Prisoner's Dilemma
Anatol Rapoport and Albert M. Chammah
http://www.press.umich.edu/titleDetailDesc.do?id=20269
The University of Michigan Press, 2009.

Chapter 8

Equilibrium Models with Adjustable Parameters

So far we have been examining models in which the parameters represent fixed propensities of the subjects. Whatever dynamics emerge in these models is a consequence of interaction among probabilistically determined events, the laws of interaction being determined by the fixed parameters. Now we shall consider models in which the parameters themselves undergo changes.

As a simplest example consider two subjects characterized by a single pair of parameters C_1 and C_2, these being, as before, the probabilities of responding cooperatively in a Prisoner's Dilemma game. Assume now that each subject is able to adjust his own parameter and that he seeks to adjust it so as to "maximize his expected gain." From the nature of Prisoner's Dilemma, we see that if each makes calculations based on the payoffs, then each player will set his C equal to zero, because the D response is "better" against either response of the other player. Suppose, however, that the players are sentient but not calculating beings, that is, they are governed by a homeostatic type process in which the behavioral parameters seek the most favorable gradient with respect to the resulting payoffs averaged over some interval of time.

To fix ideas, assume that C_2 is held constant and that player 1 seeks to adjust his C_1 in the way just mentioned. Player 1 will try some value of C_1 for a while, note the average payoff that accrues to him, then try another value $C_1' > C_1$. If the resulting average payoff is larger, he will increase C_1 still more on the next trial. If the resulting average payoff is smaller, he will try an adjustment in the opposite direction, i.e., will adopt $C_1'' < C_1$. In this way, he goes through a search procedure which will ultimately carry him to an optimal value of C_1 (given a specified constant value of C_2). This value will be optimal in the sense that it will correspond to at least a local maximum in the average payoff accruing to player 1.

In the meantime, player 2 will be going through the same procedure. We would expect that this type of search-and-fix adjustment would carry both C_1 and C_2 to zero. Let us show this formally. Call the expected payoff of the two players respectively G_1 and G_2. Then, in view of the payoffs R, T, S, and P associated with the four states, we have the following equations:[23]

$$G_1 = (CC)R + (CD)S + (DC)T + (DD)P, \qquad (69)$$

$$G_2 = (CC)R + (CD)T + (DC)S + (DD)P. \qquad (70)$$

In these equations the four states now represent *instantaneous probabilities*, not overall frequencies. Since in the absence of communication, we must assume the responses of the two players *at a given moment* to be independent, we may equate CC to $(C_1)(C_2)$, CD to $(C_1)(D_2)$, etc.

To find an extremum, we take the partial derivatives of G_1 and G_2 with respect to C_1 and C_2 respectively. This gives, in view of the fact that $D_i = 1 - C_i$,

$$\frac{\partial G_1}{\partial C_1} = (R - T)C_2 + (S - P)D_2, \qquad (71)$$

$$\frac{\partial G_2}{\partial C_2} = (R - T)C_1 + (S - P)D_1. \qquad (72)$$

But by definition of Prisoner's Dilemma, $R < T$ and $S < P$. Hence the partial derivatives are negative for all positive values of $C_i (i = 1,2)$, and so this "adjustment model" must lead to the stable result $C_1 = C_2 = 0$, i.e., to the persistence of the DD state.

The situation is different if the *conditional* probabilities (the propensities) are adjustable. Already the simplest yields an interesting result instead of a trivial one.

Consider two subjects with propensities x_1 and x_2 between 0 and 1, while $y_1 = y_2 = z_1 = z_2 = 0$; $w_1 = w_2 = 1$. We are again dealing with our two tempted simpletons (cf. p. 73). For simplicity of computation, we shall suppose that they are playing a game in which $S = -T$, $P = -R$ (e.g., Games III, IV or V of our experiments). From Equations (54)-(57), we now have the following expected payoffs:

$$G_1 = \frac{Rx_1x_2 + T(x_2 - x_1)}{2 + x_1 + x_2 - 3x_1x_2}, \qquad (73)$$

$$G_2 = \frac{Rx_1x_2 + T(x_1 - x_2)}{2 + x_1 + x_2 - 3x_1x_2}. \qquad (74)$$

The partial derivatives of G_1 and G_2 with respect to x_1 and x_2 respectively are:

$$\frac{\partial G_1}{\partial x_1} = \frac{x_2^2(3T + R) + 2x_2(R - T) - 2T}{(2 + x_1 + x_2 - 3x_1x_2)^2}, \qquad (75)$$

$$\frac{\partial G_2}{\partial x_2} = \frac{x_1^2(3T + R) + 2x_1(R - T) - 2T}{(2 + x_1 + x_2 - 3x_1x_2)^2}. \qquad (76)$$

