

Colman, A. M. (1998). Rationality assumptions of game theory and the backward induction paradox. In M. Oaksford & N. Chater (Eds), *Rational Models of Cognition* (pp. 353-371). Oxford: Oxford University Press.

## **Rationality Assumptions of Game Theory and the Backward Induction Paradox**

**Andrew M. Colman**  
*University of Leicester*

### **1. Backward induction**

Backward induction is a form of logical reasoning, based on the technique of mathematical induction, that is applicable to certain sequential games. I shall illustrate it with a simple version of the ancient game of nim. There are 20 matches on the table, and two players take turns in removing either one or two matches, the winner being the player who takes the last one, leaving the other player unable to move. Ernst Zermelo (1912) proved the first major theorem of game theory to the effect that, in games of this type, either one player or the other must have a guaranteed winning strategy. In this case, which player wins, and how?

The player who moves first wins. The game is easily solved by backward induction. We begin at the end: if it is my turn to move and there are either one or two matches left, then I can obviously win, because in either case I can immediately take the last match, leaving my opponent unable to move. Therefore, if it is my opponent's move and there are three matches left, I can win, because my opponent must take either one or two and will therefore have to leave either two or one, whereupon I can win immediately. Therefore, if it is my opponent's turn to move and there are six matches left, I can win, because my opponent will have to leave either four or five, and in either case I can then leave three, which as I have shown ensures that I can win. Therefore if it is my opponent's turn to move and there are nine matches left, I can win, because my opponent will have to leave either seven or eight, and in either case I can then leave six, which as I have shown, means that I can leave three on my next move, and this ensures that I can win. Continuing in the same vein, if it is my opponent's turn to move and there are 12 or 15 or 18 left, I can win, because whatever happens I can eventually leave three whatever my opponent does, and then I can win on my next move. So if there are 20 matches to start with and I move first, then I have a guaranteed winning strategy, which involves taking two matches on the first move, leaving 18, and then on successive moves leaving 15, 12, 9, 6, and 3 matches for my opponent, and finally taking the one or two that remain on my last move.

**THEOREM 1.** *In a two-person game of nim in which players take turns in removing either one or two matches from a pile until there are none left, the winner being the player who takes the last match, the player who moves first has a winning strategy if the initial number of matches is  $1 \pmod{3}$  or  $2 \pmod{3}$ , and the player who moves second has a winning strategy if the initial number of matches is  $0 \pmod{3}$ .*

*Proof.* The proof proceeds by induction on the number of matches. Call the players  $i$  and  $j$ . If it is Player  $i$ 's turn to move and there are three matches left, then Player  $j$  has a winning strategy, because Player  $i$  will have to take one or two, leaving either two or one, and in either case Player  $j$  can win immediately. For the inductive step, suppose that it is Player  $i$ 's turn to move with  $s$  matches left and Player  $j$  has a winning strategy. Then if it is Player  $i$ 's

turn to move with  $s + 3$  matches left, Player  $j$  has a winning strategy, because if Player  $i$  removes one match, Player  $j$  can remove two, and if Player  $i$  removes two matches, Player  $j$  can remove one, in either case leaving  $s$  matches on Player  $i$ 's next move. Thus, if it is Player  $i$ 's turn to move, Player  $j$  has a winning strategy if  $s = 0 \pmod{3}$ . If it is Player  $i$ 's turn to move and  $s = 1 \pmod{3}$  or  $s = 2 \pmod{3}$ , then Player  $i$  has a winning strategy that involves leaving  $s = 0 \pmod{3}$  matches at each of Player  $j$ 's moves. Therefore, the player who moves first has a winning strategy if the number of matches at the start is  $1 \pmod{3}$  or  $2 \pmod{3}$ , and the player who moves second has a winning strategy if the number of matches at the start is  $0 \pmod{3}$ .  $\square$

