2 The Complexity of Strategies and the Resolution of Conflict: An Introduction*

Ariel Rubinstein
THE HEBREW UNIVERSITY OF JERUSALEM
and LONDON SCHOOL OF ECONOMICS

1 INTRODUCTION

Whenever we have to choose a rule of behaviour we are confronted with the following dilemma: on the one hand we would like the rule to serve our goals and interests in the best possible way, and on the other hand we would like the rule to be as simple (as uncomplicated) as possible. As economists, we very often feel that it is unreasonable to allow economic agents to choose arbitrary rules of behaviour. Frequently we restrict the set of feasible rules by omitting those which are 'not simple enough'. However, it is only recently that economic theorists have begun to model explicitly the endogenous choice of complexity of rules of behaviour. It is the purpose of this paper to introduce the reader to some of these developments.

One can think of a variety of reasons for an economic agent's desire to reduce the complexity of his rules of behaviour. A more complex rule is more likely to break down; it is more difficult to train an agent who is supposed to carry it out; it may take more time to use and so on. But we do not address the sources of the advantages of simple rules in this paper. Rather, we just assume that complexity is 'costly'.

A rule of behaviour instructs an agent on what to do whenever he has to act. At each instance the action is an application of the rule. It is assumed that the agents are not restricted in their ability to choose the optimal rule of behaviour. Neither do they take into account the

*The paper follows Rubinstein (1986) and Abreu and Rubinstein (1986). My deep thanks to Dilip Abreu for his cooperation while working on our joint paper.

complexity of the 'meta-rule' by means of which they choose the optimal rule of behaviour. Thus we do not deal with the complexity of the optimal choice of the rule of behaviour.

From a wider perspective, we follow a direction of research advocated a long time ago by Herbert Simon (see, for example, Simon, 1957, 1978).[1] Simon argues that economists should expand the scope of their research to include questions of 'procedural rationality'. We have to deal not only with *what* decisions agents make but also with *how* they make them. The actual processes of decision making in organisations and the bounds of rationality of human beings should be represented in the formulation of the economic model. Although Simon has received worldwide recognition, his ideas have had a limited impact on mainstream economic theory. The reasons are quite clear to anyone who has tried to embed 'bounded rationality' ideas into economic theory. It is very difficult to formulate the decision making process. There is a sense of arbitrariness and *ad hoc*ness in any model in which the decision-making procedure is simply grafted onto a basically traditional model. I believe that in the absence of a more firmly established methodology, however, we can fruitfully address some issues of bounded rationality even in *ad hoc* models.

Our main aim is to explore the effect of introducing considerations of complexity of the rules of behaviour on the equilibrium outcomes of games. When the complexity of rules of behaviour is included in the model, is 'predicted behaviour' approximately the same as that in a standard model in which complexity is excluded?

2 THE BASIC MODEL

Since this chapter is supposed to be only an introduction to the topic I will restrict myself to a detailed discussion of a single example. Two players, she, player 1, and he, player 2, are involved in a long-term relationship. Every night she decides whether to date him or to date an alternative 'outside option' and he buys the tickets either for his favourite entertainment, 'Football', or for hers, 'Ballet'. She likes most meeting him at the Ballet (pay-off 3) and she dislikes Football most even if he is at her side (pay-off 1). The outside option is in between (pay-off 2). For him, the outside option is a dreadful event (pay-off 0). He likes dating her at the Football stadium (pay-off 3) and Ballet is a second best to not meeting her at all (pay-off 1).

Thus the basic (one shot) situation can be described by the following 2×2 game:

	Ballet	Football
Dating him	3, 1	1, 3
Outside option	2, 0	2, 0

This one-shot game has a unique equilibrium in which he buys football tickets and she chooses the outside option. This equilibrium is not Pareto Optimal since dating at the Ballet dominates the equilibrium outcome for both players.

The long-term relationship allows the players to settle on a better outcome. Assume that the game is repeated again and again at points of time 1, 2, 3 . . . *ad infinitum*. The long term situation is called a repeated game.[2] The basic idea is that in the repeated game, he might go to the Ballet in spite of his ability to make a short-term gain, because he fears her choosing the outside option were he to buy tickets for Football.

To be precise we have to spell out the long-term preferences and to explain what a long-term strategy is. An outcome of the repeated game is a sequence of one-shot outcomes. This sequence corresponds to a stream of one-shot pay-offs. It is assumed that the players are interested in the limit of the averages of their finite period pay-offs. (Discounting pay-offs does not lead to significantly different results.) Notice that in the repeated game the one-shot pay-offs have more than an ordinal meaning.

