

EVOLUTION AND THE THEORY OF GAMES

STEFAN GERITZ, HELSINKI, 2011

Transcript of my notes of 24-11-2011

30. This lecture is about iterated games. In an iterated game, the same game is played over and over again the same two opponents. Each repetition is called a ‘round’. The number of rounds may be finite, infinite or random.

Remember the Prisoner’s Dilemma (PD) with the strategies ‘defect’ (D) and ‘co-operate’ (C) and payoff matrix

PD	D	C
D	P, P	T, S
C	S, T	R, R

(payoffs to the row-player) with $S < P < R < T$ and $S + T < 2R$. ‘Defect’ is a strictly dominating strategy, and therefore (D,D) is the dominating strategy solution, even though *both* players would have a higher payoff if they played (C,C). Is this different in the Iterated Prisoner’s Dilemma (IPD)?

31. If the IPD is played exactly N times and both players know this, then the strategy ‘always defect’ is dominant. The proof is by backward induction: ‘defect’ is dominant in the last round, and so both players will defect in round N . Given that both players defect in round N , ‘defect’ is also dominant in the second-to-last round, and so both players will defect in round $N - 1$ as well, and so on. The same applies if the game length is unknown but has a known upper limit.

For cooperation to emerge between two players, the total number of rounds N must be random and unknown to the players. The most common way (but not the only way) to implement this into the model is by assuming that after each round there is a constant probability $\delta \in (0, 1)$ that there is another round. The total number of rounds then is a random variable with a geometric probability distribution and expectation

$$1 + \delta + \delta^2 + \delta^3 + \dots = \frac{1}{1 - \delta}$$

A strategy in the IPD tells for each round what action (D or C) to choose. We can distinguish between fixed strategies, random strategies and rule-based strategies. Examples of fixed strategies are:

- allD = (D,D,D,D,D,...)
- allC = (C,C,C,C,C,...)
- Alternate=(C,D,C,D,C,...)
- etc.

A random strategy is a sequence (p_1, p_2, p_3, \dots) where $p_n \in [0, 1]$ is the probability of choosing D in the n^{th} round. Obviously, the fixed strategies form a subset of the random strategies.

In a rule-based strategy, the choice of action depends on the history of the game up to that moment. For example, take ‘Tit for tat’ (TFT): “*Choose C in the first round; after that choose whatever your opponent did in the previous round.*” Successive rounds for TFT against an opponent thus look like:

TFT:	..	D	C	C	D	C	D	D	C	C	C	D	C	..
opponent:	..	D	C	C	D	C	D	D	C	C	C	D	C	..

Tit for tat is an English saying meaning “equivalent retaliation”: TFT rewards cooperation with cooperation and punishes defection with defection.

Another example of a rule-based strategy is ‘Pavlov’: “*Choose C in the first round; after that repeat the same action as in the previous round if your payoff was high (i.e., R or T); otherwise change.*” Successive rounds for Pavlov against the same opponent as above thus look like:

Pavlov:	..	C	D	D	D	C	C	D	C	C	C	C	D	D	..
opponent:	..	D	C	C	D	C	D	D	C	C	C	D	C	..	

Pavlov is a ‘win-stay, lose-switch’ strategy. Variations of Pavlov start with a defection, or recognize only T or all three P , R and T as high payoffs.

TFT and Pavlov are so-called memory-1 strategies, i.e., their rule only needs to remember the previous round. An example of a memory-2 strategy (which remembers the last two rounds) is ‘Tit for two tats’ (TFTT): “*Start with two rounds of C; after that respond with defecting only if the opponent defects twice in a row.*”

TFTT:	..	C	C	D	C	C	D	C	C	C	C	C	..
opponent:	..	D	C	D	D	C	D	C	C	C	D	C	..

In the TFT strategy, once the opponent defects, the TFT player immediately responds by defecting on the next move. If a C (with some low probability) is misinterpreted by the opponent as a D, then two TFT players may accidentally

get stuck in a $(C \times D)$ - $(D \times C)$ -cycle (or even a $(D \times D)$ -cycle), resulting in a poor outcome for both players. A TFFT player, however, will let the first defection go unchallenged as a means to avoid the above trap. Only if the opponent defects twice in a row, the tit for two tats player will respond by defecting.

Another example of a memory-2 strategy is ‘Two tits for Tat’ (TTFT): “*Start with C; after that respond with two defections to every defection of the opponent.*”

TTFT: .. D D C D D D D D C C D D ..
 opponent: .. D C C D C D D C C C D C ..

The strategy ‘Grim’ starts with C but changes to D after the very first defection by its opponent and plays D from then onwards. Grim is an example of a memory- ∞ strategy: it never forgets and never forgives.

32. How to calculate the payoffs for the IPD with, e.g., TFT against Pavlov? To make the example a bit more interesting we shall use a version of Pavlov that starts in the first round with a defection, and call this slightly more grim variant **Pavlov***. So the question is: what are the entries of the following payoff matrix?

IPD	TFT	Pavlov*
TFT	?	?
Pavlov*	?	?

TFT \times **TFT** gets immediately in a $(C \times C)$ -cycle. The overall expected payoff E_{CC} to TFT against TFT is given by

$$E_{CC} = R + \delta E_{CC}$$

and hence

$$(1) \quad E_{CC} = \frac{R}{1 - \delta}$$

where $\delta \in (0, 1)$ is the probability of a next round.