In Anderson's (1991, pp. 481-482) rational analysis of the process of solving problems involving sequences of steps, which I take to include problems involving backward induction, the solver's (subjective) probability that a given sequence of steps will achieve the desired goal is the product of the probabilities that the component steps will have their intended consequences, conditioned on the success of the prior steps. In this essentially probabilistic analysis, it is not necessarily rational to plan a long sequence of steps if there is a risk of the sequence diverging from the intended path at an early stage, and in any event the information-handling capacity of working memory places a limit on the number of steps that can be kept in mind at once. But the whole point of mathematical induction is to enable a problem solver to analyse the consequences of a whole series of steps without having to keep them in mind. According to Anderson's general framework for rational analysis of problem solving (see also Anderson, 1990), human cognition is optimized to the environment, and a problem solver searches among potential plans that might achieve a solution. If the value of the solution goal is  $G$ , and every plan has a subjective probability of success  $P$  and a cost  $C$ , the problem solver selects the plan that maximizes the expected value  $PG - C$  and implements the plan provided its expected value is positive. Mathematical induction greatly reduces the cost  $C$  of solving certain types of problems involving sequences of steps, and this ought to increase its expected value in sequential games such as nim.

In the paragraphs that follow I shall discuss the application of backward induction to game theory. The choices of human decision makers in sequential games are not always rational in the sense of following the prescriptions of backward induction. According to Anderson's (1991, p. 483) analysis, we should not expect problem solvers to make optimal moves in all cases, because the theory predicts that they will stop searching for optimality and settle for satisficing moves (moves that are satisfactory or that suffice) if the expected value of the search falls below a threshold. Game theory, on the other hand, conventionally assumes that players are perfectly rational and knowledgeable. As I shall show, the game-theoretic knowledge and rationality assumptions are incoherent, but an attempt to eliminate the incoherence by replacing strict logical reasoning with common-sense (nonmonotonic) reasoning leads inexorably to a contradiction.

## 2. Application to game theory

Figure 1 shows the well known Prisoner's Dilemma Game (PDG). Player I chooses between row  $C$  and row  $D$ , Player II simultaneously chooses between column  $C$  and column  $D$ , and the pair of numbers in each cell are the payoffs to Player I and Player II respectively for the corresponding pair of strategy choices.

		II	
		C	D
I	C	3, 3	1, 4
	D	4, 1	2, 2

Figure 1. Prisoner's Dilemma Game (PDG)

It is generally agreed, among game theorists and decision theorists at least, that in a PDG rational players will choose their defecting strategies, which are labelled *D* in Figure 1. The reason for this is that the *D* strategies are strictly dominant for both players, in the sense that a player receives a higher payoff by choosing *D* than by choosing *C* whatever the other player chooses---whether the other player chooses *C* or *D*. It is true that both players would be better off if they both chose their cooperative *C* strategies, and therein lies the dilemma of the PDG---in technical terminology, the *DD* outcome is Pareto non-optimal---but the *D* strategies are nevertheless the uniquely rational choices in a one-shot PDG.

How should rational players behave in the finitely iterated PDG, that is, in a finite sequence of *n* PDGs played between the same players? Luce and Raiffa (1957, pp. 97-102) put forward a powerful argument that rational players will choose *D* at every move, irrespective of the size of *n*, provided that both players know *n* in advance. The argument is based on backward induction. Suppose the players know that the game is to be iterated exactly twice. Both players will notice that their *D* strategies are strictly dominant in a one-shot version of the game. This implies that the outcome of the second round is bound to be *DD*, because it is in effect a one-shot PDG, given that there are no moves to follow and thus no indirect effects to consider. But if the outcome of the second round is perfectly determined, then the first round is also in effect a one-shot PDG, because influences on the following round can be discounted. Therefore both players will choose *D* on both rounds. This backward induction argument generalizes straightforwardly to any finite number of iterations.

Even before the backward induction argument for sequential games was formalized by Aumann (1995), its validity had been widely accepted, although the conclusion seems counter-intuitive in the finitely iterated PDG, given that both players would do better in the long run if they both played more cooperatively. In fact, Luce and Raiffa (1957) themselves seem to have had difficulty believing their own proof. They confessed that 'If we were to play this game we would *not* take the [*D*] strategy at every move!' (p. 100, italics in original), and they commented that choosing *D* on every move 'is not "reasonable" in the sense that we predict that most intelligent people would not play accordingly' (p. 101), although they offered no persuasive justification for this opinion.