A long-term strategy is a plan of what action to choose at any one point in time based on the information gathered by the player up to that point. It is assumed that at the end of each period the players have perfect recall of the history of the relationship. Thus, a player may base his action at time t on the entire list of the preceding $t - 1$ outcomes.

Notice that player 1 can enforce on player 2 a level of pay-off 0 (by choosing the outside option). Player 2 can ensure that player 1 will not get more than 2 by buying Football tickets. These levels are called the min-max levels.

Let (U_1, U_2) be an arbitrary point in the triangle formed by the one-shot game pay-off vectors (3, 1) (1, 3) and (2, 0). Using the time dimension as a coordination device the players are able to form a sequence of one-shot game outcomes such that the averages of the pay-offs will converge to (U_1, U_2) (Figure 2.1).

The characterisation of the long-term pay-offs of equilibria for this repeated game is very simple. Take any sequence of one-shot outcomes in which the sequence of averages converges to a pair of utilities (U_1, U_2) such that $U_1 \geqslant 2$ and $U_2 \geqslant 0$ (that is, U_1 and U_2 are both above the

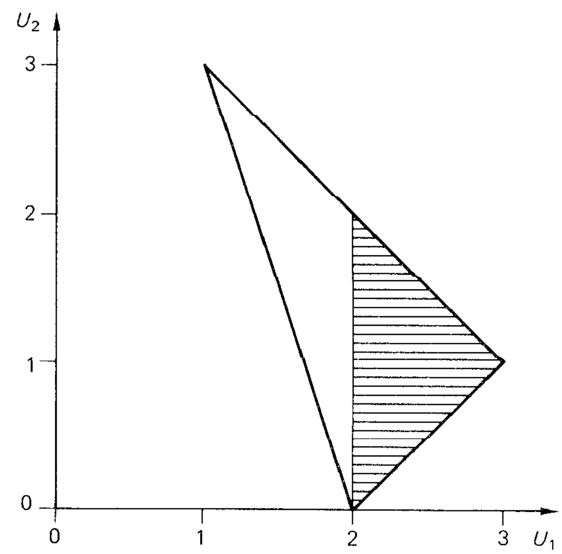

Figure 2.1 The pay-offs set

minmax levels). Consider the strategies that would follow this sequence in the event that no player deviated and would react to any deviation by playing the one-shot equilibrium for ever after. These strategies are called 'grim strategies'. They provide an equilibrium for the repeated game with the limit of the means. (It is easy to see that for all (U_1, U_2) where $U_1 > 2$ and $U_2 > 0$ there is a discount factor large enough for a pair of grim strategies to be a Nash Equilibrium with the discounted average pay-offs (U_1, U_2).)

The grim strategies are not the only Nash Equilibria of the repeated game. One could think of many other 'types' of such equilibria. For example, consider the arrangement to go to the Ballet on odd days and to Football on even days. The arrangement might be supported by his grim threat and by her threat that if he buys tickets for Football on a Ballet day she will punish him by choosing the outside option for n periods where n is the number of occasions on which he broke the arrangement in the past. Our intuition is that the latter equilibrium is much more complicated than the grim strategy equilibrium.

3 THE MACHINES GAME

We depart now from the traditional definition of a strategy. Henceforth a player is limited to carrying out his strategy by use of a 'machine' (finite automaton).[3] A machine for player 1 (and analogously for player 2) includes four elements:

1. A set of *States*. The set may be any finite set. The names of the states in the set are meaningless.
2. An *Initial State*. The initial state has to be one of the elements of the set of states.
3. An *Output Function* which specifies an action, either D or O, for every state. The interpretation of the output function is that whenever the machine is at a certain state the machine plays the one-shot action corresponding to the state.
4. A *Transition Function* which spells out how the machine moves from one state to another. At each period the machine receives as input the action that player 2 chose, namely either B or F and then it moves into a new state. The new state is determined by the transition function depending on the current state and the input received.

One can interpret the machine as a mechanical tool for carrying out a strategy. Less naively, it is possible to think of a machine as an abstraction of the process by which the repeated game rule of behaviour is implemented.

