in M). For a concrete example, take $A = \omega$, $B = \aleph_1$ and M some set that contains A, B, all natural numbers, all sets of natural numbers of size 1, all sets of natural numbers of size 2, all ordered pairs of natural numbers and the identity on ω.

Common Mistakes: This was the most frequently missed True/False question. One reason it might be tricky is that it is actually true if M is transitive (though even this requires a bit of thought because "bijection" does not necessarily mean quite the same thing in M that it does "in the real world"). It also might be tricky because it can be unintuitive to think about what is "true according to M" and because this topic was introduced only at the end of the course.

Question 2 (5 points)

Write the formula "R is a binary relation" in the language of set theory—i.e. using only variables, logical symbols and the symbols $\in$ and $=$.

Solution: Recall that a binary relation is just a set of ordered pairs. So we want to say that every element of R is an ordered pair. Substituting in the formal definition of "ordered pair," this gives us

$$\forall p \left[p \in R \implies \exists a \, \exists b \left(\forall x \, (x \in p \iff (\forall y \, (y \in x \iff y = a) \right.\right.$$
$$\left.\left. \lor \, \forall y \, (y \in x \iff (y = a \lor y = b)))))\right)\right].$$

Another way to express this is

$$\forall p \left[p \in R \implies \exists a \, \exists b \, \exists c \, \exists d \, \forall x \, ((x \in p \iff (x = c \lor x = d)) \land (x \in c \iff x = a)\right.$$
$$\left. \land \, (x \in d \iff (x = a \lor x = b)))\right]$$

Common Mistakes: There were a few common types of mistakes for this problem.

- Some people used $\iff$ instead of $\implies$ after "$p \in R$." However, this would mean that R contains *all* ordered pairs rather than just that every element of R is an ordered pair.
- Some people switched the order of $\exists a \, \exists b$ and $\forall x$ in the formula. However this would just imply that p is a set of sets of size 1 and 2, which is not quite right.
- Some people wrote something like the second formula above, but replaced one or more of the $\iff$'s with $\land$'s, for example writing $\exists x \, (x \in p \land (x = c \lor x = d))$. But this just implies p contains either c or d, not that $p = \{c, d\}$.

Common Mistakes: The following were not strictly mistakes, but seem worth pointing out.

- A few people required that there were some sets X and Y such that $R \subseteq X \times Y$. It is possible to prove in ZFC that every binary relation has this property, but it was not technically part of our definition of "binary relation" and is not necessary here.
- Many people wrote formulas that were much more complicated than necessary. For example,

some people tried to split into different cases depending on whether $a = b$. However, this is not necessary and just makes the formula harder to read.

Question 3 (5 points)

Let $0_{\mathbb{R}}$ denote the additive identity of the real numbers. Using the definition of $\mathbb{R}$ that we gave in class, list all elements in the transitive closure of $0_{\mathbb{R}}$.

Solution: Recall that, formally, $0_{\mathbb{R}} = \{q \in \mathbb{Q} \mid q < 0_{\mathbb{Q}}\}$. Also recall that we showed in class that $\operatorname{tc}(x) = x \cup \bigcup x \cup \bigcup\bigcup x \cup \ldots$. So we need to calculate $\bigcup 0_{\mathbb{R}}$, $\bigcup\bigcup 0_{\mathbb{R}}$ and so on. If we do this, we get

$$\begin{aligned}\operatorname{tc}(0_{\mathbb{R}}) =& \{q \in \mathbb{Q} \mid q < 0_{\mathbb{Q}}\}\\ &\cup \{\langle n, m\rangle \in \mathbb{Z} \times \mathbb{Z} \mid (n < 0_{\mathbb{Z}} \text{ and } m > 0_{\mathbb{Z}}) \text{ or } (n > 0_{\mathbb{Z}} \text{ and } m < 0_{\mathbb{Z}})\}\\ &\cup \{\{n\} \mid n \in \mathbb{Z} \text{ and } n \neq 0_{\mathbb{Z}}\}\\ &\cup \{\{n, m\} \mid n, m \in \mathbb{Z} \text{ and } ((n < 0_{\mathbb{Z}} \text{ and } m > 0_{\mathbb{Z}}) \text{ or } (n > 0_{\mathbb{Z}} \text{ and } m < 0_{\mathbb{Z}}))\}\\ &\cup \{n \in \mathbb{Z} \mid n \neq 0_{\mathbb{Z}}\}\\ &\cup \{\langle k, l\rangle \in \omega \times \omega \mid k \neq l\}\\ &\cup \{\{k\} \mid k \in \omega\}\\ &\cup \{\{k, l\} \mid k, l \in \omega \text{ and } k \neq l\}\\ &\cup \omega.\end{aligned}$$

A few points here deserve a bit more explanation. First, note that if q is a rational number less than 0 then q must be a set of nonzero integers. Thus the integer $0_{\mathbb{Z}}$ is *not* part of the transitive closure of $0_{\mathbb{R}}$. Next, if n is a nonzero integer then all of its elements must have the form $\langle k, l\rangle$ for $k, l \in \omega$ such that $k \neq l$. Moreover, each such pair is contained in soome nonzero integer. Finally, ω is a transitive set, so the "unfolding" process stops there.

Common Mistakes: The most common mistake was to either not realize that $0_{\mathbb{Z}}$ should not be in the transitive closure or to realize that but fail to "propagate" that realization—e.g. to not include $0_{\mathbb{Z}}$ but still include sets like $\langle k, k\rangle$ where $k \in \omega$. Some people also left out some sets—e.g. by going directly from ordered pairs of integers to integers.

Comment: When grading this question, all I was really looking for was the list of elements of $\operatorname{tc}(0_{\mathbb{R}})$, not the surrounding explanation. However, many people wrote quite a bit of explanation (up to an entire typed page). Perhaps that's because it was somewhat unclear exactly what the question was asking for or perhaps its because the solutions to the midterm exams provided explanations even for questions that didn't ask for it. In case the latter factor was responsible, let me clarify that I try to give extra explanation on the solutions to make them more helpful to read—they are not intended to be examples of what you need to actually write on the exam itself.