Since the late 1950s, well over 1000 experiments designed to investigate the behaviour of players in the PDG and related games have appeared in print (see Colman, 1995, pp. 134-160 for a comprehensive review). The evidence shows that in finitely iterated PDGs, players seldom conform to the prescriptions of the backward induction argument. What is typically observed (e.g., Andreoni & Miller, 1993; Cooper, DeJong, & Forsythe, 1996; Rapoport & Chammah, 1965) is a substantial proportion of *C* choices, often exceeding 30 per cent, even

among players who have clearly understood the strategic properties of the game. A substantial minority of players choose  $C$  even on their terminal moves. There is evidence, however, that experience with finitely iterated PDGs tends to lead to an increase in  $D$  choices (Selten & Stoecker, 1986). Possible explanations for violations of backward induction in the PDG will be discussed later.

### 3. Chain-store game

The problem at the heart of backward induction was brought to prominence through the work of Reinhard Selten (1978) on the multi-person Chain-store game. Here is a tidied-up version of it. A chain-store has branches in 20 cities, and in each city it faces one potential challenger in the guise of

another store selling the same goods. The challengers decide one by one in order,  $A, B, \dots, T$ , whether to enter the market in their respective cities, and whenever one of them decides to enter, the chain-store responds with either a *predatory pricing* move involving slashing the prices of its products to prevent the challenger from building up a customer base, or a *cooperative* move of sharing the customers with the challenger, as shown in Figure 2.

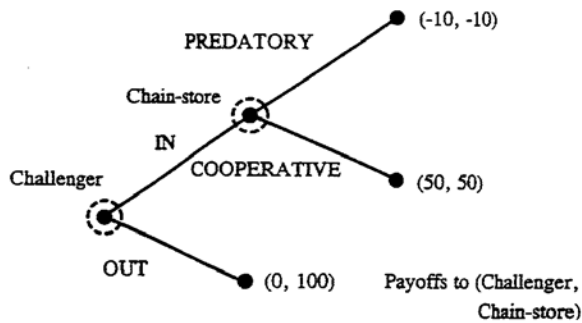


Figure 2. Chain-store game

Figure 2 shows a subgame involving the Chain-store and a single challenger. Beginning at the initial decision node circled and labelled Challenger on the left of the figure, Challenger first chooses whether or not to compete with Chain-store, so Challenger's two possible moves (represented by arcs) are labelled OUT and IN. If Challenger chooses OUT, then the terminal node at the bottom of the figure is reached: the subgame involving those two players ends and the pair of numbers in parentheses at the terminal node show the payoffs to the Challenger and the Chain-store respectively. Thus if Challenger stays out of the market, Challenger's payoff is (obviously) zero and Chain-store's is 100 units, representing maximum profits in that city. If Challenger chooses IN, then the next decision node (circled and labelled Chain-store) is reached: Chain-store responds with either COOPERATIVE or PREDATORY. If Challenger chooses IN and Chain-store responds with COOPERATIVE, they end up with equal market shares and profits, so that each receives a payoff of 50, shown in parentheses at the second terminal node. Finally, if Challenger chooses IN and Chain-store responds with PREDATORY, then both players suffer small losses, equivalent to 10 per cent of the profits that Chain-store could have expected in that city if unchallenged, so that each player receives a payoff of -10 units, as shown at the final terminal node.

After Challenger  $A$  has made a move and Chain-store has responded (if Challenger  $A$ 's choice was IN), it is Challenger  $B$ 's turn to choose IN or OUT, and Chain-store responds according to the same game tree, and so the game continues until the twentieth challenger  $T$

has had a turn and Chain-store has responded if necessary.

It seems intuitively obvious that the chain-store has a strong motive to respond aggressively to the early challengers in order to deter the later ones. Selten (1978) called this *deterrence theory*. But according to a simple backward induction argument, this theory is wrong and any attempt at deterrence is futile. The backward induction argument begins with the situation after 19 of the 20 challengers have made their moves, and the last one, Challenger *T*, is ready to move. Irrespective of what has gone before, it is obvious that Challenger *T* will choose IN, because in that case Chain-store will choose COOPERATIVE--at that stage there will be no rational reason for Chain-store to choose PREDATORY, because that would give Chain-store a worse payoff and there is no future challenger to deter. Challenger *T* knows that Chain-store will not respond to IN with PREDATORY, so Challenger *T* will choose IN, because IN followed by COOPERATIVE from Chain-store gives Challenger *T* a much better payoff than OUT. Challenger *T* will enter the market and Chain-store will respond with COOPERATIVE in this terminal subgame. This conclusion is not affected by what Challengers *A*, *B*, ..., *S* have done at earlier stages of the Chain-store game, nor by how Chain-store responded to those earlier moves.

