Control and Cybernetics

vol. **22** (1993) No. 4

Non–linear intertemporal dilemma models

by

Qiu Yuhuang and Wu Jianzhong

Institute of Automation
Academia Sinica
P.O. Box 2728
Beijing, 100080
P. R. China

In social dilemmas a short–term or individual need is pitted against a long–term or group need. Many scientists study social dilemmas and models of social dilemmas. So far, models of intertemporal social dilemmas are linear. In this paper, a nonlinear intertemporal model is proposed to describe and analyse a typical social dilemma. A recurrence formula is also given for individual profits in any time interval. The reinforcement mechanism of 'social traps' has been described quantitatively. Two social dilemma models are described algorithmically. The second model provides a model of the overuse of resources.

Keywords: Intertemporal dillemas, non–linear models, dilemma model.

1. Introduction

Hardin (1968) pointed out the following problem: People can abuse a resource for personal interest, bringing about the ruin of the scarce resource, or the "tragedy of the commons". Many scholars have researched this so–called social dilemma.

Intertemporal models of social dilemmas give some assistance in understanding complex social behaviours, as in social conflicts, armament races, environmental pollution, overfishing, and so on.

In an article on Social Traps, Platt (1973) described the social dilemma with the Skinnerian theory of reinforcement. According to his theory, social dilemmas can be divided into three categories. One of these is the one–person trap, in which short–term goods result in long–term bads. This kind of dilemma is expressed with nonlinear models in this paper. Two social dilemma models are given.

2. A nonlinear model of an intertemporal dilemma

An intertemporal social dilemma occurs when each optimal choice will result in long–term bads. For example, with smoking cigarettes or taking drugs, the short–term benefits can lead to long–term results. One begins to smoke cigarettes for curiosity or social reasons. Gradually, changes take place in the body; this addiction is reinforced again and again, and finally serious damage occurs. So intertemporal dilemmas have two properties: (1) there is a strategy which yields the best profit in each step; (2) choosing this strategy in all steps result in a deficient outcome.

A nonlinear model of an intertemporal social dilemma is described as follows:

One must choose between two strategies, C and D, in every step of time sequence. D expresses behaviour such as defection, smoking, taking drugs, polluting the environment, overfishing, and so on. On the contrary, C expresses cooperation, not smoking or giving up smoking, not taking drugs or giving up drugs, and so on. U_i expresses the choice at the ith step. $U_i = 1$ means choosing D strategy, and $U_i = 0$ choosing C strategy.

Assuming that the personal profit function at ith step is J_k,

$$J_k = \begin{cases} \alpha/(P_k + L_c) & U_k = 0 \quad (C) \\ \alpha/(P_k + L_d) & U_k = 1 \quad (D) \end{cases} \tag{1}$$

where $L_c, L_d > 0$, $L_c - L_d > 1$, $\alpha > L_c^2$, $P_k = \sum_{i=0}^{k} U_i$, i.e. P_k is the number of steps in which D is chosen.

A baby does not smoke or take drugs. At the beginning, one does not want to take drugs. So the first step choice is cooperation, i.e. $U_0 = 0$, and $P_0 = 0$. Therefore, the initial value of the profit function $J_0 = \alpha/L_c$ lies on the cooperating profit curve.

Now let us examine the ith step to see what will happen.

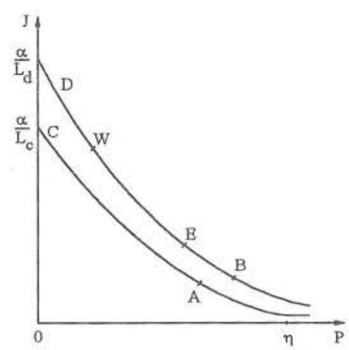

Fig. 1

Case 1. Suppose cooperation, C, was chosen at the last step, i.e. $U_k = 0$, the value of personal profit function will not be changed; if D is chosen, i.e. $U_{k-1} = 0$ (see Fig. 1 at point A). If C is chosen, i.e. $U_k = 1$, we obtain the increment of $\alpha/(P_k + L_d) - \alpha/(P_{k-1} + L_c)$ of the value of personal functions (at point B). So, when $U_{k-1} = 0$, we have

$$J_k = \begin{cases} J_{k-1} & U_k = 0 \quad (C) \\ J_{k-1} + \alpha/(P_k + L_d) - \alpha/(P_{k-1} + L_c) & U_k = 1 \quad (D) \end{cases} \tag{2}$$

or

$$J_k = J_{k-1} + [\alpha/(P_k + L_d) - \alpha/(P_{k-1} + L_c)]U_k \tag{3}$$

Case 2. Suppose defection, D, was chosen at the last step, i.e. $U_{k-1} = 1$ (see Fig. 1 at point E). If C is chosen, i.e. $U_k = 0$, the value of personal profit function will be decreased by $\alpha/(P_{k-1} + L_d) - \alpha/(P_{k-1} + L_c)$ (at point A). If D is chosen, i.e. $U_k = 1$, the value of personal profit function will be decreased by $\alpha/(P_{k-1} + L_d) - \alpha/(P_k + L_d)$ (at point B). So, when $U_k = 1$, we have

$$J_k = \begin{cases} J_{k-1} - \alpha/(P_{k-1} + L_d) + \alpha/(P_{k-1} + L_c) & U_k = 0 \quad (C) \\ J_{k-1} + \alpha/(P_k + L_d) - \alpha/(P_{k-1} + L_d) & U_k = 1 \quad (D) \end{cases} \tag{4}$$

or

$$\begin{aligned} J_k = J_{k-1} &+ [\alpha/(P_{k-1} + L_c) - \alpha/(P_{k-1} + L_d)](1 - U_k) + \\ &+ [\alpha/(P_k + L_d) - \alpha/(P_{k-1} + L_d)]U_k \end{aligned} \tag{5}$$