Examples and Constructions.

For each question in this section, provide the example or construction requested. You do not need to provide any justification for your answers.

Question 4 (8 points)

One day you meet some aliens from the planet Orbifulx, and you learn that everything in their development of mathematics revolves around circles. Therefore, rather than either ordered or unordered pairs, they prefer to consider triplets of objects arranged in a circle, which we will refer to as "circular triples." Two circular triples are considered the same if one can be rotated (but not reflected) so that it is equal to the other. Your task is to convince the Orbifulxians that their concept of "circular triples" can be constructed within set theory.

To be more precise, let $\circ(x, y, z)\circ$ denote the circular triple consisting of x, y, and z arranged in a circle so that if we go around the circle in clockwise order, starting from x, we will encounter x, y and z in that order. So $\circ(x, y, z)\circ = \circ(y, z, x)\circ$ no matter what x, y and z are, but $\circ(x, y, z)\circ = \circ(y, x, z)\circ$ if and only if at least two of x, y, z are equal. Explain how to define $\circ(x, y, z)\circ$ for all sets x, y, z such that it will behave as the Orbifulxians expect.

Solution: There were two main valid solutions.

Solution 1. Define $\circ(x, y, z)\circ = \{\langle x, y, z\rangle, \langle y, z, x\rangle, \langle z, x, y\rangle\}$.

Solution 2. Define $\circ(x, y, z)\circ = \{\langle x, y\rangle, \langle y, z\rangle, \langle z, x\rangle\}$.

Common Mistakes: Several people tried to define $\circ(x, y, z)\circ$ as an equivalence class in a certain equivalence relation. The main problem with this is that without further justification, it does not look like a valid construction of a set within ZFC. In fact, if you define an equivalence relation on ordered triples in the obvious way then the equivalence classes actually will be sets, and in fact they will be exactly the sets defined in the first solution above. However, you should not assume this fact is obvious to the Orbifulxians.

There is actually a second technical issue with this approach: the definition of "equivalence relation" that we gave in class technically required equivalence relations to be sets, not proper classes. We talked informally about things like cardinality as being an equivalence relation, but we never made this formal.

For these reasons, solutions that tried to define an equivalence relation on ordered triples were not given full credit (though unless other mistakes were made, such solutions were given close to full credit).

Question 5 (8 points)

Suppose F is a class function defined on the ordinals and α is an ordinal. Say that F *stabilizes at* α if α is the least ordinal such that for all $\beta > \alpha$, $F(\beta) = F(\alpha)$. Give an example of a class function $F\colon \mathbf{Ord} \to \{0, 1\}$ which does not stabilize at any ordinal.

Solution: There are many possible valid solutions. Probably the simplest one is to define

$$F(\alpha) = \begin{cases} 0 & \text{if } \alpha = 0 \text{ or } \alpha \text{ is a limit ordinal} \\ 1 & \text{if } \alpha \text{ is a successor ordinal.} \end{cases}$$

Question 6 (8 points)

Recall the definition of *stabilizes at* from the previous question. Given any function $g\colon \mathcal{P}(\omega) \to \mathcal{P}(\omega)$, define a class function $F_g\colon \mathbf{Ord} \to \mathcal{P}(\omega)$ by transfinite recursion as follows:

$$\begin{aligned} \text{Zero case:} \quad & F_g(0) = \omega \\ \text{Successor case:} \quad & F_g(\alpha+1) = g(F_g(\alpha)) \\ \text{Limit case:} \quad & F_g(\beta) = \bigcap_{\alpha<\beta} F_g(\alpha). \end{aligned}$$

Give an example of a function $g\colon \mathcal{P}(\omega) \to \mathcal{P}(\omega)$ such that F_g stabilizes at ω^2 and for all $A \subseteq \omega$, $g(A) \subseteq A$.

Solution: There are many valid examples here, though you have to be a little bit careful (as I unfortunately found out when writing midterm 2).

Here's one example that works. First, for any set $\varnothing \neq A \subseteq \omega \setminus \{0,1\}$, define $p(A)$ to be the least prime number which divides some element of A. Second, for any set $\varnothing \neq A \subseteq \omega \setminus \{0,1\}$, define $n(A)$ to be the least element of A which is divisible by $p(A)$. Now define $g\colon \mathcal{P}(\omega) \to \mathcal{P}(\omega)$ by

$$g(A) = \begin{cases} A \setminus \{0,1\} & \text{if } 0 \in A \text{ or } 1 \in A \\ A \setminus \{n(A)\} & \text{if } \varnothing \neq A \subseteq \omega \setminus \{0,1\} \\ \varnothing & \text{if } A = \varnothing. \end{cases}$$

It is clear that for all $A \subseteq \omega$, $g(A) \subseteq A$. Now note that we have

$$\begin{aligned} F_g(0) &= \{0,1,2,3,4,5,6,7,8,\ldots\} \\ F_g(1) &= \{2,3,4,5,6,7,8,\ldots\} \\ F_g(2) &= \{3,4,5,6,7,8,\ldots\} \\ F_g(3) &= \{3,5,6,7,8,\ldots\} \\ &\vdots \\ F_g(\omega) &= \{3,5,7,9,11,\ldots\} \\ F_g(\omega+1) &= \{5,7,9,11,\ldots\} \\ F_g(\omega+2) &= \{5,7,11,\ldots\} \\ &\vdots \end{aligned}$$

Hopefully the pattern is clear. After ω steps, all multiples of 2 have been removed. After $\omega \times 2$ steps, all multiples of 3 have been removed. More generally, after $\omega \times n$ steps, all multiples of the n^{th} smallest prime number have been removed. Thus, after $\omega \times \omega = \omega^2$ steps, all numbers have been removed and F_g stabilizes to $\varnothing$.