Now consider the stage at which the 19th challenger, Challenger *S*, is about to make a move. If *S* were the last challenger, then it would be rational for *S* to choose IN, because Chain-store would be sure to respond with COOPERATIVE, as we have seen. The only difference now is that it is not the last challenger: Challenger *S* may think that Chain-store will choose PREDATORY in order to deter the challenger to follow on the next round. But we have already seen that predatory pricing will not deter the challenger to follow, so Challenger *S* will choose IN and Chain-store will respond with COOPERATIVE. It follows that the 18th challenger, Challenger *R*, who knows all this and is rational, will also choose IN, and in that case Chain-store will also respond with COOPERATIVE, and by backward induction this applies to every challenger including the first, Challenger *A*. This leads to the conclusion that Challengers *A*, *B*, ..., *T* will all choose IN and in each case Chain-store will respond with COOPERATIVE. The inescapable conclusion from the backward induction argument is that every challenger will choose IN and Chain-store will always respond with cooperative rather than predatory pricing moves. This combination of strategies or outcome is the so-called subgame perfect equilibrium of the game. In an echo of Luce and Raiffa's (1957) loss of confidence in their earlier backward induction argument in the PDG, Selten found his own conclusion hard to swallow:

If I had to play the game in the role of [the chain-store], I would follow deterrence theory. I would be very surprised if it failed to work. From my discussions with friends and colleagues, I get the impression that most people share this inclination. In fact, up to now I met nobody who said that he would behave according to induction theory. My experience suggests that mathematically trained persons recognize the logical validity of the induction argument, but they refuse to accept it as a guide to practical behavior. (Selten, 1978, pp. 132-133)

One hesitates to criticize someone who has won a Nobel Prize for his work on game theory, but one does not hesitate for long. Either the backward induction argument is valid, in which case one should accept its conclusions, or it is invalid, in which case one should explain why the conclusions are not worthy of belief.

This is not merely an abstract puzzle: it is important to know whether deterrence works. In penology, it is sometimes argued that the threat of capital punishment, even if it is never a rational choice in any *particular* case, is an effective deterrent to future criminals. In some countries, including Iran and China, capital punishment is used prodigiously. Other examples

of the Chain-store game that spring to mind include various versions of the doctrine of military deterrence, numerous economic situations in which a series of aspirant challengers take on a monopolist, political conflicts in which underground political groups try to wrest concessions from governments by planting bombs or taking hostages, and criminal intrigues in which potential blackmailers consider trying to extort money from wealthy individuals or companies. In each of these cases, provided that the situation recurs a finite number of times, the payoff structure of the game is clear enough for the backward induction argument to apply, at least apparently, and therefore for the paradox to emerge.

Although there are no published experiments on choices in the Chain-store game, there is abundant anecdotal evidence that people do sometimes follow the prescription of backward induction in real-life situations with the corresponding strategic structure. Even in the animal kingdom, a dominant individual often defends itself against a series of challengers by cooperative *conventional fighting* rather than predatory or escalated fighting (Dawkins, 1989; Krebs & Davies, 1991; Lazarus, 1982, 1987, 1994). A large carnivore usually refrains from attacking a challenger as ferociously as it is capable of doing when in predatory mode. A dominant bighorn ram leaps at an adversary head-on rather than charging its flanks, which would cause much more damage. A male fiddler crab defending a burrow never injures a challenger, although its huge claws are powerful enough to crush the challenger's abdomen. A dominant rattlesnake wrestles with a challenger but never bites; a dominant male deer locks antlers with a challenger rather than piercing its body; in some species, a dominant antelope actually kneels down while engaging in combat with a challenger. In each of these cases and many others besides, the animals are in effect responding cooperatively rather than in a predatory manner to serial challenges to their possession or control of resources such as food, territory, or mates, and they appear to follow the logic of backward induction rather than forward deterrence.

