Iterated Prisoner's Dilemma

The theory of learning in games studies the equilibrium concepts dictated by various simple learning mechanisms. That is, while the Nash equilibrium is based on the assumption of perfectly rational agents, in learning in games the assumption is that the agents use some kind of algorithm. The theory determines the equilibrium strategy that will be arrived at by the various learning mechanisms and maps these equilibria to the standard solution concepts, if possible. Many learning mechanisms have been studied. The most common of them are explained in the next few sub-sections. We saw how in the prisoner's dilemma the dominant strategy was to defect but both agents could have received a higher utility by cooperating. One way to try to get out of this conundrum, and to better simulate real-world interactions, is to let two agents play the same game some number of times. This new game is known as the *iterated prisoner's dilemma*. In general, we refer to these type of games as *repeated games*. Repeated games with a finite horizon are a special case of extended form games. They have, however, sometimes been studied independently.

One way to analyse repeated games that last for a finite number of periods is to backtrack from the end. Let's say you are playing an iterated prisoner's dilemma which will last for 50 rounds. You know that at the last round you will defect because that is the dominant strategy and there is no need to be nice to the other player as that is the last game. Of course, since you realized this then you also know that the other player realized the same thing and so he will also defect at the last round. As such, you know that you have nothing to gain by cooperating at round 49 so you will defect at 49, and so will he. This backward induction can be repeated any number of times leading us to the conclusion that for any finite number of games the rational strategy is to always defect. However, people don't act like this.

We can also formally prove a cooperative equilibrium for the iterated prisoner's dilemma if instead of a fixed known number of interactions there is always a small probability that every interaction will be the last interaction. That is, the agents never know if this will be their last interaction or not. In this scenario we can show that the dominant strategy for the iterated prisoner's dilemma is to cooperate, with a certain probability.

Axelrod performed some experiments on the iterated prisoner's dilemma. He sent out an email asking people to submit Fortran programs that played the prisoner's dilemma against each other for 200 rounds. The winner was the one that accumulated the most points. Many entries were submitted. They included the following strategies:

- *ALL-D:* always play defect;
- *RANDOM:* pick action randomly;
- *TIT-FOR-TAT:* cooperate in the first round, then do whatever the other player did last time;
- *TESTER:* defect on the first round. If other player defects then play tit-for-tat. If he cooperated then cooperate for two rounds then defect;
- *JOSS:* play tit-for-tat but 10% of the time defect instead of cooperating.

The *tit-for-tat* strategy won the tournament. It still made less than *ALL-D* when playing against it but, overall, it won more than any other strategy. It was successful because it had the opportunity to play against other programs that were inclined to cooperate. It cooperated with those that could cooperate and defected against the rest. This result, while intriguing, is not theoretically robust. For example a tit-for-tat strategy would lose against a strategy that plays tit-for-tat but defects on the last round. Still, the *tit-for-tat* strategy has been widely used and is considered to be a simple yet robust strategy.

References

J. M. Vidal, *Fundamentals of Multiagent Systems with NetLogo examples*, http://jmvidal.cse.sc.edu/papers/mas.pdf, 2010.

R. Axelrod, *The Evolution of Cooperation*, Basic Books, 1984.