Comment: There is another sort of example which may feel somehow like "cheating" but which does not violate any of the constraints of the problem and which is a completely valid solution. Namely, pick some bijection $p\colon \omega \to \omega \times \omega$. Then define $g\colon \mathcal{P}(\omega) \to \mathcal{P}(\omega)$ as follows

$$g(A) = \begin{cases} \varnothing & \text{if } A = \varnothing \\ A \setminus \{p^{-1}(\alpha)\} & \text{if } A \neq \varnothing \text{ and } \alpha \text{ is the least element of } p[A] \end{cases}$$

where "least element of $p[A]$" refers to the standard ordering on $\omega \times \omega$ which we defined in class (i.e. the reverse lexicographic ordering). The idea is basically just to think of A as a subset of $\omega \times \omega$ and ensure that in step α of the transfinite recursion, we will remove α from the set.

Common Mistakes: Some solutions to this question did not actually give well-defined functions on all of $\mathcal{P}(\omega)$. For example, some people tried to define $g(A)$ in terms of the least prime contained in A, or the least power of a prime. However, there are plenty of sets $A \subseteq \omega$ which do not contain any primes or do not contain any powers of primes so such definitions only make sense if they include a case for handling these sorts of A's.

Long Answer Questions.

For each question in this section, provide a complete proof.

Question 7 (12 points)

Let A denote the set of all functions $\omega \to \omega$. Let $\sim$ be the binary relation on A defined by

$$f \sim g \iff \{n \in \omega \mid f(n) \neq g(n)\} \text{ is finite.}$$

Note that $\sim$ is an equivalence relation on A (you do not need to prove this). Show that $A/{\sim}$ is uncountable.

Solution: There are a few different valid solutions.

Solution 1: Diagonalization. One solution is to use a diagonalization argument. Suppose $A/{\sim}$ is countable. Then we can enumerate its elements with some sequence $C_0, C_1, C_2, \ldots$. Using the Axiom of Choice we can pick representatives $f_0 \in C_0, f_1 \in C_1, \ldots$. We will now construct a function $g\colon \omega \to \omega$ such that for all $n \in \omega$, $g \nsim f_n$. This shows that $[g]_\sim$ is not equal to any of the C_n's and thus that the C_n's do not actually enumerate $A/{\sim}$. Define g as follows

$$g(n) = 1 + \max_{i \leq n} f_i(n).$$

Now let n be any natural number. Note that for all $m \geq n$, $g(m) > f_n(m)$ and hence g and f_n disagree at infinitely many places. Thus $g \nsim f_n$, as desired.

Solution 2: Computing sizes of equivalence classes. Another solution is to show that each equivalence class in $A/{\sim}$ is countable. Let's first see why this is enough to finish the proof. By the Axiom of Choice, any countable union of countable sets is countable. Thus if $A/{\sim}$ is countable then so is its union, which is just A itself. But we have already shown in class that A is not countable.

Now let's actually show that each equivalence class in A is countable. Fix an equivalence class $C \in A/{\sim}$ and a representative $f \in C$. Recall that $\omega^{<\omega}$ denotes the set of finite sequences of natural numbers. We will show there is a surjection from $\omega^{<\omega}$ to C. Since we showed in class that $\omega^{<\omega}$ is countable, this is enough to show C is countable. The surjection can be defined as follows. For any $\sigma \in \omega^{<\omega}$ define a function $f_\sigma\colon \omega \to \omega$ by

$$f_\sigma(n) = \begin{cases} \sigma(n) & \text{if } n \in \mathrm{dom}(\sigma) \\ f(n) & \text{otherwise.} \end{cases}$$

We now claim that the map $\sigma \mapsto f_\sigma$ is a surjection from $\omega^{<\omega}$ onto C. Let g be an element of C. Since $g \sim f$, there is some largest natural number n such that $g(n) \neq f(n)$. Let $\sigma = g \restriction (n+1)$ and note that $g = f_\sigma$.

Solution 3: An injection from 2^ω. A third solution is to just directly build an injection from 2^ω into $A/{\sim}$. The basic idea of the injection is that we can map distinct elements of 2^ω to elements of A

in distinct equivalence classes by "repeating each bit infinitely many times." Let's now make this precise.

Pick a bijection $p\colon \omega \to \omega \times \omega$. Also for any $x = \langle n, m\rangle \in \omega \times \omega$, let x_0 denote the first element of x (i.e. n) and let x_1 denote the second element of x (i.e. m). Now for any $f \in 2^\omega$ define a function $\tilde{f}\colon \omega \to \omega$ as follows

$$\tilde{f}(n) = f(p(n)_0).$$

In other words, $\tilde{f}$ repeats each bit of f infinitely many times—i.e. $f(0)$ is repeated at $\tilde{f}(p^{-1}(\langle 0, 0\rangle))$, $\tilde{f}(p^{-1}(\langle 0, 1\rangle))$, $\tilde{f}(p^{-1}(\langle 0, 2\rangle))$ and so on.

Define a map $2^\omega \to A/\!\sim$ by $f \mapsto [\tilde{f}]_\sim$. We claim this map is injective. To see why, suppose $f \neq g$ are elements of 2^ω. Let $n \in \omega$ be such that $f(n) \neq g(n)$. Thus for all $m \in \omega$, $\tilde{f}(p^{-1}(\langle n, m\rangle)) = f(n) \neq g(n) = \tilde{g}(p^{-1}(\langle n, m\rangle))$. The point is that $\tilde{f}$ and $\tilde{g}$ disagree at infinitely many points and therefore $\tilde{f} \nsim \tilde{g}$. Since $\tilde{f} \nsim \tilde{g}$, we have $[\tilde{f}]_\sim \neq [\tilde{g}]_\sim$ as desired.

Comment: There are a couple interesting points about the three proofs above.

First, the first two proofs require the Axiom of Choice, but the third one does not.