To demonstrate the concept of a machine, let us look at a few machines for player 1. The first machine carries out the 'grim' strategy. The machine starts at the state \$ in which it plays D and it stays there unless it observes the action F. Then the machine moves to the other state, which is an absorbing state (the machine stays there whatever player 2's action). At £ the machine plays the action O. Notice that the grim strategy can be carried out also by more complicated machines. Figure 2.2 shows the simplest machine needed to carry out the grim strategy.

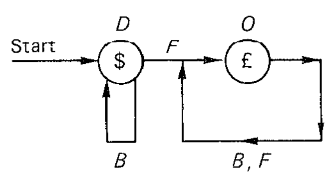

Figure 2.2 The grim strategy machine

In the next example, the machine alternates between playing D and O independently of player 2's moves (Figure 2.3):

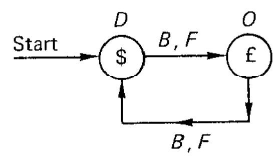

Figure 2.3 The alternating machine

The next machine is programmed to play D as long as the other player plays B and to play O for 3 periods if player 2 plays F while M_1 is at the state \$. Notice the need for at least 4 states to carry out this strategy (Figure 2.4).

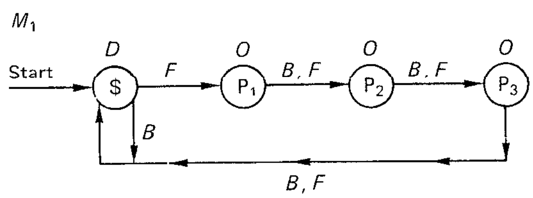

Figure 2.4 The three-period tit-for-tat machine

The last example is a machine for player 2 (Figure 2.5).

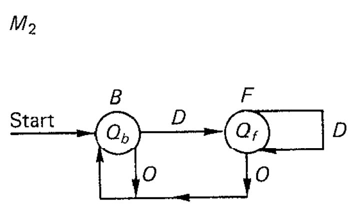

Figure 2.5 An example of a machine for player 2

The machine M_2 starts by playing B. Then player 1's D is followed by the state Q_f (and playing F) and O is followed by the state Q_b (and playing B).

Not every strategy can be executed by a finite machine. Consider for example player 1's strategy to play one D and one O, followed by one D

and two Os, followed by one D and three Os and so on independently of player 2's actions. This strategy cannot be carried out by a finite state machine.

We now move on to describe how the repeated game is played by a pair of machines. The description is demonstrated on the above machines M_1 and M_2.

The machines start at the states \$ and Q_b. The outcome of the first round of the repeated games is (D, B) since the output function of M_1 assigns the action D to the state \$ and the output function of M_2 assigns B to the state Q_b. The next period states are determined by the transition functions. The transition function of M_1 leaves the machine at \$ after it observes that player 2 played B. The transition function of M_2 transfers the machine from Q_b to Q_f as a response to the input D. Thus period 2's pair of states is (\$, Q_f). The output functions determine period 2's outcome (D, F) and M_1 now moves from \$ to P_1 while M_2 stays at Q_f.

time	M_1's state	M_2's state	the one-shot outcome	pay-offs
1	\$	Q_b	(D, B)	(3, 1)
2	\$	Q_f	(D, F)	(1, 3)
3	P_1	Q_f	(O, F)	(2, 0)
4	P_2	Q_b	(O, B)	(2, 0)
5	P_3	Q_b	(O, B)	(2, 0)
6	\$	Q_b		

At the sixth period the pair of states are the same as at period 1 and then the play of the repeated game starts to repeat itself. Because of the finiteness of the set of states and the Markovian structure of the machines all pairs of machines must eventually enter into a cycle, although not necessarily immediately, as in the above example.

4 COMPLEXITY

Although not all repeated game strategies can be executed by a finite machine, the restriction to strategies which can be is not significant in itself. It is made here to enable us to make the next conceptual departure from the conventional repeated game literature. We are about to include explicitly the complexity of a strategy in the players' optimising calculations.

First we have to define the term 'complexity of a machine'. There are many possibilities available. We will make do with a very naive and simple measure of complexity. *The complexity of a machine* is defined by *the number of states* in the machine. Thus the complexity of the transition function is ignored.

Given a pair of machines a player gets a stream of pay-offs which is evaluated by him according to the limit of the averages. Notice that the limit of the averages is always well-defined since the sequence of pairs of machine states must eventually enter a cycle. The limit of the means of a player's pay-offs is equal to the average of the player's pay-off in the cycle.