Combining case 1 with case 2, we get the recursive form as follows:

$$\begin{aligned} J_k = J_{k-1} &+ [\alpha/(P_k + L_d) - \alpha/(P_k + L_c)]U_k + \\ &+ [\alpha/(P_{k-1} + L_c) - \alpha/(P_{k-1} + L_d)]U_{k-1} \end{aligned} \tag{6}$$

where $k > 1$.
Therefore

$$J_k|_{U_k=0} = J_{k-1} + [\alpha/(P_{k-1} + L_c) - \alpha/(P_{k-1} + L_d)]U_{k-1} \tag{7}$$

$$\begin{aligned} J_k|_{U_k=1} = J_{k-1} &+ [\alpha/(P_k + L_d) - \alpha/(P_{k-1} + L_c)]U_k + \\ &+ [\alpha/(P_{k-1} + L_c) - \alpha/(P_{k-1} + L_d)]U_{k-1} \end{aligned} \tag{8}$$

DEFINITION 1 *Define Platt's reinforcement as follows:*

$$J_k|_{U_k=1} - J_k|_{U_K=0} = \alpha/(P_k + L_d) - \alpha/(P_{k-1} + L_c) \tag{9}$$

When $\alpha/(P_k + L_d) > \alpha/(P_{k-1} + L_c)$, choosing D is better than choosing C, but, personal profit is decreased as the number of choices increases. Choosing D is always better than choosing C, so long as $L_c - L_d > 1$.

DEFINITION 2 *When $L_c - L_d > 1$, form (1) or (6) is said to be a one-person trap.*

Of course, we can think of "one-person" as any individual decision maker, who could be a group.

PROPERTY 1 In the one-person trap, choosing strategy D is better than choosing strategy C at every step.

PROPERTY 2 In the one-person trap, the value of the profit function decreases as the number of times strategy D is chosen increases. Gradually, the curve D becomes nearly horizontal.

PROPERTY 3 In the one-person trap, the more times strategy D is chosen, the less the profit function is reduced. (i.e. the smaller

$$\alpha / \left(L_d + \sum_{i=1}^{k} U_i \right) - \alpha / \left(L_d + \sum_{i=1}^{k+1} U_i \right)).$$

The value of the profit function finally tends to zero.

We can explain these properties as follows:
In the one-person trap, we have $L_c - L_d > 1$. One is motivated by the reinforcement of $[\alpha/(P_k + L_d) - \alpha/(P_k + L_c)]$ at every step to choose strategy D and get into the trap. The more often strategy D is chosen, the smaller the reduction in value of the one-person profit function. 1001 boxes of cigarettes do as much damage to the body as 1000 boxes. The addiction is reinforced, and eventually serious damage occurs.

The point $(\sqrt{\alpha} - L_d, \sqrt{\alpha})$ of curve D is the vertex of the hyperbola D. When $0 < P_k < \sqrt{\alpha} - L_d$, the curve of the profit function makes a fast descent. When $\sqrt{\alpha} - L_d < P_k$, the curve makes a slow descent. Gradually, one becomes addicted to drugs. We have assumed $\alpha > L_c^2$, so we have definition 3.

DEFINITION 3 *The point $(\sqrt{\alpha} - L_d, \sqrt{\alpha})$ on the curve D is the addiction threshold.*

The larger L_d, the easier to become addicted; and the smaller α, the easier to become addicted. But note that $\sqrt{\alpha} > L_d$, i.e. $L_d/\sqrt{\alpha} < 1$, is required. There are different α and L_d for different drugs. It is easy to show property 4.

PROPERTY 4 As $L_d/\sqrt{\alpha}$ increases and approaches 1, it gets easier to become addicted.

