HAT PROBLEMS AND THE AXIOM OF CHOICE

GLENN SUN*
UCLA MATH CIRCLE ADVANCED 1

NOVEMBER 20, 2022

Before we introduce the main event, try your hand at the following classic brain teasers!

Problem 1. There are 2 people who are each wearing a hat. Each hat is either red or blue. The goal is to guess the color of your own hat. All players must announce their color simultaneously. Talking is not allowed, but the players can agree on a strategy beforehand.

Can you come up with a strategy in which at least 1 person always guesses correctly?

Problem 2. The setup is the same, except there are now n people and instead of a color, the hats are labeled with integers $0, \ldots, n-1$. (As before, some labels may repeat and others may not be appear at all.) Can you come up with a strategy in which at least 1 person always guesses correctly? (Hint: Say that person i is wearing x_i. Consider $x_0 + \cdots + x_{n-1} \bmod n$.)

*Adapted from an earlier Math Circle worksheet, *Hat Problems and the Axiom of Choice* by Aaron Anderson and Glenn Sun.

Now, the main question we will tackle today is: instead of n people and n labels, we have an infinite line of people whose hats are labeled by $\mathbb{N}$. Do you think it's still possible to have a strategy in which at least 1 person always guesses correctly? (We are assuming that every person can see every other person in the line, and that a person can do any number of computations in an instant. The problem is that the strategy for a finite n does not generalize to the infinite case even under these assumptions.)

If you'll excuse a short historical digression, let me tell you a brief history of Set Theory. You are probably familiar with the idea of axioms, statements that are just "obviously true" and should be accepted. Several centuries ago, each part of math used its own axioms. For example, you may be familiar with Euclid's 5 special axioms just for geometry (although as a side note, we actually need a few more than 5). In the early 1900s, a man named Zermelo introduced a list of axioms that govern set theory: what sets are and how they can be constructed. Today, we almost always define mathematical objects using sets, so this is a very foundational theory.

One of the axioms that Zermelo introduced was the Axiom of Choice. I'll get to what it says later, but in the decades after Zermelo's paper, the Axiom of Choice was highly controversial. Although it seemed necessary to prove many statements that felt "obviously true," it could also be used to prove many statements that felt "obviously false." Mathematicians even debated for a long time if it was consistent: meaning it doesn't lead to contradictions. In the mid 1930s, Gödel proved that the Axiom of Choice is consistent with the rest of Zermelo's axioms. Since then, most mathematicians have no problem using the Axiom of Choice whenever necessary, although there are efforts to see if weaker forms of the axiom are enough to prove certain theorems.

The reason I bring this up is because surprisingly, we need the Axiom of Choice to resolve our infinite hat puzzle! Let's first cover some vocabulary which will be necessary to understand what the Axiom of Choice says.

Definition 1 (equivalence relation)**.** A relation $\sim$ over a set S is something that says if any two elements of S are related ($a \sim b$) or unrelated ($a \not\sim b$). An equivalence relation is a special kind of relation that further satisfies three properties:

1. Reflexivity: for all $a \in S$, we have $a \sim a$.
2. Symmetry: for all $a, b \in S$, if $a \sim b$, then $b \sim a$.
3. Transitivity: for all $a, b, c \in S$, if $a \sim b$ and $b \sim c$, then $a \sim c$.

One example of an equivalence relation is just equality. Take any set S. Then,

1. Reflexivity: for all $a \in S$, we indeed have $a = a$.
2. Symmetry: for all $a, b \in S$, if $a = b$, then indeed $b = a$.
3. Transitivity: for all $a, b, c \in S$, if $a = b$ and $b = c$, then indeed $a = c$.

Most equivalence relations aren't this straightforward, but still fairly straightforward. One useful way to think about equivalence relations is just "kind of equal".

Problem 3. Which of the following are equivalence relations and why?

1. The relation $\equiv \pmod{n}$ over the integers $\mathbb{Z}$, where $a \equiv b \pmod{n}$ means a and b have the same remainders mod n.

