CMPS 272	**Jason Rohrer**
Class Lecture Notes	**January 29, 2002**

1 Social creatures

1.1 Four quadrants

Evolution is driven by fitness and competition for limited resources. Resources are often fellow creatures. Figure 1 shows the fitness impact of various actions. Figure 2 on page 2 shows the division between efficient and inefficient actions.

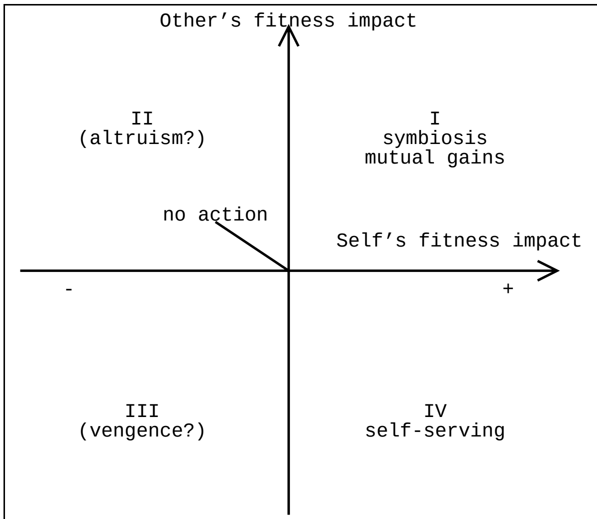

Figure 1: The four quadrants of alternative action outcomes.

Direct evolution favors actions producing outcomes in IV- and opposes actions in II+, though this is not socially efficient. Social creatures thrive on devices that encourage actions in II+ and discourage actions in IV-, and they evolve mechanisms to achieve this.

1.2 Devices: r, δ

Relatedness is one mechanism that encourages social behavior. If r is the coefficient of relatedness between two individuals, then evolution favors actions with impact above the line through $(0, 0)$ with slope $-r$. A given action with self-benefit x and other-benefit y is efficient if $x + y > 0$. Thus the relatedness device reconciles social efficiency with evolution if $r = 1$, but it falls short for smaller r. For humans who interact, r is typically close to 0. Thus, relatedness for diploid organisms with small families does not work very well to motivate social behavior.

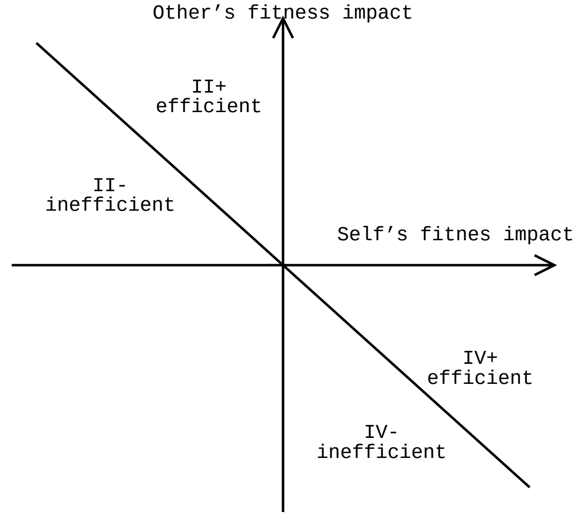

Figure 2: Efficient actions.

Reciprocity is another mechanism encouraging social behavior. We discuss reciprocity and the δ coefficient in the next section. A third mechanism, emotional valence, will be discussed in another lecture.

2 Repeated games

2.1 Subgame perfection

In an extensive form game, we have subgame perfection if the Nash equilibrium (NE) of the entire game is such that it induces a NE on each subtree that is itself an extensive form game. Figure 3 on page 3 shows an extensive form game and its three subgames. When #1 and #2 choose actions, they must each reason about the credible and non-credible threats in the future actions of the other player. NE that are subgame perfect (or, which is the same in games of complete information, obtained by back ward induction) involve only credible threats.