Second, the third proof actually proves more than the first two. The first two just show that $A/\!\sim$ is not countable, but the third shows that its cardinality is at least $|2^\omega|$ (this is only implied by uncountability if you assume the continuum hypothesis). This raises a question: what is the exact cardinality of $|A/\!\sim|$? Using the Axiom of Choice, it is easy to show it is exactly $|2^\omega| = |\omega^\omega|$—there is an obvious surjection from ω^ω to $A/\!\sim$ and using Choice this can be converted to an injection going the other way. Thus we have $|2^\omega| \leq |A/\!\sim| \leq |\omega^\omega| = |2^\omega|$ so by the Cantor-Schroeder-Bernstein theorem, they are equal. Without the Axiom of Choice, the situation is different: it is actually not possible to show that $|A/\!\sim| \leq |2^\omega|$. This demonstrates a rather bizarre possibility when working without the Axiom of Choice. Namely, it is possible that a quotient of a set has strictly larger cardinality than the set itself.

Question 8 (10 points)

Without using the Axiom of Choice, show that for any infinite sets A and B, $|A \times B| \leq |A^B|$.

Solution: We need to show that there is an injective function $A \times B \to A^B$. In other words, we need to explain how to associate a function $B \to A$ to each pair $\langle a, b\rangle \in A \times B$ in such a way that distinct pairs are associated with distinct functions. One natural way to do this is to send the pair $\langle a, b\rangle$ to a function which takes value a at b and is constant (and not equal to a) everywhere else. Since B is infinite, if we are given such a function then we can tell what a and b are as follows: first look for an element of the domain of the function whose image is different from all other elements of the domain (that element will be b) and then check where that element gets mapped to (that will be a). The only problem with that for each $a \in A$, we need to pick some value of A not equal to a. Superficially, this seems to require the Axiom of Choice. However, if we are more careful then it can actually be done without using Choice.

Let's now do this more carefully. Since A is infinite, it has at least two elements. So we may pick $a_0, a_1 \in A$ such that $a_0 \neq a_1$. Now for any $\langle a, b\rangle \in A \times B$, define a function $f_{a,b}\colon B \to A$ as follows

$$f_{a,b}(x) = \begin{cases} a & \text{if } x = b \\ a_0 & \text{if } x \neq b \text{ and } a \neq a_0 \\ a_1 & \text{if } x \neq b \text{ and } a = a_0. \end{cases}$$

We now claim that the map $\langle a, b\rangle \mapsto f_{a,b}$ is an injective function from $A \times B$ to A^B.

To see why, let $\langle a, b\rangle, \langle a', b'\rangle \in A \times B$ such that $f_{a,b} = f_{a',b'}$. Since B is infinite, we can find some $d \in B$ which is equal to neither of b and b'. We now split into two cases.

Case 1: $a = a_0$ or $a' = a_0$. Without loss of generality, assume $a = a_0$. In this case, $f_{a',b'}(d) = f_{a,b}(d) = a_1$. Since $d \neq b'$, this implies $a' = a_0$ and hence $a = a'$. Since $a_0 = f_{a,b}(b) = f_{a',b'}(b)$ and $f_{a',b'}$ is only equal to a_0 at b', we must have $b = b'$. Thus $\langle a, b\rangle = \langle a', b'\rangle$.

Case 2: $a \neq a_0$ and $a' \neq a_0$. In this case, $f_{a,b}(b') = f_{a',b'}(b') = a' \neq a_0$. But since $a \neq a_0$, there is only one place where $f_{a,b}$ is not equal to a_0, namely at b. Thus $b = b'$. Therefore $a = f_{a,b}(b) = f_{a,b}(b') = f_{a',b'}(b') = a'$. So $\langle a, b\rangle = \langle a', b'\rangle$ and we are done.

Common Mistakes: There were two main types of mistakes in this problem.

First, some people used essentially the construction above, but instead of using the trick with a_0 and a_1 to find some element of A not equal to a, they simply said something like "pick some element of A not equal to a." The problem is that this is implicitly using the Axiom of Choice, which the problem explicitly forbids. The point of picking a_0 and a_1 in advance is that it allows us to explicitly define which element of A we should use in each case and since we have only had to make a finite number of arbitrary choices (our choice of a_0 and a_1) rather than infinitely many, we do not have to use the Axiom of Choice.

Second, some people gave an alternate construction that was not injective. For example, some people tried to define $f_{a,b}$ as follows. First, pick some $a_0 \in A$ and then define

$$f_{a,b}(x) = \begin{cases} a & \text{if } x = b \\ a_0 & \text{otherwise.} \end{cases}$$

The trouble with this definition is that for all $b, b' \in B$, $f_{a_0,b} = f_{a_0,b'}$. By the way, this particular construction should seem a bit suspicious immediately because it never seems to use the fact that $|A| \geq 2$, even though the claim in the problem statement fails if $|A| = 1$. Other people tried more complex variations on this idea which failed to be injective for a number of different reasons.

Common Mistakes: One *uncommon* mistake was to define an injection $A \times B \to A^B$ using some injection $\omega \to A$. Unfortunately, however, the fact that $|\omega| \leq |A|$ for every infinite set A is not provable without the Axiom of Choice so this approach does not work.

Comment: This problem is quite similar to an optional homework problem from lecture 23, which asked you to show without Choice that for all infinite sets A and B, $|A \sqcup B| \leq |A \times B|$. If the problem above appealed to you then I recommend you spend a few moments thinking about this one as well. It's not quite as obvious as it at first appears.

Question 9 (12 points)

Show that the Axiom schema of Separation is not provable from the Axioms of Extensionality, Empty Set and Powerset. In other words, show that there is some instance of the Axiom of Separation that is not provable from those three Axioms.