2. The relation $\leq$ over the real numbers $\mathbb{R}$, where $a \leq b$ means exactly that.

3. The relation $\sim$ over triangles, where $\triangle ABC \sim \triangle A'B'C'$ means all corresponding angles are equal (i.e. $m\angle A = m\angle A'$, $m\angle B = m\angle B'$, and $m\angle C = m\angle C'$.)

4. The relation $\sim$ over pairs of integers, where $(a, b) \sim (c, d)$ means $ad = bc$.

5. The relation $\sim$ over pairs of real numbers, where $(a, b) \sim (c, d)$ means the distance from (a, b) to (c, d) is an integer (using the standard distance formula).

Definition 2 (equivalence class)**.** Given an equivalence relation $\sim$ on a set S, the equivalence class of $a \in S$ is the set $[a]_\sim = \{b \in S \mid a \sim b\}$. The set of all equivalence classes of $\sim$ is denoted $S/\sim$ (read "S modulo $\sim$").

The equivalence class of a is "everything kind of equal to a." A good picture to have is that the equivalence classes form a partition of S: every element of S belongs to exactly one of the equivalence classes.

Problem 4. For every equivalence relation in problem 3, describe $S/\sim$ in words, and determine how many equivalence classes there are.

Now, we can state the Axiom of Choice.

Definition 3 (choice function, representative element)**.** Given an equivalence relation $\sim$ on a set S, a choice function is any function $f : S/\sim \to S$ such that $f([a]_\sim) \in [a]_\sim$ for all $a \in S$. In other words, the choice function picks an element from every equivalence class. A chosen element is called the representative element of its class.

The axiom of choice says that every equivalence relation has a choice function.

In other words, there is always a way to pick a representative element from every equivalence class (even if there are infinitely many equivalence classes). In even simpler words, you can pick an object out of every box. Doesn't this sound obvious?

Problem 5. For each of the equivalence relations in problem 3, give an example of a choice function, explicitly specifying a representative element for each equivalence class.

However, this seemingly innocent axiom has major consequences, as we will now see. Remember: the Axiom of Choice says that a choice function exists, even if it's not obvious how to explicitly write it. Let's solve our infinite hat puzzle now.

Problem 6. Define an equivalence relation $\sim$ on infinite sequences of natural numbers so that two sequences are equivalent if they are eventually the same.

1. Are the sequences $27, 0, 3, 1, 4, 1, 5, 9, \ldots$ and $5, 2, 3, 1, 4, 1, 5, 9, \ldots$ (digits of pi starting at the third position) equivalent?

2. Are the sequences $0, 1, 0, 1, \ldots$ and $1, 0, 1, 0, \ldots$ equivalent?

3. Check that $\sim$ really is an equivalence relation.

Problem 7. Use the Axiom of Choice on the above equivalence relation, devise a strategy for the infinite hat problem.

Our original goal was that at least 1 person guesses correctly, but with this strategy, you should be able to get that *all but finitely many* people guess correctly!

Problem 8. (Challenge) Here's a related problem: everything is the same as before, but now, the players stand in a line and they guess their numbers in $\mathbb{N}$ in order (they hear others' guesses). Give a strategy in which *at most one* player guesses incorrectly!

Hint: Think about last week's proofs that about things being countable.

Problem 9. (Bonus) Going back to the finite case of n players and labels $0, \ldots, n-1$, but now arranged in a line and guessing in order, can you find a similar strategy to solve this modification optimally?

With these hat problems, we saw how the Axiom of Choice could lead to some very strange results: somehow, *almost all* players can guess the number on their head, even though they can't see it. This is better than the finite case! Next, we will show an instance where the Axiom of Choice is *necessary* for some very obvious-looking statements to be true.

First, let's recall some properties of functions. Recall that a function f specifies an input set (domain) X, an output space (codomain) Y, and a rule that maps each $x \in X$ to a unique output $y \in Y$. The output is denoted by $f(x)$.

