Prisoner's Dilemma: The Case of the Stolen Candies

UF Math Circle - T.Truong

February 2025

Consider the classic story of the Prisoner's Dilemma. Alice and Bob just committed a heinous crime: they stole candies from the candy jar. Mom caught them sneaking out of the kitchen, so Mom told Alice to go to her room and told Bob to go to his room. Then, Mom went to each room separately and asked, "Did your sibling steal the candies?"
The kids both knew what will happen next:

The Rules of the Game

Each child has two choices:

- **Stay Silent (Cooperate with their sibling)**
- **Blame the other (Defect)**

Mom makes the following decision based on their answers:

- If both Alice and Bob stay silent, Mom cannot prove they stole candies and gives them only a small punishment: **no dessert for one night**.
- If one blames the other while the other stays silent, the one who blamed gets **no punishment**, while the one who stayed silent receives **three days of no desserts**. For example, If Alice blamed Bob, and Bob was silent, Mom would punish Bob significantly. Mom would think that Bob is the real culprit, and he would end up with three days of no desserts. Alice, by contrast, would have no punishment, because Mom would think Alice told the truth. The same for if Bob blamed Alice. Mom would think Alice is the real culprit, and Alice would end up with three days of no desserts. Bob, by contrast, would have no punishment, because Mom would think Bob told the truth.
- If both blame each other, they both receive **a punishment**: **say two nights without dessert**.

Visualizing the Consequences

We can capture the above choices and their consequences in a simple table:

	Bob Stays Silent	**Bob Blames Alice**
Alice Stays Silent		
Alice Blames Bob		

Questions

1. If Bob and Alice could coordinate, what would they jointly decide to do?
2. When Bob is alone in the room with Mom, what will he be tempted to do
3. When Alice is alone in the room with Mom, what will she be tempted to do?
4. What do you think is disappointing about all the possible outcomes?
5. Is there a pair of choices Alice and Bob could make where neither would want to change their decision after finding out what the other did?

Definition 0.1. A strategy s is dominating if regardless of your opponent's strategy, choosing s maximizes your utility.

Definition 0.2. Nash Equilibrium is a situation where no player could gain utility by changing their own strategy.

Cases

Now consider what happens if Alice and Bob find themselves in this same situation repeatedly. Can Alice develop a strategy that will encourage Bob to be cooperative?

Case 1 Suppose Alice decides that she will always be silent, no matter what Bob does. How can we visualize this?

1. If they are caught stealing candies ten times, what is the best outcome she might achieve?
2. If they are caught stealing candies ten times, what is the worst outcome she might achieve?

Round	**Alice**	**Bob**	**Alice Gets**	**Bob Gets**
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
Total				

Case 2 A more complicated case to consider is if Alice is silent unless, in the prior round, Bob blamed her. If Bob blamed her, she gets even by blaming him in this round. This is known as the "Tit-for-Tat" strategy. Let's practice by visualizing this strategy:

Now, let's consider if Bob knows that Alice is playing tit-for-tat, what should he do in each round?

Even though Bob and Alice's dominating strategy is to blame the other person, in a multi-round prisoner's dilemma game, the optimal strategy is to play tit-for-tat and cooperate.

Round	**Alice**	**Bob**	**Alice Gets**	**Bob Gets**
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
Total				

This explains why people are evolutionarily altruistic, even though it seems like it would have a negative impact on their utility.

Round	**Alice**	**Bob**	**Alice Gets**	**Bob Gets**
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
Total				