In an equilibrium of the machine game a player chooses a 'best response' against the other player's machine. A player's preference depends only on two numbers:

1. the repeated game pay-off, and
2. the machine complexity.

The preference relation is assumed to be monotonic in the two numbers: increasing with the repeated game pay-off and decreasing with the complexity. Sometimes we will be interested in the model where the preferences are lexicographic, in the sense that a player's first priority is the repeated game pay-off, and only secondarily does a player care about the complexity of the machine. The model with lexicographic preferences is the 'closest' possible model to the standard model without complexity.

5 EXAMPLES

In this section we will look at two examples of pairs of machines: in the first example her machine, M_1, is the two-state machine which carries out the 'grim strategy' and his machine, M_2, is the one-state machine which plays B (see Figure 2.6). The machine M_2 is a best response against the machine M_1. Even by using a more complex machine, player 2 cannot achieve a higher repeated game pay-off. Player 1 cannot achieve a higher repeated game payoff but she is able to reduce the number of states in her machine by dropping £ without reducing the repeated game pay-off. Given M_2, the state £ is used only to threaten player 2. But in equilibrium the threat is redundant and player 1 can omit £. Thus, this pair of machines is not a Nash Equilibrium in the machine game. The general conclusion we note is that *in equilibrium all states must be used at least once.*

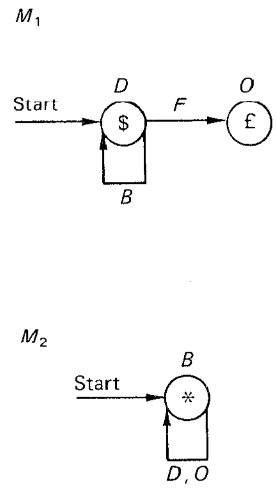

Figure 2.6 A pair of machines which are not in equilibrium

The next pair of machines is a Nash Equilibrium in the machine games if the players do not give too much weight to complexity (see Figure 2.7).

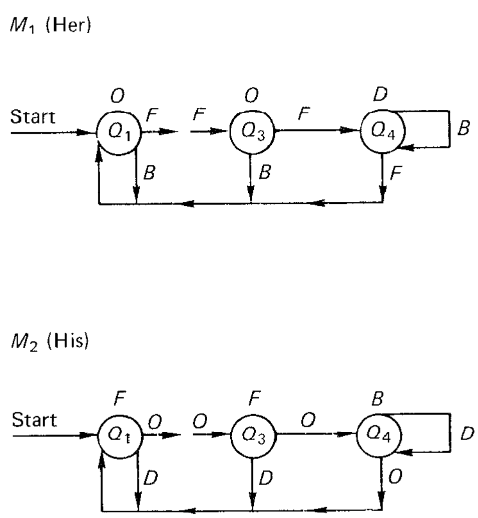

Figure 2.7 An example of an equilibrium with the outcome (D, B)

Here the players start by 'showing' their ability to punish. After the 'display of threat' the players move to the more 'cooperative' phase where they date for the Ballet. After they reach the 'cooperative' phase they punish a deviator by moving back to the initial state. The length of the cycle in the play of the machine game is 1 and the length of the 'introductory period' is 3 (she needs 3 periods to erase his gain from playing F instead of B). If player 1 used only a one- or two-state machine he could achieve a limit of averages of at most 2. The machine M_1 is her best response if the pay-off 3 with complexity 3 is preferred by her to the pay-off 2 with complexity 1. As to player 2, one-, two- or three-state machines will give him a repeated game pay-off of at most 0.75, and if the pay-off 0.75 with complexity 1 is not preferred to the pay-off 1 with complexity 4 then M_2 is a best response against M_1.

6 THE STRUCTURE OF EQUILIBRIUM IN THE MACHINES GAME

We are able to derive several conclusions about the structure of machines in equilibrium. The following properties are true in general for all repeated games. A full presentation of the results and proofs appears in Abreu and Rubinstein (1986).

First, notice that in equilibrium *the number of states in both machines must be equal*. The reason for this is that a Markovian problem with m states has an optimal stationary solution. Given player i's machine of size m, player j does not need more than an m-state machine to achieve the best repeated game pay-off.

We have noticed already that all the states in the machines must be used at least once in the course of playing the game. Some of the states appear in the cycle and some do not. It can be shown that *all non-cycle states are used only once and appear consequently in the beginning of the play* of the game. After that, only cycle states are used (although a deviation of the opponent may activate a non-cycle state). Finally *the length of the cycle is equal to the number of cycle states*. In other words a state does not appear twice in the cycle.

