

## The Prisoner's Dilemma

Doug Lichtman

Math Circle

5/10/20

A few weeks ago, we studied trade, which is an interaction where two parties can act in their own self-interest and yet **both** end up better off. Sadly, life does not always work that way.

Consider the classic story of the Prisoner's Dilemma.

Alice and Bob just committed a horrible crime: they stole cookies from the cookie 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 cookies?"

The kids knew what would happen next.

- If **both kids are silent**, Mom will punish each lightly. Mom knows that the cookies were stolen, but, without testimony, she will go easy on each child. Alice and Bob would thus each lose television for one day.
- 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 lost television. Alice, by contrast, would have no punishment, because Mom would think Alice told the truth.
- If **Bob blamed Alice, and Alice was silent**, the opposite situation would play out. Mom would think Alice is the real culprit, and Alice would end up with three days of lost television. Bob, by contrast, would have no punishment, because Mom would think Bob told the truth.
- Lastly, if **Bob blames Alice, and Alice blames Bob**, both kids would get in trouble. Alice and Bob would each lose two days of television.

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

	Bob Silent ↓	Bob Blame Alice ↓
ALICE Silent →	1, 1	3, 0
ALICE Blame → Bob	0, 3	2, 2

Red arrows point from the (2,2) cell to the (1,2) cell labeled "Alice" and to the (2,1) cell labeled "Bob".

### QUESTIONS

1. If Bob and Alice could coordinate, what would they jointly decide to do?

Silent & Silent

2. When Bob is alone in the room with Mom, what will he be tempted to do?

Blame!  
If Alice was silent, Bob should blame.  
If Alice blamed, Bob should blame.

3. When Alice is alone in the room with Mom, what will she be tempted to do?

Same thing in reverse: Blame!

4. In your own words, what is disappointing about this?

Both are worse off. "

$(1,1)$  is better than  $(2,2)$

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?

1. Suppose Alice decides that she will always be silent, no matter what Bob does.

1a. If they are caught stealing cookies ten times, what is the best outcome she might achieve?

10 days from ten  $(1,1)$

1b. If they are caught stealing cookies ten times, what is the worst outcome she might achieve?

30 days from ten  $(3,0)$

1c. What do you think Bob will do? Will he take the risk of being silent the first time, knowing it might turn out badly for him? And if he does, will he take the risk of being silent the second time also? What about the tenth and final time?

Hard to know if Bob will start silent.  
But in the last round, Bob will blame.

1d. What if Bob realizes that Alice's strategy is to always stay silent. What might Bob do?

Round	Alice	Bob	Alice Gets	Bob Gets
1	S	B	3	0
2	S	B	3	0
3	S	B	3	0
4	S	B	3	0
5	S	B	3	0
6	S	B	3	0
7	S	B	3	0
8	S	B	3	0
9	S	B	3	0
10	S	B	3	0
<b>TOTAL</b>			<b>30</b>	<b>0</b>

2. There are a large number of possible strategies that Alice can play. Above, we considered the possibility that she will always stay silent. A more complicated option would be for Alice to play **tit-for-tat**. In that strategy, Alice is silent unless, in the prior round, Bob blamed her. If Bob blamed her, she "gets even" by blaming him in this round.

2a. If Bob knows that Alice is playing tit-for-tat, what should he do in each round?

Round	Alice	Bob	Alice Gets	Bob Gets
1	S	S	1	1
2	S	S	1	1
3	S	S	1	1
4	S	S	1	1
5	S	S	1	1
6	S	S	1	1
7	S	S	1	1
8	S	S	1	1
9	S	S	1	1
10	S	B	3	0
<b>TOTAL</b>			<b>12</b>	<b>9</b>

2b. What other strategies should Alice consider? For instance, should Alice maybe forgive Bob if he blames her once, but after that play strict tit-for-tat? Would that be helpful?

Maybe. If she thinks Bob might just make a mistake, she might want to forgive him instead of triggering a tit-for-tat war