Solution: In class, we only learned one technique to prove independence results in set theory. Namely, if you want to prove that one set of sentences T cannot prove some other sentence φ, then find a set which is a model of every sentence in T but which is not a model of φ. So that's what we'll do here. In particular, we will show that ω is a model of the Axioms of Extensionality, Empty Set and Powerset, but is not a model of some instance of Separation.

Extensionality: We showed in class that every transitive set is a model of the Axiom of Extensionality and we also showed that ω is transitive.

Empty Set: We showed in class that every nonempty set is a model of the Axiom of Empty Set.

Powerset: We claim that for all $n \in \omega$, $\omega \vDash \mathcal{P}(n) = (n+1)$. It is worth pointing out here that $n+1$ is not *really* the powerset of n—it is just the case that it looks like it is from the perspective of ω (basically because ω is missing most subsets of n). To see why this is true, recall that the formal definition of "$x = \mathcal{P}(n)$" is

$$\forall y\,(y \in x \iff \forall z\,(z \in y \implies z \in n)).$$

Relativized to ω, this becomes

$$\forall y \in \omega\,(y \in x \iff \forall z \in \omega\,(z \in y \implies z \in n)).$$

Note that for $m \in \omega$, if $m \leq n$ then $m \subseteq n$ and hence $\forall z \in \omega(z \in m \implies z \in n)$ and if $m > n$ then $n \in m \cap \omega$ but $n \notin n \cap \omega$. Therefore, the only sets in ω which satisfy the final clause in the formula above are $0, 1, \ldots, n$. And since $n+1$ is exactly the set containing all of these, it will satisfy the formula above.

Failure of Separation: Let $\phi(x, y)$ denote the formula $x = y$. We will show that ω fails to model the instance of Separation using the formula ϕ with parameter 1 applied to the set $2 = \{0, 1\}$. This instance of Separation states the following

$$\exists z\,\forall x(x \in z \iff (x \in 2 \wedge x = 1)).$$

Relativized to ω, this becomes

$$\exists z \in \omega\,\forall x \in \omega(x \in z \iff (x \in 2 \wedge x = 1)).$$

In other words, it asserts that there is some $z \in \omega$ such that $z \cap \omega = \{1\}$. However, since every element of ω that contains 1 also contains 0, this statement is false.

Comment: You might have noticed that deciding what sentences are true according to some model M can require a lot of care and attention to detail. In general, the assertion that $M \vDash \phi$ may be quite different than the assertion that ϕ itself holds. In addition, determining whether $M \vDash \phi$ seems to require writing out ϕ in the language of set theory, which can result in some rather long formulas. So how do set theorists handle this sort of complexity when proving independence results? There are two, complementary, answers.

First, set theorists have developed a lot of intuition for what it means for various common formulas to hold or not hold in a model M and which statements hold of M if and only if they hold "in the real world." Second, this intuition is supplemented with a number of precisely stated theorems stating that for certain sorts of models M and certain sorts of formulas ϕ, $M \vDash \phi$ if and only if ϕ holds. Such theorems are typically referred to as "absoluteness theorems."

Here's one example of such a theorem. Suppose M is a transitive set and ϕ is a formula with parameters from M such that all quantifiers in ϕ have the form $\forall x \in y$ or $\exists x \in y$ (such quantifiers are called *bounded*). Then $M \vDash \phi$ if and only if ϕ holds. If you look back at the proof we gave in class for the independence of the Axiom of Replacement, you'll see that applying this theorem could have simplified several parts of the proof.

Common Mistakes: Several people gave proofs that made incorrect assumptions about when a model "believes" some fact about a set. For example, it is tempting to assume that if $y = \mathcal{P}(x)$ and $x, y \in M$ then $M \vDash$ "$y = \mathcal{P}(x)$" but this is false in general (the problem is that there may be some sets which M "thinks" are subsets of x but which are actually not subsets of x and hence not contained in y). Thus to show that M is a model of the Axiom of Powerset, it is not enough to check that M is closed under the powerset operation.

Common Mistakes: Several people described sets M that either were not models of the Axiom of Powerset or were models of the Axiom of Separation (or both). In one case, someone gave a set M which was not a model of Extensionality or Powerset, but was a model of Separation.

One common response of this sort was to define $M = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \ldots\}$. However, this is not a model of the Axiom of Powerset. To see why, note that $M \vDash \varnothing \subseteq \{\varnothing\}$ and $M \vDash \{\varnothing\} \subseteq \{\varnothing\}$. Thus any set which M believes is the powerset of $\{\varnothing\}$ must contain both $\varnothing$ and $\{\varnothing\}$, but M does not contain any such set. This also *is* a model of the Axiom of Separation (as a small exercise, try figuring out why that is).

Common Mistakes: Several people also tried to give heuristic arguments for why Separation is not provable from Extensionality, Empty Set and Powerset. For example, if you start with the empty set and iterate taking the powerset then you can only ever get bigger sets, not smaller subsets of sets you already have. Unfortunately, these arguments, while perhaps persuasive at an informal level, do not rigorously prove anything. They can be made into formal mathematical arguments by building some set M which makes precise the informal argument that "by iterating powerset you can never capture smaller subsets of sets you already have."

Question 10 (16 points)

Recall that a *graph* is simply a pair consisting of a set V, called the *set of vertices*, and a set $E \subseteq V \times V$, called the *set of edges*. If x is a set, the *membership graph* of x is the graph whose set of vertices is x and whose set of edges is the set $\{\langle y, z\rangle \mid y, z \in x \text{ and } y \in z\}$. A graph is *realizable* if it is isomorphic to the membership graph of some transitive set.

A graph is *rankable* if there is some way of assigning ordinals to the vertices such that ordinal assignments always increase along edges—i.e. $G = \langle V, E\rangle$ is rankable if there is some function $f\colon V \to \mathbf{Ord}$ such that for all edges $\langle u, v\rangle \in E$, $f(u) < f(v)$. A graph $G = \langle V, E\rangle$ is *rigid* if for all u, v in V,

$$\{w \in V \mid \langle w, u\rangle \in E\} = \{w \in V \mid \langle w, v\rangle \in E\} \implies u = v.$$

Prove that a graph is realizable if and only if it is rankable and rigid.