TFT \times **Pavlov*** gets into a $(C \times D)$ - $(D \times D)$ - $(D \times C)$ -cycle. Let now E_{CD} and E_{DD} and E_{DC} denote the overall payoffs to TFT if the cycle is started in $(C \times D)$ or $(D \times D)$ or $(D \times C)$, respectively. Then we have

$$\begin{cases} E_{CD} = S + \delta E_{DD} \\ E_{DD} = P + \delta E_{DC} \\ E_{DC} = T + \delta E_{CD} \end{cases}$$

which is readily solved for E_{CD} , E_{DD} and E_{DC} . Of these we only need the first one, because that’s how a **TFT** \times **Pavlov*** contest actually starts. This gives us

$$(2) \quad E_{CD} = \frac{S + \delta P + \delta^2 T}{1 - \delta^3}$$

Pavlov* \times **TFT** gets into a (D \times C)-(D \times D)-(C \times D)-cycle. If E_{DC} and E_{DD} and E_{CD} denote the overall payoffs to Pavlov* for different starting points in the cycle, we have

$$\begin{cases} E_{DC} = T + \delta E_{DD} \\ E_{DD} = P + \delta E_{CD} \\ E_{CD} = S + \delta E_{DC} \end{cases}$$

from which we solve

$$(3) \quad E_{DC} = \frac{T + \delta P + \delta^2 S}{1 - \delta^3}$$

Pavlov* \times **Pavlov*** gives (D \times D) in the first round followed by a (C \times C)-cycle. With E_{DD} and E_{CC} denoting the overall payoffs to Pavlov* for different starting points, we have

$$\begin{cases} E_{DD} = P + \delta E_{CC} \\ E_{CC} = R + \delta E_{CC} \end{cases}$$

from which we solve

$$(4) \quad E_{DD} = P + \frac{\delta R}{1 - \delta}$$

For the overall game we collect the payoffs (1)-(4) in the payoff matrix:

IPD	TFT	Pavlov*
TFT	$\frac{R}{1-\delta}, \frac{R}{1-\delta}$	$\frac{S+\delta P+\delta^2 T}{1-\delta^3}, \frac{T+\delta P+\delta^2 S}{1-\delta^3}$
Pavlov*	$\frac{T+\delta P+\delta^2 S}{1-\delta^3}, \frac{S+\delta P+\delta^2 T}{1-\delta^3}$	$P + \frac{\delta R}{1-\delta}, P + \frac{\delta R}{1-\delta}$

We conclude that TFT is an ESS if

$$\frac{R}{1-\delta} > \frac{T + \delta P + \delta^2 S}{1 - \delta^3}$$

and that Pavlov is an ESS if

$$P + \frac{\delta R}{1-\delta} > \frac{S + \delta P + \delta^2 T}{1 - \delta^3}$$

One readily shows that TFT is an ESS for sufficiently large $\delta \in (0, 1)$, i.e., if the number of rounds tends to be high. I'm not entirely sure, but I think that Pavlov* is always an ESS against TFT. This need no longer be true, however, if we consider additional strategies which may turn out to be better.

32. Now that we know how to calculate payoffs in the IPD, let's do a little tournament in the style of Axelrod & Hamilton (*Science* (1981) **211**, 1390-1396): we play allD (the champion of the IPD with a fixed number of rounds) against TFT.

TFT×**TFT** gets immediately in a (C×C)-cycle, and the overall expected payoff E_{CC} to TFT against TFT (as calculated previously) is

$$E_{CC} = \frac{R}{1 - \delta}$$

TFT×**allD** gives (C×D) in the first round followed by a (D×D)-cycle. The overall payoff when starting with (C×D) is calculated as illustrated in the previous section and turns out to be

$$E_{CD} = S + \frac{\delta R}{1 - \delta}$$

allD×**TFT** gives (D×C) in the first round followed by a (D×D)-cycle. The overall payoff when starting with (D×C) is

$$E_{CD} = T + \frac{\delta R}{1 - \delta}$$

allD×**allD** immediately settles down in a (D×D)-cycle, and so the overall payoff to allD against allD is

$$E_{CC} = \frac{P}{1 - \delta}$$

For the payoff matrix of the IPD with strategies TFT and allD we thus have

IPD	TFT	allD
TFT	$\frac{R}{1-\delta}, \frac{R}{1-\delta}$	$S + \frac{\delta P}{1-\delta}, T + \frac{\delta P}{1-\delta}$
allD	$T + \frac{\delta P}{1-\delta}, S + \frac{\delta P}{1-\delta}$	$\frac{P}{1-\delta}, \frac{P}{1-\delta}$

TFT is an ESS whenever $\frac{R}{1-\delta} > T + \frac{\delta P}{1-\delta}$, i.e., whenever $1 - \frac{R-P}{T} < \delta < 1$. In other words, TFT is an ESS against allD if δ is large and the game, on average, continues for many rounds. AllD is an ESS whenever $\frac{P}{1-\delta} > S + \frac{\delta P}{1-\delta}$, i.e., whenever $P > S$, which, by assumption, is always true.

33. Normally, allC against allD is an all-time loser. But suppose that we equip allC with the rule “*Quit whenever you receive the sucker’s payoff S*”, and denote this new variant **allC***. Like TFT, allC* has a means of punishing a defector, not by reciprocating the defection, but by simply quitting the game. An (allC*×allC*)-contest lasts an average of $(1 - \delta)^{-1}$ rounds, as does a (allD×allD)-contest, but a (allC*×allD)-contest lasts only one round. The payoff matrix is

IPD	allC*	allD
allC*	$\frac{R}{1-\delta}, \frac{R}{1-\delta}$	S, T
allD	T, S	$\frac{P}{1-\delta}, \frac{P}{1-\delta}$

One readily sees that allD against allC* is always an ESS, and allC* is an ESS against allD whenever $\frac{R}{1-\delta} > T$, i.e., whenever $1 - \frac{R}{T} < \delta < 1$, i.e., whenever the average number of rounds is sufficiently high.