EVOLUTION AND THE THEORY OF GAMES

STEFAN GERITZ, HELSINKI, 2011

Transcript of my notes of 29-11-2011

34. In the 1970s, Robert Axelrod organized two computer contests for the iterated Prisoner's Dilemma (IPD) in which people could participate by submitting a strategy in the form of a computer program. Each strategy was paired with each other strategy and scored on the total points accumulated through the tournament. The very simple cooperative strategy Tit-for-Tat (see previous lecture) turned up as the winner.

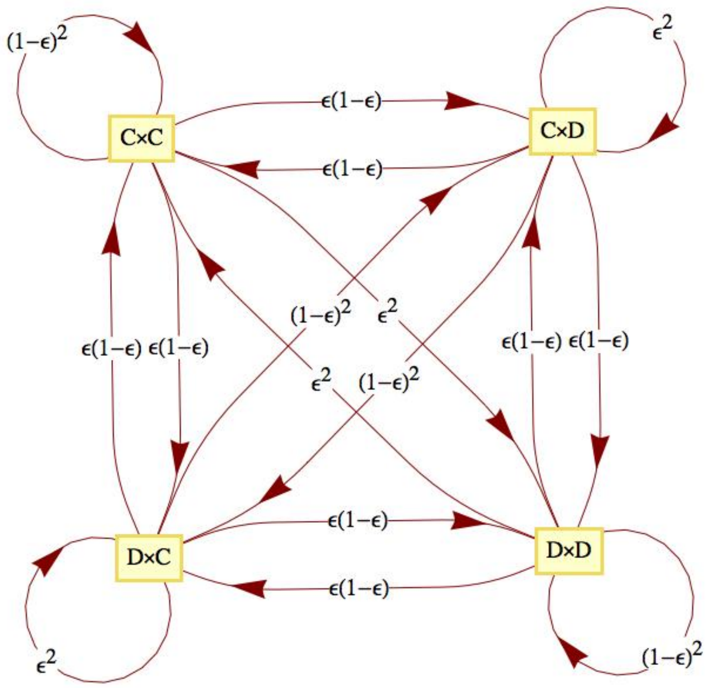
`(See, e.g., the Axelrod Tournament Demonstration Software website`
`http://www2.econ.iastate.edu/tesfatsi/demos/axelrod/axelrodt.htm)`

A weakness of TFT shows up if players can make mistakes, e.g., if they can misinterpret the opponent's move in a noisy environment. For example, in a TFT×TFT contest, if one player mistakes his opponent's 'cooperation' (C) for a 'defection' (D), the game gets trapped in a D×C→C×D→D×C etc. cycle with the corresponding low payoff to both players, at least until the next mistake occurs.

In another computer simulation with stochastic strategies in a noisy environment, Nowak and Sigmund (*Nature* (1992) **355,** 250-253) found that "Generous TFT" (GTFT) was the winner. GTFT is like TFT but "correct" mistakes by being cooperative with a given small probability after his opponent's defection. We'll now see how this works.

35. Depending on the current moves (C or D), and as long as we are dealing with memory-1 strategies (see previous lecture), the IPD can be in one of only four states: C×C, C×D, D×C or D×D. Contests with deterministic strategies (fixed or rule-based) in a deterministic setting therefore can never become more complicated than a four-cycle or less. In a stochastic setting, i.e., in a noisy environment which generates mistakes in the perception of the opponent's move, such limitations no longer hold.

As an example, consider TFT×TFT with a small $\varepsilon > 0$ probability of misinterpretation the opponent's move. The transition probabilities between the various states from one round to the next (conditioned on that there will be a next round) are given by the following graph:



If we put the different states of the game in the order {C×C, C×D, D×C, D×D}, then *the matrix of the graph* becomes

$$\mathbf{A}_\varepsilon = \begin{pmatrix} (1-\varepsilon)^2 & \varepsilon(1-\varepsilon) & \varepsilon(1-\varepsilon) & \varepsilon^2 \\ \varepsilon(1-\varepsilon) & \varepsilon^2 & (1-\varepsilon)^2 & \varepsilon(1-\varepsilon) \\ \varepsilon(1-\varepsilon) & (1-\varepsilon)^2 & \varepsilon^2 & \varepsilon(1-\varepsilon) \\ \varepsilon^2 & \varepsilon(1-\varepsilon) & \varepsilon(1-\varepsilon) & (1-\varepsilon)^2 \end{pmatrix}$$

