The Demon's Game

By Khiyali K. Pillalamarri

Submission for the MITx Philosophy Award/High School Philosophy Award
7 September 2018

Part I - An objection to Professor Rayo's treatment of the Demon's Game

In the online course *Paradox and Infinity*, Professor Agustín Rayo detailed the following paradox, which he called the Demon's Game.

There are a countably infinite number of people: P_1, P_2, P_3, etc. The people are all assumed to be "fully rational." A demon approaches this group of people, and announces that he will ask each individual person to say either *aye* or *nay*.

If a finite number *n* of players say *aye*, he will give each member of the group -- regardless of their answer -- *n* gold coins. For example, if five people answer *aye* and the rest say *nay*, every member of the group will get 5 gold coins. However, if an infinite number of players say *aye*, the demon will not give out any coins.

Professor Rayo did not explicitly specify the objective of the players. Therefore, I assume that the goal of any given person in the group is to earn as many gold coins as possible.

Professor Rayo noted that if the players were given time to confer with one another before answering, they could simply pick a number *n*, assign persons P_1 through P_n to say *aye* and assign everyone else to say *nay*. Then each player would receive *n* coins.

However, said Rayo, the demon knows that if he were to let the players make any such plan, he would lose a lot of money. Therefore, he isolates the players and asks each one individually to say *aye* or *nay*.

This, said Rayo, made the Demon's Game a paradox. His logic was as follows:

"[I]magine yourself as a member of the group. You are isolated from your colleagues, and you are pondering your answer. Should you say aye *or should you say* nay*? You know that your decision can have no effect on other people's decisions, and reason as follows:*

> *If infinitely many of my colleagues answer* aye *nobody will get any coins, regardless of what I decide to do. So my decision can only make a difference to the outcome on the assumption that at most finitely many of my colleagues answer* aye. *But in that case I should definitely answer* aye. *For doing so will result in an additional gold coin for everyone, including myself! (If I were to answer* nay*, on the other hand, I wouldn't be helping anyone.)*

So: answering aye *couldn't possibly make things worse, and could very well make things better. The rational thing to do is therefore to answer* aye*!*

Of course, other members of the group are in exactly the same situation as you. So what is rational for you is also rational for them.

That is why the demon has nothing to fear.

As long as every member of the group behaves rationally, everyone will answer aye*, and the demon won't have to cough up any money."*

I believe that Professor Rayo is incorrect in saying that one should answer *aye* whenever "at most finitely many of [one's] colleagues answer *aye*."

Now, I do agree that one should answer *aye* if zero other people do. (If one were to say *nay* instead, then there would be no *aye* answers, and no player would receive coins.) However, I argue that **if a finite, nonzero number of one's colleagues say *aye*, it is no more logical for one to answer *aye* than it is to answer *nay***. As long as at least one person says *aye*, and as long as it is not the case that an infinite number of people say *aye*, it does not matter how many people say *aye*. I will now explain why.

Suppose exactly one person says *aye*. Then, each member of the group will receive 1 coin. Thus, the group *as a whole* will get as many coins as there are people. Since there are as many people as there are natural numbers, the group as a whole will get $|\mathbb{N}|$ coins. To generalize: if there are n *ayes*, where $n \in \mathbb{N}$ and $n > 0$, the group as a whole will get $n + n + n + ... = |\mathbb{N}|$ coins.

We have determined that as long as a finite, nonzero number of people answer *aye*, the group will get $|\mathbb{N}|$ gold coins. After the demon leaves, these gold coins may be redistributed among the members of the group as follows:

Recall that the players are numbered P_1, P_2, P_3, etc. Let all of the players combine their coins into a joint pile. Then, let them execute the following steps.

1) P_1 takes a coin from the pile.
2) P_1 and P_2 each take a coin from the pile. Now P_1 has 2 coins and P_2 has 1 coin.
3) P_1, P2, and P3 each take a coin from the pile. Now P1 has 1 coin, P2 has 2 coins, and P3 has 1 coin.
4) P1-P4 take a coin each.
5) P1-P5 take a coin each.

Because this method simultaneously increases a) the number of people with coins, and b) the number of coins each person has, it will result in each person getting $|\mathbb{N}|$ gold coins -- far more than the demon offered.

Part II -- Can we get out of the paradox?

We have determined that it does not matter how many *ayes* there are as long as there are a finite, nonzero number. However, does this make a given player's decision any easier?

Suppose that it is your turn to answer *aye* or *nay*. There are three possibilities:

- When the answers are tallied, infinitely many of the other players will have said *aye*. Regardless of your own answer, you do not receive any gold coins.

- When the answers are tallied, a finite, nonzero number of the other players will have said *aye*. Regardless of your own answer, you receive (after the redistribution system described in Part I) $|\mathbb{N}|$ gold coins.

- When the answers are tallied, none of the other players will have said *aye*. In this case, your answer does make a difference. If you answer *nay* like everyone else, you will not receive gold coins. However, if you answer *aye*, you will receive $|\mathbb{N}|$ gold coins (again, after redistribution).

Therefore, it is logical to answer *aye*. Because everyone else is in the same situation as you are, all of them will say *aye* as well. Consequently, no one will receive any coins, and we will be back in the same predicament Professor Rayo described: each individual person will act logically, and yet will end up empty-handed.

I think the problem is this: what is logical for one person is logical for *every* person. Therefore, since everyone is in the same situation, everyone will give the same answer. However, for players to earn coins, different players must give different answers.

But there is no way of distinguishing between players. Though different people have different features -- for example, each person was given a number P_n at the start of this essay -- these distinguishing features are irrelevant.

This is because the players have not been able to confer with one another. If they had been given the opportunity to do so, they could have made a plan, such as "Let P_1 say *aye* and the rest of us *nay*." But because they could not do this -- because they could not assign *meanings* to distinguishing features, the distinguishing features are irrelevant. (For example, above, I apply the meaning, "If you are P_1, it means you say *aye* and if you are not P_1 it means you say *nay*," to the feature of each person having a number.)

Thus, everyone is effectively identical, and everyone is faced with the same problem -- saying *aye* or *nay*. Additionally, we assumed earlier that every player was perfectly rational.

Two rational thinkers cannot begin at the same situation, use the same rules of logic, and yet reach different conclusions.

Therefore, the players will all be forced to give the same answer, and thus fail to earn gold coins. There is no way out.