Solution: ($\implies$) Suppose $G = \langle V, E\rangle$ is a realizable graph. So there is some transitive set x and some bijection $f\colon V \to x$ such that for all $u, v \in V$, $\langle u, v\rangle \in E$ if and only if $f(u) \in f(v)$. We need to show G is rankable and rigid.

To see that G is rankable, consider the function $g\colon V \to \mathbf{Ord}$ defined by $g(v) = \operatorname{rank}(f(v))$. Note that if $\langle u, v\rangle \in E$ then $f(u) \in f(v)$ and hence $\operatorname{rank}(f(u)) < \operatorname{rank}(f(v))$. Thus g increases along any edge in E.

To see that G is rigid, suppose we have $u, v \in V$ such that $\{w \in V \mid \langle w, u\rangle \in E\} = \{w \in V \mid \langle w, v\rangle \in E\}$. Since f is an isomorphism, this implies that $\{y \in x \mid y \in f(u)\} = \{y \in x \mid y \in f(v)\}$ or, in other words, that $f(u) \cap x = f(v) \cap x$. But since x is transitive and $f(u), f(v) \in x$, we have $f(u), f(v) \subseteq x$ and thus $f(u) = f(u) \cap x = f(v) \cap x = f(v)$. And since f is a bijection, $f(u) = f(v)$ implies that $u = v$.

($\Longleftarrow$) Suppose $G = \langle V, E\rangle$ is a graph which is both rankable and rigid and let $f\colon V \to \mathbf{Ord}$ be a ranking function. By the Axiom of Replacement, we can assume that the codomain of f is actually some fixed ordinal, α and that f is an actual function, rather than a class function. The idea now is to assign sets to vertices in V by transfinite recursion, on step β assigning sets to all vertices v such that $f(v) = \beta$ (none of these vertices can be neighbors of each other, so it is okay to assign all of them simultaneously). At each step, there is actually one valid choice of which set to assign to each vertex (if we wish to end up mapping G to the membership graph on a transitive set) and rigidity will guarantee that the assignment is injective.

More formally, we will use transfinite recursion to define a function $\beta \mapsto g_\beta$ such that for each $\beta < \alpha$, g_β is a function with domain $\{v \in V \mid f(v) = \beta\}$. We can then stitch all of these functions together to get a single function defined on all of V. So suppose that we have already defined g_γ for all $\gamma < \beta$. Then we define g_β as follows.

$$g_\beta(v) = \{g_{f(u)}(u) \mid f(u) < \beta \text{ and } \langle u, v\rangle \in E\}.$$

Note that by the Axiom of Replacement, g_β is really a function (rather than a class function). By the way, if you wanted to do this transfinite induction completely formally then you would have to define g_β to be the unique function defined as above or $\perp$ if no such function exists and then prove by transfinite induction that such a function does exist at every step of the transfinite recursion.

Now define a function $g = \bigcup_{\beta<\alpha} g_\beta$. Note that by the Axiom of Replacement, g is actually a set, not a class. Also note that since each $v \in V$ is only in the domain of a single g_β, g really is a function, rather than just a binary relation. Furthermore, for any $v \in V$, we have $g(v) = g_{f(v)}(v)$. Similar remarks apply to $\bigcup_{\gamma\le\beta} g_\gamma$ for any $\beta < \alpha$. We will now establish several properties of the function g.

Claim 1. $\operatorname{range}(g)$ is transitive.

Proof. Suppose $y \in x \in \operatorname{range}(g)$. Thus for some $v \in V$, $x = g_{f(v)}(v)$. And by definition of $g_{f(v)}(v)$, this means there is some $u \in V$ such that $y = g_{f(u)}(u)$. Thus $y \in \operatorname{range}(g)$ and so $\operatorname{range}(g)$ is transitive. □

Claim 2. g is injective.

Proof. We will show by transfinite recursion that for each $\beta < \alpha$, $\bigcup_{\gamma\le\beta} g_\gamma$ is injective. Fix $\beta < \alpha$ and suppose for induction that this holds for each $\gamma < \beta$. Let $u \ne v$ be vertices in the domain of $\bigcup_{\gamma\le\beta} g_\gamma$ (so in particular, $f(u), f(v) \le \beta$). By rigidity, $u \ne v$ implies that for some $w \in V$ either $\langle w, u\rangle \in E$ and $\langle w, v\rangle \notin E$ or vice-versa. Without loss of generality, suppose that the former case holds. Since $f(w) < f(u)$ we have $g_{f(w)}(w) \in g_{f(u)}(u)$. We now claim that $g_{f(w)}(w) \notin g_{f(v)}(v)$, which implies that $g_{f(u)}(u) \ne g_{f(v)}(v)$.

Suppose not. Thus there is some $z \in V$ such that $\langle z, v\rangle \in E$ and $g_{f(z)}(z) = g_{f(w)}(w)$. Note that both $f(z)$ and $f(w)$ must be strictly less than β and thus by the inductive hypothesis applied to $\max(f(z), f(w))$, $z = w$. But this implies that $\langle w, v\rangle \in E$, which contradicts our choice of w. □

Claim 3. g is an isomorphism from G to the membership graph of range(g).

Proof. Every function is a surjection onto its range and we have already shown that g is injective. Thus it remains to show that for all $u, v \in V$, $\langle u, v\rangle \in E$ if and only if $g(u) \in g(v)$.

The forward direction is clear: if $\langle u, v\rangle \in E$ then $f(u) < f(v)$ and so by definition of g, $g(u) = g_{f(u)}(u) \in g(v) = g_{f(v)}(v)$.