DEFINITION 4 *$L_d/\sqrt{\alpha}$ is the addiction index.*

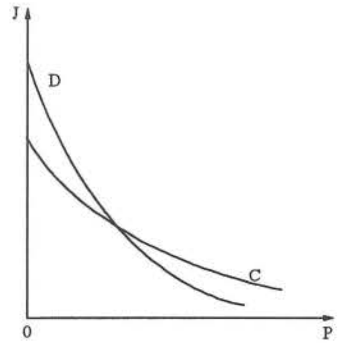

Fig. 2

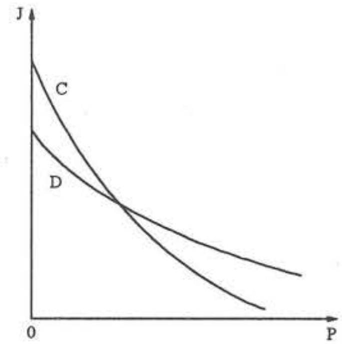

Fig. 3

The above mentioned two curves C and D do not intersect. We can also model two curves that do intersect (see Fig. 2 and Fig. 3).

$$J_k = \begin{cases} \alpha_c/(P_k + L_c) + \beta_c & U_k = 0 \quad (C) \quad \alpha_c, L_c, \beta_c > 0 \\ \alpha_d/(P_k + L_d) & U_k = 1 \quad (D) \quad \alpha_d, L_d > 0 \end{cases} \tag{10}$$

and

$$J_k = \begin{cases} \alpha_c/(P_k + L_c) & U_k = 0 \quad (C) \quad \alpha_c, L_c > 0 \\ \alpha_d/(P_k + L_d) + \beta_d & U_k = 1 \quad (D) \quad \alpha_d, L_d, \beta_d > 0 \end{cases} \tag{11}$$

Expression (10) represents the extricable trap, and expression (11) the drug trap (Wu, Lu and Zheng, (1990)). In (11), reinforcement tends toward β_d as the number of times strategy D is chosen increases.

3. Algorithmic models

Now we study the general case. Assume that f_c and f_d are the cooperation profit function and the defection profit function, respectively. They are assumed to be nonincreasing (linear or nonlinear).

Let

$$J_k = \begin{cases} f_c(P_k) & U_k = 0 \\ f_d(P_k) & U_k = 1 \end{cases} \tag{12}$$

$$P_k = \sum_{i=0}^{k} U_i$$

Model 1.

Step 1: $k = 0$; $U_0 = 0$; $P_0 = 0$; $J_0 = f_c(0)$;
Step 2: $k = k + 1$;
Step 3: If C is chosen, then $U_k = 0$; else $U_k = 1$;
Step 4: $P_k = P_k + U_k$;
Step 5: If $U_k = 0$, then $J_k = f_c(P_k)$; else $J_k = f_d(P_k)$;
Step 6: return to step 2.

Of course, model 1 is an infinite cycle. We can limit the number of steps, k, to force the process to stop. If cooperation, C, was chosen at the last step, choosing cooperation again at this step does not change the value of personal profit. In the model 2 we make the value of personal profit increase. One who chooses cooperation, C, again, will be rewarded with increased personal profit.
Model 2.

Step 1: $k = 0$; $U_0 = 0$; $P_0 = 0$; $J_0 = f_c(0)$;
Step 2: $k = k + 1$;
Step 3: If D is chosen, then $U_k = 1$ and go to step 5; else go to step 4;
Step 4: (a) If C was chosen at the previous step and $P_{k-1} > \eta$ or $P_{k-1} = 0$, then $U_k = 0$ and go to step 5; (The meaning of η is discussed below).
(b) If C was chosen at the previous step and $P_{k-1} \leq \eta$, then $U_k = -1$ and go to step 5;
(c) If D was chosen at the previous period, then $U_k = 0$;
Step 5: $P_k = P_k + U_k$;
Step 6: If $U_k = 1$ then $J_k = f_d(P_k)$ else $J_k = f_c(P_k)$;
Step 7: return to step 2.

For illustration (see Fig.1), consider fishing strategies. Assume that at the beginning one chose cooperating C, $U_0 = 0$, and obtained α/L_c of fish. This is not overfishing, so the remaining fish produce future generations. If one chooses cooperating C again at next step, $U_1 = 0$, from step 4(a) of model 2, and from step 5, $P_1 = P_0 + U_1 = 0$, so profit equals α/L_c of fish. If one chose defecting, $U_1 = 1$, one obtains more fish than by cooperating. Because choosing strategy D is better, one chooses it again and again. P_k will reach η and after that P_k will exceed η. At that time, if one chooses cooperating C for some reason, P_k is constant because of step 4(a). This means that fish resources are destroyed. The remaining fish cannot produce the original value α/L_c. So long as $P_k < \eta$, if one chooses cooperation repeatedly, P_k moves to the left because of step 4(b) of model 2. The remaining fish can provide the original value α/L_c.

Thus η is defined as the recovery threshold.

4. Conclusion

In this paper, we describe intertemporal social dilemmas using nonlinear models, and analysed several situations. A recurrence formula was given for profit in an intertemporal dilemma for any time interval. Two general models of a social dilemma are proposed. Model 2 provides a model of overuse of resources. An intertemporal nonlinear model might be much better, but this is a matter under study.

References

HARDIN G. (1968) The tragedy of the commons, *Science*, **166**, 1103–1107.

PLATT J. (1973) Social trap, *American Psychology*, **28**, 641–651.

AXELROD R., WILLIAM D.H. (1981) The Evolution of Cooperation, *Science*, **212**, 1390–1396.

WU J.Z., LU Z.Y. AND ZHENG Y.P. (1990) A New Game Theoretical Model of Social Dilemma, IFAC World Congress, Tallin, U.S.S.R.