Durvaa Tiwari

Isn't Mathematics The Issue? : An insight into the Prisoner's Dilemma

1. Introduction

We like to believe that our decisions are our own. That when the moment comes, we'll always choose what is right, or at least what's best. But what if I told you that mathematics had predicted your decisions a long time ago? The Prisoner's Dilemma is not just a theoretical problem in mathematics, but a reflection of the invisible calculations humans perform every day. We assume that being rational leads to the best outcome. The Prisoner's Dilemma shows that this is not always true.
Sometimes logical decisions lead to worse outcomes.

2. *The Dilemma*

In game theory, the prisoner's dilemma is a thought experiment involving two rational agents. When caught and imprisoned, they are separated and unable to communicate. Each is given a choice: remain silent or testify against the other.

If they both choose to remain silent, each serves one year in prison. If one of them testifies while the other stays silent, the one who testifies is set free, while the other has to serve three years in prison. And if they both testify, both must serve two years in prison.

Two agents. Two choices. No way to communicate and one decision that can alter their lives forever.

3. The Mathematics

At first glance, we might think the best option is for both to stay silent.
But when we analyze this mathematically, our viewpoints might shift. If we assign values to each year spent in prison, we might be able to judge this situation more precisely.
For instance, we may represent the outcomes as (1,1), (3,0), (0,3), and (2,2), where lower values represent better outcomes.

	A stays silent	A testifies
B stays silent	(1,1)	(0,3)
B testifies	(3,0)	(2,2)

Now, let's consider the perspective of one agent. If the other stays silent, he can either stay silent and serve a year or betray and become free. If, however, the other agent does testify, he can either stay silent and serve 3 years in prison, or he can testify and serve 2 years instead.

In both cases, betraying seems like the better option with a higher payoff. This means that betrayal becomes the dominant strategy: the option that appears best regardless of the other agent's decision.

When both players choose to betray, neither can improve their outcome by changing their decision alone. This creates what is known as a Nash equilibrium, a situation where no player could gain more by changing their own strategy (holding all other players' strategies fixed) in a game.

The results are stable, but not optimal. They both act logically and yet end up in a worse situation than if they had just trusted each other.

Here is where the paradox emerges. If both agents remain silent, they might both get a lighter sentence. Yet, when faced with uncertainty, both of

them turn towards betrayal. Acting rationally, each chooses the option that seems the safest from an individual perspective.
The dilemma lies in the conflict between personal reasoning and collective benefit. What looks like the most sensible decision for one person can lead to an outcome that is far from ideal for all. In this way, the Prisoner's Dilemma highlights a subtle flaw in our understanding of our rationality: that making the best choice for ourselves does not always mean the best result overall.

This leads us to an uncomfortable realization about human nature. Mathematics does not fail in this situation; it shows a limitation in the way we think. It doesn't misguide us; it exposes that our idea of what is 'right' is still incomplete.
The discomfort lies in what it reveals about ourselves. It challenges the belief that logic is enough, it suggests that trust, uncertainty, and human expectations play a far bigger role than what we like to admit.

4. Beyond The Courtroom

But this dilemma doesn't only exist in hypothetical scenarios and crime investigations.
This happens far more often than we realize. Imagine a group of students before an exam, they might all share notes and benefit equally. However, one might hesitate to do so in case someone else might receive an unfair advantage, the other might think the same. Ironically, when everyone makes this 'rational' decision, nobody receives a benefit they could all have gotten.
Similar scenarios happen in environmental well-being, where leaders stall climate change activities, waiting for other countries. In traffic, every driver choosing the faster shortcut makes it slower for everyone. Even our behaviour online, where silence spreads misinformation when no one is ready to challenge it.

5. Is Mathematics the issue?

So is mathematics the problem? Is it failing us by leading us down the wrong road, all in the name of 'logic'?
No. Mathematics isn't prescribing the wrong path; it's the mirror we hold up to the path we are on. It exposes the human reality, the gaps.
The Prisoner's Dilemma doesn't suggest a flaw in logic. It proves wrong the assumption that individual logic, applied in isolation, is enough.

So why do we expect fellow humans to betray us?
Perhaps because mathematics tells us to. Not as a prescription, but as a warning.
What makes this dilemma particularly striking is how naturally it reflects human behaviour. It is a built-in system for us.

Even in our personal relationships; the modern term 'situationships' can be explained through this dilemma: one may hesitate to be truthful or vulnerable, often unsure of the other's intentions, leading to commitment issues. Here, too, the same mathematics plays out.
Even on a global scale, conflicts between nations also work on the same logic. Cooperation could lead to mutual stability, however distrust and uncertainty, and in general human greed, can lead to defensive decisions.

6. A Universal Pattern

But this dilemma isn't uniquely human. Nature set this trap long before we arrived. Bacteria competing for resources, animals deciding to cooperate in a hunt, even trees growing in height to compete for more sunlight; all of this falls in the same mathematical structure, the same equilibrium, and the same results.
No courtroom, no prison, no choices. Just the same pattern, repeating, for centuries.

This suggests that the Prisoner's Dilemma isn't just the failure of human reasoning. It's way deeper than that. It's a structure of the universe, a situation which emerges in competing interests and no guaranteed trust. And mathematics didn't invent this trap, it just had the honesty to show us that it had always been there.

7. The Way Out

And yet.
Humans have found their way out. Institutions. Contracts. Reputation. They didn't eliminate self-interest; they reshaped it. They didn't try to become more rational or make betrayal the blind choice; instead, they made cooperation the norm. By building and changing these structures, they shifted the mathematics of it entirely and changed the game itself.

Which means that the lesson isn't despair. It's design. If the structure of our system pushes even good people to bad outcomes, then the problem isn't the people, it's the structure of the world we live in.

In the end, the Prisoner's Dilemma isn't about making logical decisions. It's about understanding the limits of logic itself.

8. The Conclusion

So we return to the original question: can Mathematics actually predict human decisions?
In a way, yes.
Not because it can control our choices and manipulate our decisions. But because it understands the structure of the world we live in, in which those choices are made.

Yet this does not make us powerless.
If mathematics can predict our choices, it can also show us how to change them.
To simply state, it isn't whether maths can predict our decisions, it's about whether we are willing to change the rules that make those predictions come true.
Because once we truly get the game, we don't just play it.
We begin to change it.

References

a. Flood, Merrill, and Melvin Dresher. “Prisoner's dilemma.” *Wikipedia*, 1950, https://en.wikipedia.org/wiki/Prisoner%27s_dilemma. Accessed 10 April 2026.

b. Nash, John Forbes, and Nathaniel T. Bacon. “Nash equilibrium.” *Wikipedia*, https://en.wikipedia.org/wiki/Nash_equilibrium. Accessed 13 April 2026.