Thus, in equilibrium, during the introductory phase before cycle states are used, and in the cycle itself there is a one-to-one correspondence between his and her states. This means that in equilibrium in any period one machine 'knows' the state that the other machine is in at the same time (except during a possible intermediate phase after the introductory phase and before the cycle begins).

These results have a dramatic consequence for the set of Nash Equilibrium outcomes in the machine game. In any equilibrium the one-shot outcomes are either (D, B) and (O, F) only, or (D, F) and (O, B) only. Therefore if there is a period in which the couple dates at the Ballet they never date at the Football stadium. In equilibrium there is a one-to-one correspondence between player 1's actions and player 2's actions. The proof of this assertion is beyond the scope of this paper but I would still like to provide some intuition by examining the following pair of machines (see Figure 2.8): The play of the game by this pair of machines starts with

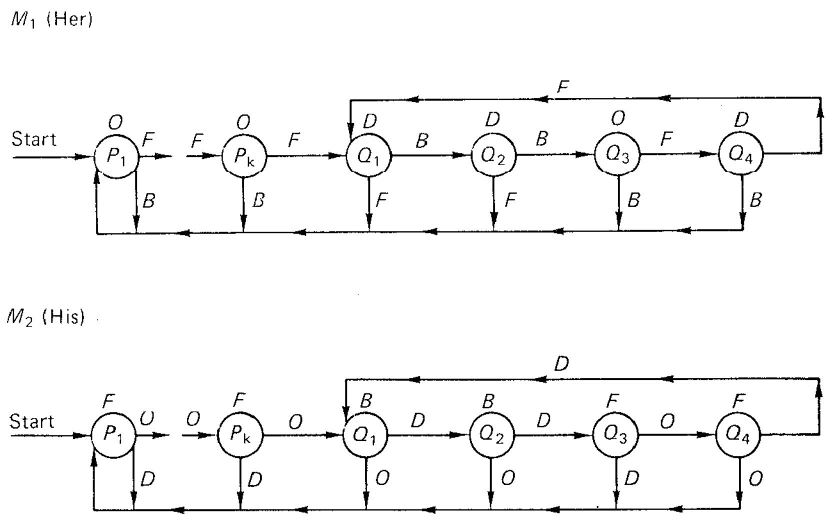

Figure 2.8 Another pair of machines which is not an equilibrium

K periods of 'threat display' (K 'large' enough). Then the machines enter into a cycle of length 4 in which the outcomes are (D, B), (D, B), (O, F) and (D, F). Any deviation from the pattern of behaviour in the cycle causes the opponent's machine to move to its starting point and by doing so to inflict a punishment of K periods on the deviator. The average payoffs are 2.25 for player 1 and 1.25 for player 2. It is easy to verify that the machines form a Nash Equilibrium in the repeated game with the limit of the means but do not form an equilibrium in the machine game. Inspect M_1. Player 1 needs Q_1, Q_2 and Q_4 in order to know when to play the outside option, O. However she can execute the procedure of counting up to

3 without Q_1. The output of player 1's machine at Q_1 and at Q_4 is the same (D). However when M_1 is at Q_1 the machine M_2 plays B and when M_1 is at Q_4 the machine M_2 plays F. She can save one state by replacing Q_1 and Q_4 by one state that does the same job as the two states do in M_1.

The revised machine is shown in Figure 2.9.

The New M_1

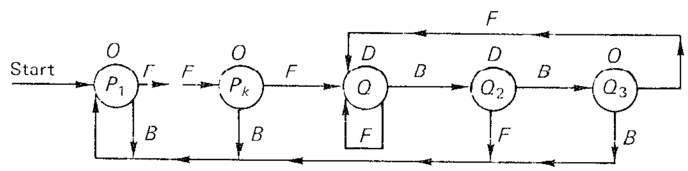

Figure 2.9 A profitable deviating machine (for the Figure 2.8 example)

By using the new machine player 1 does not change the sequence of played outcomes, yet her machine is more 'economical' than the original M_1. The pair comprising the new machine and M_2 is not an equilibrium since player 2 may use the fact that player 1 does not monitor his behaviour when he is at state Q and he may deviate profitably to the $K+1$ state machine that forms the one-period cycle with the outcome (D, F).