Let further $\mathbf{r}$ be the vector of payoffs to the row-player at the different states of the game and $\mathbf{e}$ the vector of expected payoffs to the row-player calculated over all rounds for the contest if started in the respective states, i.e.,

$$\mathbf{r} = \begin{pmatrix} R \\ S \\ T \\ P \end{pmatrix} \quad \text{and} \quad \mathbf{e} = \begin{pmatrix} E_{C\times C} \\ E_{C\times D} \\ E_{D\times C} \\ E_{D\times D} \end{pmatrix}$$

Furthermore, let $\delta \in (0,1)$ be the probability that there will be a next round at all. Then, in matrix notation, we have

$$\mathbf{e} = \mathbf{r} + \delta \mathbf{A}_\varepsilon \cdot \mathbf{e}$$

which can be solved formally as

$$\mathbf{e} = (\mathbf{I} - \delta \mathbf{A}_\varepsilon)^{-1} \cdot \mathbf{r}$$

where $\mathbf{I}$ is the identity matrix. Since TFT×TFT starts with C×C, we're actually only interested in $E_{C\times C}$, which turns out to be

$$(*) \quad E_{C\times C} = \frac{1}{4}\left(\frac{P+R+S+T}{1-\delta} + \frac{P+R-S-T}{1-\delta(1-2\varepsilon)^2} - \frac{2(P-R)}{1-\delta(1-2\varepsilon)}\right)$$

36. What is the proportion of time TFT×TFT spends in each of the states C×C, C×D, D×C and D×D? Let $\mathbf{p}_n$ be the probability distribution over the various states. Then, as long as the game is on, $\mathbf{p}_n$ satisfies the recurrence equation

$$\mathbf{p}_{n+1} = \mathbf{A}_\varepsilon \cdot \mathbf{p}_n$$

with initial condition $\mathbf{p}_1 = (1,0,0,0)^{\mathrm{T}}$. This gives

$$\mathbf{p}_n = \mathbf{A}_\varepsilon^{n-1} \cdot \mathbf{p}_1$$

The expected number of rounds spent in the different states is

$$\mathbf{p}_1 + \delta \mathbf{p}_2 + \delta^2 \mathbf{p}_3 + \cdots = \left(\sum_{k=0}^{\infty} (\delta \mathbf{A}_\varepsilon)^k\right) \mathbf{p}_1 = (\mathbf{I} - \delta \mathbf{A}_\varepsilon)^{-1} \cdot \mathbf{p}_1$$

Division by the expected number of rounds, which is $(1-\delta)^{-1}$, gives the expected proportion of time spent in each state

$$\bar{\mathbf{p}}_\varepsilon \stackrel{\text{def}}{=} (1-\delta)(\mathbf{I} - \delta \mathbf{A}_\varepsilon)^{-1} \cdot \mathbf{p}_1$$

which, if written out in full, gives

$$\bar{\mathbf{p}}_\varepsilon = \begin{pmatrix} \frac{1}{4}\left(1 + \frac{1-\delta}{1-\delta(1-2\varepsilon)^2} - \frac{2(1-\delta)}{1-\delta(1-2\varepsilon)}\right) \\ \\ \frac{\delta\varepsilon(1-\varepsilon)}{1-\delta(1-2\varepsilon)^2} \\ \\ \frac{\delta\varepsilon(1-\varepsilon)}{1-\delta(1-2\varepsilon)^2} \\ \\ \frac{\delta\varepsilon^2(1-\delta(1-2\varepsilon))}{(1-\delta(1-2\varepsilon)^2)(1-\delta(1-2\varepsilon))} \end{pmatrix}$$

Expansion of $\bar{\mathbf{p}}_\varepsilon$ into terms of different orders of ε gives an idea of the distribution if mistakes are rare:

$$\bar{\mathbf{p}}_\varepsilon = \begin{pmatrix} 1 - \frac{2\delta\varepsilon}{1-\delta} + \mathrm{O}(\varepsilon^2) \\ \\ \frac{\delta\varepsilon}{1-\delta} + \mathrm{O}(\varepsilon^2) \\ \\ \frac{\delta\varepsilon}{1-\delta} + \mathrm{O}(\varepsilon^2) \\ \\ \mathrm{O}(\varepsilon^2) \end{pmatrix}$$

Thus, for small ε, TFT×TFT spends most of its time in mutual cooperation and only a tiny fraction of its time in mutual defection. How do these results compare to, e.g., a GTFT×TFT contest? That's what we'll look at next.

37.

To be continued . . .