If an equilibrium exists, it must be where the numerators of both expressions (75) and (76) vanish. Both the numerators[24] represent the same parabola, which crosses the horizontal axis at

$$x = x^* = \frac{T - R \pm \sqrt{R^2 + 7T^2}}{3T + R}. \qquad (77)$$

For this value to have meaning in our context, we must have $0 \leq x \leq 1$. Hence only the positive square roots need to be considered. Furthermore, we must have:[25]

$$T - R + \sqrt{7T^2 + R^2} \leq 3T + R, \qquad (78)$$

$$7T^2 + R^2 \leq [2(T + R)]^2, \tag{79}$$

$$7T^2 + R^2 \leq 4T^2 + 8TR + 4R^2, \tag{80}$$

$$3T^2 - 8TR - 3R^2 \leq 0. \tag{81}$$

Inequality (81) holds if and only if $T \leq 3R$.

Our first conclusion, therefore, is that an equilibrium state will exist between 0 and 1 if and only if $T \leq 3R$.[26]

Next we observe that this equilibrium value of x^* is the larger, the larger the magnitude of T. To see this, differentiate (77) partially with respect to T:

$$\frac{\partial x^*}{\partial T} = \frac{4RQ + 7RT - 3R^2}{Q} \tag{82}$$

where we have denoted $\sqrt{7T^2 + R^2}$ by Q. Since $7RT > 3R^2$, $\partial x^*/\partial T > 0$.

It seems strange at first that the equilibrium value of x^* should increase with T inasmuch as we intuitively associate larger values of T with *smaller* frequencies of cooperative responses, while the latter increase with both x_1 and x_2. A closer look at the nature of the equilibrium dispels our doubts concerning the intuitive acceptability of this result. Observe that the equilibrium obtains where the parabola represented by the numerators of expressions (75) and (76) has its *larger* root (cf. Figure 27). Moreover, the parabola is concave upward, since the coefficient of its squared term is positive. It follows that as, say, x_1 increases to a value larger than x^*, the partial derivative $\partial G_1/\partial x_1$ becomes positive and the conditional probability x_1 is adjusted to a still larger value. Mutatis mutandis, if x_1 falls below x^*, $\partial G_1/\partial x_1$ becomes negative, and so the parameter will be adjusted to a still lower value. Hence, not homeostasis but a positive feedback is operating in this system. The equilibrium at x^* is unstable and cannot persist. The value x^* is, in fact, a

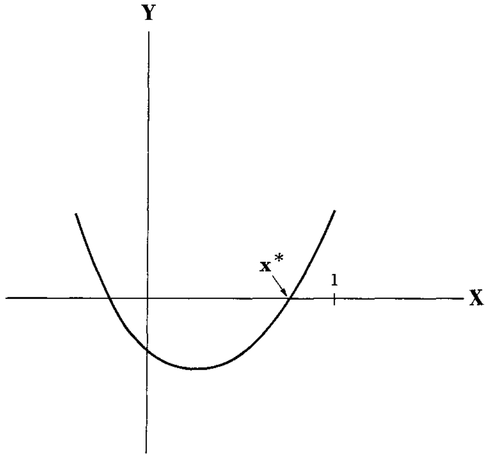

Figure 27. The parabola $y = (3T + R)X^2 + 2X(R - T) - 2T$, where $T < 3R$.

point on a "watershed" in (x_1x_2)-space. When x_1 and x_2 fall short of this watershed, they tend to become smaller: the system tends toward noncooperation. If x_1 and x_2 pass over the "hump," they will tend to become still larger: the system will tend toward cooperation.

Now it is clear why x^* increases with T. The larger the temptation, the more it takes to push the system "over the hump" toward increasing cooperation. If $T > 3R$, it is altogether impossible to do so. In this case $x^* > 1$, hence it is nonexistent as a probability, and the system will go toward noncooperation whatever are the values of x_1 and x_2.

Note that in this very special case $x_1 = x_2 = 0$, $C_1 = C_2 = \frac{1}{2}$. This can be seen from Equations (54) and (57). The reason for this is the high value of w $(w = 1)$, which insures that every time our two simpletons find themselves in the DD state, they cooperate on the next play. The restriction was applied solely in the interest of simplifying the mathematics, since we are using this model for illustrative purposes only.

In summary, we have described a model which generates a "Richardson Effect," i.e., an unstable situation in which both the tendency to cooperate and to defect are "self-propelling," governed by a positive feedback.