7 A CHARACTERISATION OF THE EQUILIBRIUM OUTCOMES

In the previous section a result was given where the set of outcomes which appear in the play of any equilibrium must be either a subset of $\{(D,B), (O,F)\}$ or a subset of $\{(D,F), (O,B)\}$. It is easy to exclude the possibility that an equilibrium play in the machines game includes the outcomes (D, F) and (O, B) only. For any pair of machines in which the outcomes are only from among these elements, player 1's repeated game pay-off is at most 2. She is able to achieve the repeated game pay-off 2 by a one-state machine which plays O. Thus if M_1 is not this one-state machine player 1 can deviate profitably. If M_1 is the one-state machine which plays O then player 2's machine must be the one-state machine which plays B and then player 1 can deviate profitably to the one-state machine which plays D. Therefore in equilibrium the one-shot outcomes played must include (D,B) and (O,F) only.

The set of equilibrium outcomes has now been reduced dramatically. The exact characterisation of equilibrium outcomes depends on the players' tradeoff between the repeated game pay-off and the complexity of the machines. In the case of lexicographic orderings (when the players' consideration of complexity is only secondary to the repeated game pay-off) for any two integers m and n there is an equilibrium with a cycle of length $m + n$ in which (D, B) appears m times and (O, F) appears n times. To demonstrate this result we will look at the next pair of machines in which the players alternate in the cycle between (D, B) and (O, F), $(m = n = 1)$ (see Figure 2.10).

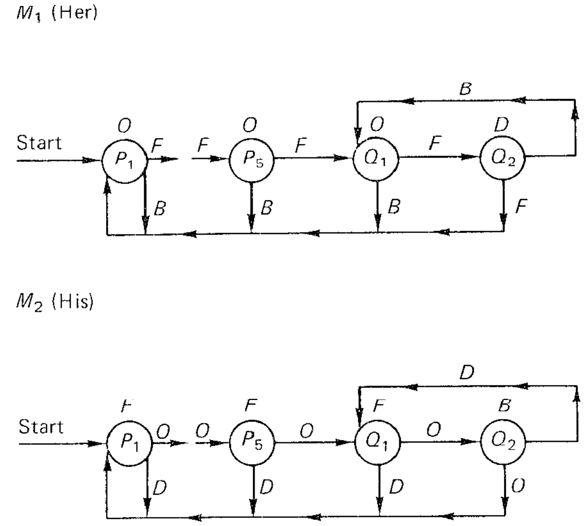

Figure 2.10 An equilibrium with a combination of outcomes (D, B) and (O, F)

Notice the following:

1. A deviation by player 2 during the cycle is punished for 5 periods, which is long enough to offset his gain from the deviation.
2. If the order of the outcomes in the cycle is reversed (first playing (D, B) and then playing (O, F)) then the pair of machines ceases to be an equilibrium and each of the players can save a state by omitting P_5 and at P_4 transiting to Q_2 as the response to O or to F.
3. The length of the punishment depends on the mixture of outcomes we would like to sustain, that is on m and n.

4. It can be shown that it is impossible to sustain an equilibrium in which (D, B) is one of the outcomes without having an introductory phase to the play of the game.

For more general preference relations the above is an equilibrium in the machines game if player 1 prefers a repeated game pay-off 1.5 with complexity 7 to the pay-off 2 with complexity 1 and if player 2 prefers a pay-off 0.5 with complexity 7 to a pay-off 3/7 with complexity 1.

8 FINAL REMARKS

The complexity of behaviour in a repeated game has three major components:

1. The complexity of the routine. A path of outcomes in which the players have to change their actions in the cycle in an 'irregular' order requires the machines to have many states.
2. The complexity of punishment. The need to threaten the opponent with the carrying out of a punishment if he deviates might require holding extra states.
3. The complexity of monitoring. A player might need states for monitoring the behaviour of the opponent.

In the above analysis, due to the complexity of punishment the players do not hold special states for punishing the opponents, and the complexity of monitoring prevents equilibria in which the couple will date at every period and the man will switch back and forth from Ballet to Football.

In the repeated game, mixtures of (D, B) and (D, F) could be sustained in equilibrium. In the machine game any non-degenerate mixture of (D, B) and (D, F) requires player 2 to make changes in his actions from B to F and from F to B. An arrangement whereby she monitors him is unstable because it requires her to maintain special states. Relying on him is unstable because he might gain by changing the mixture of B and F without being detected.