Definition 4 (injective, surjective, bijective)**.** A function $f : X \to Y$ is:

$$\left.\begin{array}{r}\text{injective}\\\text{bijective}\\\text{surjective}\end{array}\right\}\text{ if for all } y \in Y\text{, there exists }\left\{\begin{array}{c}\text{at most 1}\\\text{exactly 1}\\\text{at least 1}\end{array}\right\}x \in X\text{ such that } f(x) = y.$$

We covered these last week with slightly different but equivalent definitions. This definition above emphasizes that injectivity and surjectivity are really complementary ideas.

Problem 10. Categorize the following functions as injective, surjective, bijective, or none of the above. Drawing their graphs might help!

1. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x$.

2. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$.

3. $f : \mathbb{R} \to [0, \infty)$ given by $f(x) = x^2$.

4. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x(x + 1)(x - 1)$.

5. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = e^x$.

6. $f : \mathbb{R} \to (0, \infty)$ given by $f(x) = e^x$.

Definition 5 (left inverse, right inverse, inverse)**.** For a function $f : X \to Y$, a function $g : Y \to X$ is called a

$$\left.\begin{array}{r}\text{left inverse}\\ \text{inverse}\\ \text{right inverse}\end{array}\right\} \text{ if } \begin{cases} g(f(x)) = x \text{ for all } x \in X \\ \text{both} \\ f(g(y)) = y \text{ for all } y \in Y \end{cases}.$$

Problem 11. Determine if the functions below have a left inverse, inverse, right inverse, or none. If they do, say if it's unique, and give an example using a formula.

1. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x$.

2. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$.

3. $f : \mathbb{R} \to [0, \infty)$ given by $f(x) = x^2$.

4. $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = e^x$.

Problem 12. Prove the following statements. For the first two, you won't need anything too fancy. For the third, you will need the Axiom of Choice!

1. A function has a left inverse if and only if it is injective.

2. A function has an inverse if and only if it is bijective.

3. A function has a right inverse if and only if it is surjective.

Problem 13. Assume the negation of the Axiom of Choice. That is, assume that there exists an equivalence relation $\sim$ on a set S for which no choice function exists. Give an example of a surjective function with no right inverse.

Together, these last two problems show that you really do need the Axiom of Choice to prove some pretty fundamental, obvious-sounding statements. At the same time, it produces counterintuitive strategies for infinite hat puzzles. That's what makes the Axiom of Choice fun! Now, in our last section, we'll see one more counterintuitive application of the Axiom of Choice.

The Banach–Tarski paradox is the following theorem. Published in 1924, this was notably before Gödel's proof that the Axiom of Choice is consistent with the rest of set theory, so mathematicians were at first very worried about the Axiom of Choice when they heard this theorem.

Theorem 6 (Banach–Tarski theorem)**.** It's possible to chop a solid ball into 5 pieces, rotate them around without deforming the pieces, and put them back together to make two perfect balls, each of the same volume as the original ball (with no holes).

Note that the part about not deforming pieces is important: it's very easy to do this if you allow deforming the pieces, as a ball is an uncountably infinite set of points. (Think about last week's packet.) The fact that you can do this with *rigid motions*, which one would think behave nicely, is extremely surprising! The whole theorem is a bit too much for us to prove, but we can talk about some ideas.

The Banach–Tarski paradox is possible because the Axiom of Choice lets you define sets that are so weird, that it is impossible to calculate their volume. To see what this means, let's assume for a bit that every subset of $\mathbb{R}^3$ has a well-defined volume, and that volume satisfies a few obvious axioms. Then we will show from these assumptions that the Banach-Tarski paradox is impossible. This means that if you assume the Axiom of Choice, you can't have a well-defined volume for every subset of $\mathbb{R}^3$.

Definition 7. A volume measure is a function V that maps subsets of $\mathbb{R}^3$ to $[0,\infty)\cup\{\infty\}$, satisfying the following "obviously true" facts:

- The empty set has volume 0: $V(\emptyset) = 0$.
- If $A_1, A_2, A_3, \dots \subset \mathbb{R}^3$ are pairwise disjoint sets, then $V(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} V(A_i)$. (Any sum that includes ∞ is just ∞.)
- The volume of the unit sphere is finite and positive.
- If $A \subseteq \mathbb{R}^3$ has a defined volume, and B is the set formed by moving or rotating A, then $V(A) = V(B)$.