2.2 Stage game

The stage game is the game played in each round of the repeated game and can be any normal form or extensive form game. For example, consider the symmetric prisoner's dilemma (PD) with payoff matrix

$$\begin{matrix} c \\ d \end{matrix} \begin{pmatrix} 5 & -3 \\ 8 & 0 \end{pmatrix}$$

2.3 Repeated game ingredients

Repeated games have the following features:

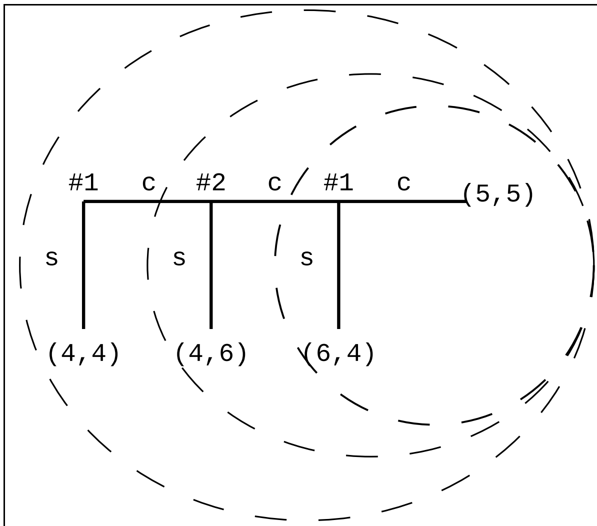

Figure 3: An extensive form game with its three subgames.

- a stage game, as described above
- a series of stages, for example, $t = 1, 2, \cdots, T$, where $T \leq \infty$
- histories: all players know all previous actions and outcomes
- strategies: functions mapping histories to actions (possibly mixed)
- payoffs that are a function of each stage in the game. For example, we might take the average, minimum, or maximum payoff. We typically take a discounted present value (PV), which is a form of weighted average.

2.4 Discounting

For PV, we have a discount factor $\delta \in (0, 1]$. Let $t = 0$ be the current time. We want to reason about future payoffs. If $f_{i,t}$ is the payoff for player i at future time t, we define our payoff as

$$\pi_i = \sum_{t=1}^{T} \delta^t f_{i,t}.$$

Note that δ summarizes:

- the time value of resources (or preferences). For example, if r is an interest rate, we have

$$\delta = \frac{1}{1+r}.$$

- the probability of continuing. If $p < 1$ is the probability of continuing, with interest rate r, we have

$$\delta = \frac{p}{1+r}.$$

Intuitively, we prefer payoffs sooner rather than later. Money received now and invested for a year is worth more than that same amount of money promised for next year. The same idea holds if the game can terminate at random in the future: the more likely the termination, the less we can count on future payoffs.

2.5 Some repeated game strategies

2.5.1 Uncontingent

For example, always play C, or alternate D and C, regardless of history.

2.5.2 History-conscious

Tit for tat (TFT):

- play C at time $t = 1$
- play opponent's $t - 1$ strategy at time $t > 1$

Grim trigger strategy:

- play C until opponent first plays D
- play D forever after

More complicated examples are also possible.

3 Repeated social dilemmas and reciprocity

3.1 Unraveling when $T < \infty$

Consider the repeated PD for $T < \infty$, say $T = 7$, and examine the players' choices at $t = 7$. D is the dominant strategy for both players in this one-stage subgame, so the NE will be (D, D).

Now consider the players' choices at $t = 6$. Both players realize that (D, D) will be played at $t = 7$ no matter what, so the future payoffs cannot be influenced by their current actions. Thus, we still have a one-stage maximization problem. The NE for the subgame $t = 6$ is (D, D). By induction, we can carry this argument all the way back to $t = 1, \cdots, 7$, so the all-D strategy is the unique subgame-perfect NE, and it is the only NE for this repeated game. Of course, we are assuming that both players are rational and use backward induction to select their strategies.

In real life, people tend to cooperate up to the last few stages of the game. If they are given the chance to play the multi-stage game again, they start defecting sooner. However, no matter how many repeated sets they play, they seldom defect the whole time. People, in general, are not completely rational.

Theoretically, cooperation unravels to (D, D) when the known T is finite. Cooperation is possible for infinite T, however.

3.2 Cooperation when $T = \infty$

Consider the repeated PD for $T = \infty$ and $\delta < 1$. We claim that (Grim, Grim), or (G, G), is a NE that induces cooperation for the entire game, if δ is suitably high (*e.g.*, $\delta = 0.9$). A sketch of the proof follows.