For the backwards direction, suppose that $g(u) \in g(v)$. Then by definition of g there must be some $w \in V$ such that $\langle w, v\rangle \in E$ and $g(u) = g(w)$. But since g is injective, this implies that $u = w$ and thus that $\langle u, v\rangle \in E$. □

Comment: The reverse direction of this equivalence is essentially a construction called the "Mostowski collapse" of a well-founded directed graph. It is used frequently in certain parts of set theory, often so automatically that its use may not even be mentioned.

Comment: It is possible to show that if G is a rankable, rigid graph then not only is it isomorphic to the membership graph of a transitive set, that transitive set and the isomorphism are both unique. This can be shown by using transfinite induction to show that if h is an isomorphism from G to the membership graph of a transitive set then for every v in V, $h(v)$ is equal to $g(v)$ as defined above.

Comment: I initially intended to use the following, slightly different version of this problem. Suppose you define "realizable" as "isomorphic to the membership graph of *any* set (not necessarily transitive)." Then it is possible to show that a graph is realizable in this sense if and only if it is rankable (with no rigidity requirement). The proof is essentially the same except to ensure that g remains injective you have to add a unique extra element to each $g(v)$. Moreover, you cannot add these extra elements at the end of the construction—they need to be added during the transfinite recursion so that $g(v)$ does not change after you put it into other $g(u)$'s. Making sure that these extra elements really are distinct, not just from each other but from all the $g(v)$'s (some of which have not even been built yet) can be somewhat tricky. Thus I decided that version of the problem, despite having a simpler statement, was not a good choice for the exam.

Common Mistakes: Several people claimed that the ranking function $f\colon V \to \mathbf{Ord}$ on G already is an isomorphism from G to the membership graph of a transitive set. There are several problems with this.

- The range of f may not even be transitive. The problem is that f may "skip" some ordinals and thus its range may be not an ordinal itself, but rather a proper subset of an ordinal. As an example, consider the membership graph on ω. A valid ranking function on this graph is to assign the ordinal $2n$ to the vertex n. But the range of this ranking function is not all of ω, but just the set of even natural numbers.
- f might not be injective. The only requirement on f is that ordinals increase along edges.

But if there are no edges between two vertices then they might well be assigned the same ordinal. As an example, consider the membership graph on V_3 where the ordinal assigned to each element is its rank. Then $\{\varnothing, \{\varnothing\}\}$ and $\{\{\varnothing\}\}$ are assigned the same ordinal despite not being equal.

- G might not look like a line. Note that if G is isomorphic to the membership graph of an ordinal then for every pair of vertices u, v in G, either $\langle u, v\rangle \in E$ or $\langle v, u\rangle \in G$. But this is certainly not true of all rankable rigid graphs. Another way to think about this is to note that the only condition we have on f is that $\langle u, v\rangle \in E$ implies that $f(u) < f(v)$, but this does not imply that if $f(u) < f(v)$ then $\langle u, v\rangle \in E$.

Some people tried to modify the function f in some way such that it is injective and doesn't skip ordinals. However, the third point above is still a problem. And in fact, it's simply not true that every rankable, rigid graph is isomorphic to the membership graph of an ordinal, so this approach is fundamentally unworkable.

Common Mistakes: To show that a realizable graph is rankable, some people tried to define a ranking function using some version of transfinite recursion. The problem with this approach is that it's not clear what well-order is being used to carry out the transfinite recursion.

It is actually possible to get this approach to work, but it requires some extra work plus some facts that we did not cover in class. Let me explain. It turns out that it is possible to extend transfinite recursion to work on more general structures than just well-orders. In particular, transfinite induction and recursion work on any well-founded partial order (which is a partial order where every set has a minimal element). The transitive closure (in the sense of binary relations) of the membership relation on a transitive set can be shown to be a well-founded partial order and thus it is possible to define functions by transfinite recursion on this partial order. However, we did not cover transfinite recursion on well-founded partial orders in class so without further justification this approach was not given full credit.

Common Mistakes: A number of people gave proofs that realizable implies rigid that never mentioned transitivity. But this is an important part of the proof and the implication does not hold without it—there are non-transitive sets whose membership graphs are not rigid. Therefore such solutions did not receive full credit.

Extra Credit Questions.

The following question is optional. If you find a correct solution you will receive two points of extra credit.

Question 11 (2 points (bonus))

Suppose $H\colon \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a continuous function and $f : \mathbb{R} \to \mathbb{R}$ is a function (not necessarily continuous) such that for all $x \in \mathbb{R}$,

$$|f(x) - x| \leq H(x) - H(f(x)).$$

Intuitively, you can think of H as the "potential energy" of a point and the equation above says that f can only move x very far if it decreases the potential energy of x a lot. Show that there is some $x \in \mathbb{R}$ such that $f(x) = x$.

Hint: There is a way to solve this question that uses ideas we learned in class.

Solution: Every fixed point theorem in mathematics is proved by iterating the function. This one is no different, except that the iteration required is of transfinite length.

Recall that $\aleph_1$ is the first uncountable ordinal. Define a sequence $\{x_\alpha\}_{\alpha\in\aleph_1}$ of points in $\mathbb{R}$ by transfinite recursion as follows.

$$\begin{aligned} \text{Zero case:} \quad & x_0 = 0 \\ \text{Successor case:} \quad & x_{\alpha+1} = f(x_\alpha) \\ \text{Limit case:} \quad & x_\alpha = \lim_{\beta<\alpha} x_\beta. \end{aligned}$$

