444 Lecture 8.2 - Axelrod Tournament

Brian Weatherson

Overview

This lecture covers some of the lessons from the Iterated Prisoners' Dilemma tournaments that Michigan professor Robert Axelrod ran in the early 1980s.

Four Papers

- Effective Choice in the Prisoner's Dilemma, Journal of Conflict Resolution 24 (1980): 3-25.
- More Effective Choice in the Prisoner's Dilemma, Journal of Conflict Resolution 24 (1980): 379-403.
- The Emergence of Cooperation among Egoists, The American Political Science Review 75 (1981): 306-318.
- The Evolution of Cooperation with William Hamilton, Science 211 (1981): 1390-1396.

The First Tournament

- Axelrod advertised the first round of his tournament, and called for submissions.
- This was far from trivial in pre-internet days, and he only got 13 submissions.
- In the first tournament he said that k would be 100, but no one actually exploited that fact.

The Winner

Tit-for-Tat

Tit-for-Tat

Two rules.

1. Play C at round 1.
2. In all subsequent rounds, do whatever the other player just did.

The Second Tournament

- So Axelrod wrote this up, including saying who won.
- He called for more submissions, and now got 66.
- Some of these were typed, some came to Ann Arbor on the huge magnetic disks that were used way back then.
- He ran the tournament again, this time with a random number of rounds.

The Second Tournament

- So Axelrod wrote this up, including saying who won.
- He called for more submissions, and now got 66.
- Some of these were typed, some came to Ann Arbor on the huge magnetic disks that were used way back then.
- He ran the tournament again, this time with a random number of rounds.
- And Tit-for-Tat won again.

Logic and Victory

- This doesn’t mean Tit-for-Tat is the best strategy.
- Indeed, in each tournament it was easy in retrospect to describe strategies that would have beaten everyone, including TFT, if they had been entered.
- But still, it’s pretty impressive.

Four Features

Tit-for-Tat has five striking characteristics, each of which was positively correlated with success in the tournaments.

- Nice
- Provocable
- Forgiving
- Not envious
- Simple

Nice

The clearest distinction in the tournament was between strategies that were Nice and those that were Nasty.

- By definition, a strategy is Nice iff it is never the first to defect.
- You don’t have to be very nice in the intuitive sense to count as Nice.

Grim Trigger

Here is one nice strategy, one Axelrod calls Grim Trigger.

1. Cooperate on move 1.
2. If the other player ever defects, defect on every subsequent move.

This strategy did really badly; it was the worst Nice strategy in round 2. But still many Nasty strategies did worse.

Nice Strategies

- In the evolutionary versions of the game, there can be a tendency for strategies to tend towards being Nice.
- Then evolution stops, because when two Nice strategies meet, the payout is inevitably 3k to each.
- Although the best strategies are all Nice, it is how they interact with Nasty strategies that determines who wins.

Provocable

- It’s bad to get pushed around.
- Nasty strategies are always looking for how much they can get away with.
- So you want to send a clear message that defections will not be tolerated.
- Obviously TFT does that.

Forgiving

- But you don’t want to be Grim Trigger.
- It’s bad to be pushed around, but it’s not much better to end up in all defect land.
- You need a way back to all cooperate land.
- TFT has that, though notably it isn’t perfect at this.
- TFT can get into CD-DC-CD-etc cycles with a bunch of strategies.

Not Envious

- In any interaction, TFT never does better than who it is playing with.
- Yet it comes out first overall.
- This is kind of amazing.
- It just does not care at all about winning against who it is facing off with.

Not Envious To a Fault

- Note that TFT doesn’t always do that well in evolutionary games.
- This is because it might take this a bit too far.
- It doesn’t look to exploit weaknesses in opponents.

Simple

- Other strategies try to figure out what their rivals are doing.
- They normally get this wrong.

Simple

- Other strategies try to figure out what their rivals are doing.
- They normally get this wrong.
- Or they try and send complex signals.
- These are usually misinterpreted.

Simple

- Other strategies try to figure out what their rivals are doing.
- They normally get this wrong.
- Or they try and send complex signals.
- These are usually misinterpreted.
- TFT keeps things simple, and doesn’t lose points messing around looking for any edges.

Variant Games

- The most interesting variant to me is the one where a strategy only gets implemented with probability 0.99 on each move.
- Sometimes there are performance errors.
- TFT does terribly in this; it can't get out of randomly generated defection cycles.
- In this kind of game you need to be a bit more forgiving.
- But also you can try to get away with a bit more; if the other person will treat a defection as random, you can plan a few.

Rest of Day

- I'm not going to do slides/recordings about Oyun.
- If you have questions about it (and it isn't obvious), come along to class on Monday and we'll talk through how it works.