To summarise, I have tried to introduce the reader to a new model in which the complexity of a strategy is included explicitly in the players' consideration. Since there are *ad hoc* and arbitrary assumptions in the model I would hesitate to regard the model and the results as a new theory of repeated games. It is probably better to think of the work in its current stage as a modelling exercise. However, the topic is fundamental and the possible implications dramatic, so I am quite confident that it will continue to attract attention and that we will be seeing exciting results in the very near future.

Notes

1. For an introduction to Simon's ideas on 'Bounded Rationality' see for example Simon (1972) and Simon (1978). A pioneering work in the direction of connecting 'Bounded Rationality' with economic theory is Radner and Rothschild (1975).
2. For introductions to the literature of repeated games with perfect information see Aumann (1981), Rubinstein (1979) and Abreu (1983). For early attempts to use automata in economic theory see Futia (1977), Gottinger (1983), Marschak and McGuire (1971) and Varian (1975). For attempts to use 'Bounded Rationality' ideas to recover from the 'paradoxical' results in finitely repeated games see Radner (1986), Smale (1980) and Green (1982). Green (1982) is the closest in spirit to the work reported here. Recent works using Finite Automata to discuss finitely repeated games include Ben-Porath (1986), Lehrer (1986), Megiddo and Wigderson (1986) and Neyman (1985). Another related work is Kalai and Stanford (1986).
3. For a textbook on Automata Theory see Hopcroft and Ullman (1979). The idea of using finite automata in the repeated game context was first suggested in Aumann (1981).

References

Abreu, D. (1983) 'Repeated Games with Discounting', PhD Thesis, Princeton University.

Abreu, D. and A. Rubinstein (1986) 'The Structure of Nash Equilibrium in Repeated Games with Finite Automata', London School of Economics and Harvard University.

Aumann, R. J. (1981) 'Survey of Repeated Games' in *Essays in Game Theory and Mathematical Economics in Honour of Oskar Morgenstern* (Mannheim: Bibliographisches Institut), 11–42.

Ben-Porath, E. (1986) 'Repeated Games with Bounded Complexity', mimeo.

Futia, C. (1977) 'The Complexity of Economic Rules', *Journal of Mathematical Economics* **4**, 289–99.

Gottinger, H. W. (1983) *Coping with Complexity* (Dordrecht: Reidel).

Green, E. (1982) 'Internal Costs and Equilibrium; The Case of Repeated Prisoner's Dilemma', mimeo.

Hopcroft, J. E. and J. D. Ullman (1979) *Introduction to Automata Theory, Languages and Computation* (Reading, Mass: Addison Wesley).

Kalai, E. and W. Stanford (1986) 'Finite Rationality and Interpersonal Complexity in Repeated Games', Northwestern University.

Lehrer, E. (1986) 'Repeated Games with Stationary Bounded Recall Strategies', mimeo.

Marschak, T. A. and C. B. McGuire (1971) 'Economic Models for Organization Design', unpublished lecture notes.

Megiddo, N. and A. Wigderson (1986) 'On Play by Means of Computing Machines', mimeo.
Neyman, A. (1985) 'Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoners' Dilemma', *Economics Letters* **19**, 227–9.
Radner, R. (1986) 'Can Bounded Rationality Resolve the Prisoners' Dilemma?', in A. Mas-Colell and W. Hildenbrand (eds), *Essays in Honour of Gerard Debreu*, (Amsterdam: North-Holland).
Radner, R. and M. Rothschild (1975) 'On the Allocation of Effort', *Journal of Economic Theory* **10**, 358–76.
Rubinstein, A. (1979) 'Equilibrium in Supergames with the Overtaking Criterion', *Journal of Economic Theory* **21**, 1–9.
Rubinstein, A. (1986) 'Finite Automata Play the Repeated Prisoner's Dilemma', *Journal of Economic Theory* **39**, 83–96.
Simon, H. A. (1957) *Models of Man* (New-York: Wiley).
Simon, H. A. (1972) 'Theories of Bounded Rationality' in C. B. McGuire and R. Radner (eds), *Decision and Organization* (Amsterdam: North-Holland).
Simon, H. A. (1978) 'On How to Decide What to Do', *Bell Journal of Economics* **9**, 494–507.
Smale, S. (1980) 'The Prisoner's Dilemma and Dynamical Systems Associated to Non-Cooperative Games', *Econometrica* **48**, 1617–34.
Varian, H. R. (1975) 'Complexity of Social Decisions', mimeo.