The limit case of this definition deserves more explanation since it may not be clear how to define such a limit or why it exists. To explain, let's first consider what happens at ω. First note that by the properties of f and H, we have

$$H(x_0) \geq H(x_1) \geq H(x_2) \geq \ldots$$

Since $H(x_0), H(x_1), H(x_2), \ldots$ is a decreasing sequence bounded below (by 0), it converges to some value, h_ω. Moreover, we can show by induction that for all $n, m \in \omega$,

$$|f(x_n) - f(x_m)| \leq H(x_n) - H(x_m) \leq H(x_n) - h_\omega.$$

Since $H(x_n)$ converges to h_ω, this implies that the sequence $x_0, x_1, x_2, \ldots$ is a Cauchy sequence and so it converges to some value, y. We then set x_ω equal to y. Note that by continuity of H, $H(x_\omega) = h_\omega$ and thus for all $n < \omega$, $H(x_n) \geq H(x_\omega)$. Moreover, we can also use continuity of H to show that for any $n < \omega$, we have

$$|f(x_n) - f(x_\omega)| \leq \sup_{m\geq n} |f(x_n) - f(x_m)| \leq \sup_{m\geq n} (H(x_n) - H(x_m)) = H(x_n) - H(x_\omega).$$

The point is that if α is a countable limit ordinal and $f(x_\beta)$ and $H(x_\beta)$ are sufficiently well behaved for $\beta < \alpha$ then we can essentially repeat this same analysis again at α. More precisely, suppose that α is a limit ordinal and that for all $\beta < \gamma < \alpha$, $H(x_\beta) \geq H(x_\gamma)$ and $|f(x_\beta) - f(x_\gamma)| \leq H(x_\beta) - H(x_\gamma)$.

Now pick a sequence $\beta_0 \leq \beta_1 \leq \beta_2 \leq \ldots$ of ordinals less than α such that $\sup_{n<\omega} \beta_n = \alpha$ (it is possible to show that such a sequence exists for any countable limit ordinal α). We can again show that the sequence $H(x_{\beta_0}), H(x_{\beta_1}), H(x_{\beta_2}), \ldots$ converges to some value h_α and that the sequence $x_{\beta_0}, x_{\beta_1}, x_{\beta_2}, \ldots$ is Cauchy and thus converges to some x_α. And as before, continuity of H guarantees that $H(x_\alpha) = h_\alpha$ and that for any β_n, $|f(x_{\beta_n}) - f(x_\alpha)| \leq H(x_{\beta_n}) - H(x_\alpha)$. Furthermore, suppose that $\gamma < \alpha$. Thus there is some n such that $\gamma < \beta_n$. So by assumption plus what we have just noted, we have

$$H(x_\gamma) \geq H(x_{\beta_n}) \geq H(x_\alpha)$$

and

$$\begin{aligned} |f(x_\gamma) - f(x_\alpha)| &\leq |f(x_\gamma) - f(x_{\beta_n})| + |f(x_{\beta_n}) - f(x_\alpha)| \\ &\leq H(x_\gamma) - H(x_{\beta_n}) + H(x_{\beta_n}) - H(x_\alpha) \\ &= H(x_\gamma) - H(x_\alpha). \end{aligned}$$

We are now almost done. Note that if we can find some $\alpha < \aleph_1$ such that $H(x_\alpha) = H(x_{\alpha+1})$ then we are done because we have

$$|x_\alpha - f(x_\alpha)| = |x_\alpha - x_{\alpha+1}| \leq H(x_\alpha) - H(x_{\alpha+1}) = 0$$

and thus x_α is a fixed point of f.

So let's suppose that for all $\alpha < \aleph_1$, $H(x_\alpha) > H(x_{\alpha+1})$. Then for each $\alpha < \aleph_1$, pick some rational number q_α in the interval $(H(x_{\alpha+1}), H(x_\alpha))$. Since $\{H(x_\alpha)\}_{\alpha<\aleph_1}$ is a strictly decreasing sequence, all of the q_α's are distinct. Thus we have an injective map from $\aleph_1$ into $\mathbb{Q}$, which is impossible since $\mathbb{Q}$ is countable and $\aleph_1$ is not.

Comment: This is a special case of the Caristi fixed point theorem. There are other proofs known that do not rely on transfinite recursion, but they do not seem as easy or natural to me as the proof above, which is really just an elaboration of the idea that you ought to be able to find fixed points of functions by iterating the function sufficiently many times.

You may find it surprising that a theorem from real analysis uses transfinite recursion, but Cantor's original motivation for developing the theory of ordinals and transfinite recursion was actually a problem in Fourier analysis. In a funny coincidence, Paul Cohen, who also revolutionized set theory, also worked on Fourier analysis.

Comment: A funny point is that the fixed point theorem for continuous, order preserving functions on the ordinals (which was part of Homework 6's "long" question), which one might expect to involve transfinite recursion, can be proved with a length ω iteration, while the Caristi fixed point theorem, which is a theorem about real numbers and not ordinals, requires a transfinite iteration.

Common Mistakes: Two people solved this problem correctly. A number of people also submitted incorrect solutions. By far the most common error was the following. It is tempting to think that we can do something similar to the solution above, but stop after only ω many steps. And indeed, this is how several fixed point theorems are proved, such as the Banach fixed point theorem, a.k.a. the contraction mapping theorem. However, in this case it does not work. The problem is that since f is not guaranteed to be continuous, it may not be the case that $f(x_\omega) = \lim_{n<\omega} f(x_n)$ and thus we have no easy way to calculate x_ω.

Here's a concrete example to show why this fails. Suppose H is the function defined by $H(x) = |x|$ and f is defined as follows

$$f(x) = \begin{cases} 1 + \frac{1}{n+1} & \text{if } x = 1 + \frac{1}{n} \text{ for some } n \in \mathbb{N} \setminus \{0\} \\ 0 & \text{if } x = 1 \\ x & \text{otherwise.} \end{cases}$$

It is easy to verify that f and H obey the conditions in the problem statement. But if we set $x_0 = 2$ and start iterating, we will get

$$\begin{aligned} x_0 &= 2 \\ x_1 &= 1 + 1/2 \\ x_2 &= 1 + 1/3 \\ &\vdots \\ x_n &= 1 + 1/(n+1) \\ &\vdots \end{aligned}$$

Taking the limit of these we get $x_\omega = 1$. But 1 is not a fixed point of f—indeed, $f(1) = 0$.