Problem 14. Show that if $A \subset B \subset \mathbb{R}^3$ and V is a volume measure, then $V(A) \le V(B)$. (We define $x \le \infty$ for all $x \in \mathbb{R}$.)

Problem 15. The Banach–Tarski paradox says that there exist five pairwise disjoint sets, A_1, A_2, A_3, A_4, A_5, such that $A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5$ is the unit solid ball, but you can move and/or rotate each A_i to a different set B_i, such that $B_1, \ldots, B_5$ are also pairwise disjoint, and $B_1 \cup B_2 \cup B_3 \cup B_4 \cup B_5$ consists of two disjoint copies of the unit sphere.

Show that if a volume measure V exists, then this is impossible.

(Just to remind you, the conclusion from the above problem is that *volume measures don't exist*, since the Banach–Tarski paradox is a theorem. When mathematicians want to talk about volume, we have to very careful and only define volume for certain "nice" subsets of $\mathbb{R}^3$, not the crazy subsets that arise in the paradox.)

Now we will show how the Axiom of Choice implies an easier version of the Banach–Tarski paradox, in a 1-dimensional context.

Definition 8. A length measure is a function ℓ that maps subsets of $\mathbb{R}$ to $[0, \infty) \cup \{\infty\}$, satisfying the following "obviously true" facts:

- The empty set has length 0: $\ell(\emptyset) = 0$.
- If $A_1, A_2, A_3, \dots \subset \mathbb{R}^3$ are pairwise disjoint sets, then $\ell(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \ell(A_i)$. (Any sum that includes ∞ is just ∞.)
- The length of an interval $[a, b]$ is $\ell([a, b]) = b - a$.
- If $A \subseteq \mathbb{R}$ has a defined volume, and $B = \{a + r : a \in A\}$ for some $r \in \mathbb{R}$, i.e. a translation of A by r, then $V(A) = V(B)$.

We will show directly that there is a set X such that $\ell(X)$ cannot be defined, much like the 5 subsets of the sphere in the Banach-Tarski paradox. The procedure will also involve cutting up a piece of (this time 1-dimensional) space, this time into countably many pieces, moving them, and reassembling them into a bigger region of space.

Problem 16. For $x, y \in [0, 1]$, define $x \sim y$ to be true if and only if $y - x \in \mathbb{Q}$. Show that $\sim$ is an equivalence relation.

Problem 17. Using the Axiom of Choice, construct $X \subseteq [0,1]$ such that X contains exactly one representative of each equivalence class of $\sim$.

Problem 18. Prove the following.

1. Show that there is a bijection $h : \mathbb{N} \to \mathbb{Q} \cap [-1,1]$.

2. For $k \in \mathbb{N}$, let X_k be $\{x + h(k) : x \in X\}$. Show that $[0,1] \subseteq \bigcup_{k\in\mathbb{N}} X_k \subseteq [-1,2]$.

3. Show that $X_1, X_2, X_3, \ldots$ are pairwise disjoint.

Problem 19. In the following cases, compute $\ell(\bigcup_{x\in\mathbb{N}} X_k)$ and derive a contradiction. Conclude that $\ell(X)$ cannot be defined.

1. $\ell(X) = 0$.

2. $\ell(X)$ is finite and positive.

3. $\ell(X) = \infty$.

Our set X is called a *Vitali set*, and is the easiest example of a non-measurable set to construct and prove from scratch, using of course, the vital ingredient, the Axiom of Choice. This was discovered before the Banach–Tarski paradox, and Banach and Tarski were inspired by it. Unfortunately (or perhaps fortunately, if you don't like paradoxes), the Banach–Tarski result itself, of cutting up an obviously measurable set into finitely many pieces and reassembling into multiple copies of the original, or a larger copy of the original, doesn't work very well in 1 or 2 dimensions. The actual Banach–Tarski proof revolves around studying the group theory of 3-dimensional rotations and translations, to set up a very particular equivalence relation which we can then apply the Axiom of Choice to.