Lions and lambs—can you solve this classic game theory puzzle?

July 20 2017, by Amirlan Seksenbayev

Credit: Monstera Production from Pexels

How many lions does it take to kill a lamb? The answer isn't as straightforward as you might think. Not, at least, according to game theory.

Game theory is a branch of maths that studies and predicts decision-making. It often involves creating hypothetical scenarios, or "games", whereby a number of individuals called "players" or "agents" can choose from a defined set of actions according to a series of rules. Each action will have a "pay-off" and the aim is usually to find the maximum pay-off for each player in order to work out how they would likely behave.

This method has been used in a wide variety of subjects, including economics, biology, politics and psychology, and to help explain behaviour in auctions, voting and market competition. But game theory, thanks to its nature, has also given rise to some entertaining brain teasers.

One of the less famous of these puzzles involves working out how players will compete over resources, in this case hungry lions and a tasty lamb. A group of lions live on an island covered in grass but with no other animals. The lions are identical, perfectly rational and aware that all the others are rational. They are also aware that all the other lions are aware that all the others are rational, and so on. This mutual awareness is what's referred to as "common knowledge". It makes sure that no lion would take a chance or try to outsmart the others.

Naturally, the lions are extremely hungry but they do not attempt to fight each other because they are identical in physical strength and so would inevitably all end up dead. As they are all perfectly rational, each lion prefers a hungry life to a certain death. With no alternative, they can survive by eating an essentially unlimited supply of grass, but they would all prefer to consume something meatier.

One day, a lamb miraculously appears on the island. What an unfortunate creature it seems. Yet it actually has a chance of surviving this hell, depending on the number of lions (represented by the letter N). If any lion consumes the defenceless lamb, it will become too full to defend himself from the other lions.

Assuming that the lions cannot share, the challenge is to work out whether or not the lamb will survive depending on the value of N. Or, to put it another way, what is the best course of action for each lion – to eat the lamb or not eat the lamb – depending on how many others there are in the group.

Credit: AI-generated image (disclaimer)

The solution

This type of game theory problem, where you need to find a solution for a general value of N (where N is a positive whole number), is a good way of testing game theorists' logic and of demonstrating how backward induction works. Logical induction involves using evidence to form a conclusion that is probably true. Backward induction is a way of finding a well-defined answer to a problem by going back, step-by-step, to the very basic case, which can be solved by a simple logical argument.

In the lions game, the basic case would be N=1. If there was only one hungry lion on the island it would not hesitate to eat the lamb, since there are no other lions to compete with it.

Now let's see what happens in the case of N=2. Both lions conclude that if one of them eats the lamb and becomes too full to defend itself, it would be eaten by the other lion. As a result, neither of the two would attempt to eat the lamb and all three animals would live happily together eating grass on the island (if living a life solely dependent on the rationality of two hungry lions can be called happy).

For N=3, if any one of the lions eats the lamb (effectively becoming a defenceless lamb itself), it would reduce the game to the same scenario as for N=2, in which neither of the remaining lions will attempt to consume the newly defenceless lion. So the lion that is closest to the actual lamb, eats it and three lions remain on the island without attempting to murder each other.

And for N=4, if any of the lions eat the lamb, it would reduce the game to the N=3 scenario, which would mean that the lion that ate the lamb would end up being eaten itself. As none of the lions want that to happen, they leave the lamb alone.

Essentially, the outcome of the game is decided by the action of the lion closest to the lamb. For each integer N, the lion realises that eating the lamb would reduce the game to the case of N-1. If the N-1 case results in the survival of the lamb, the closest lion eats it. Otherwise, all the lions let the lamb live. So, following the logic back to the base case every time, we can conclude that the lamb will always be eaten when N is an odd number and will survive when N is an even number.

This article was originally published on The Conversation. Read the original article.

Provided by The Conversation