Suppose not. Let t be the time of the first defection, say for player 1, and show that defection at time t is irrational for player 1. Note that the payoff stream from t onward is

$$8, 0, 0, 0, 0, 0, \cdots$$

If player 1 instead played C at time t (and waited at least one more period to defect), player 1's payoff stream from t onward would be at least

$$5, 8, 0, 0, 0, 0, \cdots$$

And, for $\delta > 3/8$,

$$5\delta + 8\delta^2 > 8\delta.$$

Playing C against G at finite time t has a higher PV than playing D, so playing D is not a best response to G. Thus G is a best response to G, so (G, G) is a NE.

3.3 Folk theorem

The above results are somewhat trivialized when we realize that almost *anything* is a NE in a repeated game. Consider Figure 4 on page 6. The diamond-shaped region represents payoff vectors of all possible mixed strategies in the PD game we were considering. The shaded kite-shaped region represents all individually rational (what a player can guarantee itself) strategy mixes. It turns out that any point in the shaded region is the outcome of an NE in the repeated game. The Folk theorem, due to J. Friedman *et al.*, circa 1970, is as follows:

> Trigger strategies can support *any* individually rational outcome of the stage game if $\delta < 1$ is suitably large.

4 Direct reciprocity

Suppose that self chooses a behavior with personal cost C that benefits the other by B, where $B > C > 0$. This will evolve if there is reciprocity with a discount factor of $\delta > C/B$. In other words, if we are above the line defined by $\delta B - C > 0$, then the behavior is sustainable.

Intuitively, we can help the other now and incur a cost, but if the other is playing the same strategy, we will be helped by the other later and receive a benefit. Figure 5 shows the cost-benefit cycle in a reciprocal relationship.

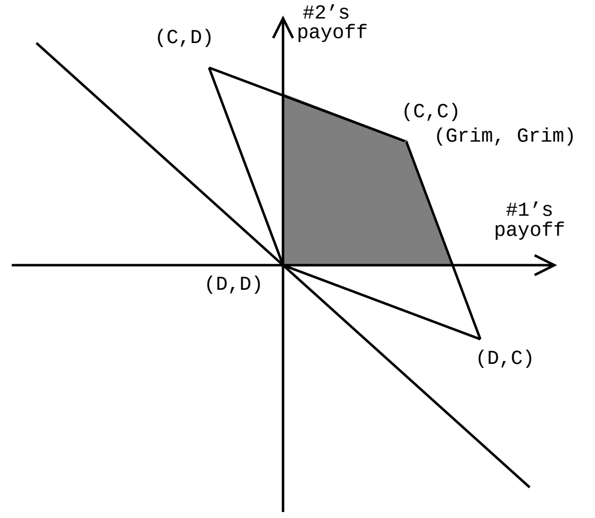

Figure 4: The region for NE's in the Folk theorem.

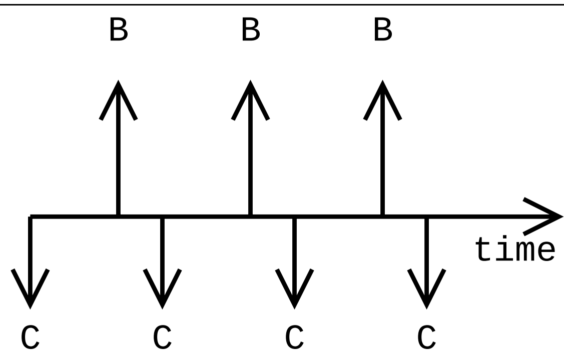

Figure 5: Direct reciprocity cycle.

If we take no action, we get 0. However, if we take cooperative action, we can get

$$-C + \delta(B-C) + \delta^2(B-C) + \cdots = -C + (B-C)\sum_{t=1}^{\infty}\delta^t, \text{ or}$$

$$-C + (B-C)\frac{\delta}{1-\delta}$$

If this value is greater that 0, taking the cooperative action is rational, *i.e.*, if

$$(B-C)\delta > (1-\delta)C,$$

or if

$$B\delta > C.$$

Using the notation of Figure 2 on page 2, $B = x$ and $C = -y$, the condition is that we are above the line $x + \delta y = 0$. Thus reciprocity in the repeated game can induce efficient social behavior exactly as can relatedness (r). But for humans (and some other social mammals such as vampire bats) $\delta > r$ and so is a more effective mechanism.