#### 4. Centipede game

The paradoxical quality of backward induction is exposed most vividly in the two-person Centipede game, first investigated by Robert Rosenthal (1981). A simple version of it has the following rules: on successive trials two players alternate in choosing whether to stop the game or to continue it. If a player stops the game, then there are no payoffs to either player at that point. But whenever a player chooses to continue, that player is fined £1 and the other player is rewarded with £10. This is a Centipede game with linearly rather than exponentially increasing payoffs, and I have interpreted it in a particular way. For different payoff parameters and alternative interpretations, see Aumann (1995); Binmore (1987); El-Gamal, McKelvey, & Palfrey (1993); Kreps (1990); McKelvey and Palfrey (1992); and Reny (1986, 1992).

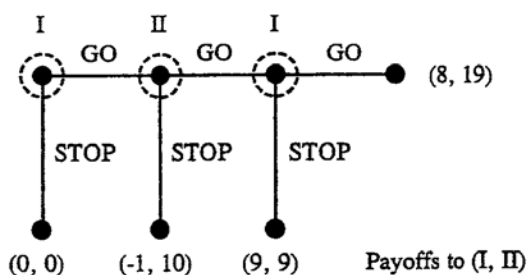


Figure 3. Centipede game

The sequence of moves runs from left to right in Figure 3. As in Figure 2, each decision node is circled and labelled with the player who makes the decision, and each terminal node is labelled in parentheses with the payoffs to Player I and Player II in that order if the game stops there. To begin, Player I chooses whether to STOP the game, in which case both players receive zero payoffs, or to GO, in which case Player I is fined £1 and Player II receives £10 and the game continues (Figure 3 shows only the final payoffs at terminal nodes). Then it is Player II's turn to move. Player II can choose to STOP the game, in which case no further payments are made and final payoffs are -1 to Player I and 10 to Player II, or to GO, in which case Player II is fined £1 and Player I receives £10 and the game continues. Then it is Player I's turn to make another move. Once again, Player I can choose to STOP, with no further payments, or to GO, in which case Player I is fined another £1 and Player II receives another £10 and the game continues. After Player I's second move, whether it is STOP or GO, the game ends. This is a short version; it is really a tripede rather than a centipede, because it has only three legs. Longer versions of this game are possible, of course, and it would look more like a centipede if it had 100 decision nodes, in which case the payoffs to both players would then mount up to large sums if there were a lot of GO choices.

Backward induction leads to the surprising conclusion that Player I should stop the game at the very first move and be content with a zero payoff. Suppose the game has reached the third decision node at which Player I is about to make the last move of the game. Player I can choose to STOP or to GO, but the only rational choice at that point is to STOP and pocket the payoff of £9, rather than to choose GO and receive a lesser payoff of £8. This means that at the immediately preceding decision node, Player II effectively chooses between stopping the game and receiving £10 or continuing it and receiving £9 when Player I, who is assumed to be rational, responds by stopping the game on the following move; Player II will therefore choose STOP at the second decision node. This in turn means that Player I, at the first decision node of the game, effectively chooses between stopping the game immediately and receiving a zero payoff or choosing GO and losing £1 when Player II, assumedly rational, stops the game at the second decision node. Player I will therefore choose STOP immediately. Even if the Centipede has 100 feet and fabulous payoffs towards its head, the backward induction argument leads inexorably to the same conclusion: Player I, if rational, will stop the game at the very first move, in spite of the fact that vastly better payoffs are guaranteed for both players if they both choose GO moves. The outcome that results from Player I choosing STOP at the first decision node is the unique subgame perfect equilibrium point of the Centipede game, and this is extremely puzzling.

Experimental evidence (El Gamal, McKelvey, & Palfrey, 1993; McKelvey & Palfrey, 1992) shows that human decision makers seldom follow the logic of backward induction. In a four-legged Centipede game with a maximum payoff of \$6.40 at its head, only 7 per cent of McKelvey and Palfrey's players stopped the game at the first decision node. In a high-payoff four-legged Centipede game with a maximum payoff of \$25.60 at its head, 15 per cent stopped the game at the first decision node. And in a six-legged Centipede game with a maximum payoff of \$25.60 at its head, just under 1 per cent stopped the game at the first decision node. Even at the last decision node, only between 69 per cent of players (in the high-payoff four-legged Centipede) and 85 per cent of players (in the low-payoff four-legged Centipede) chose STOP. A significant proportion of the players who participated in the experiment---between 6 and 7 per cent---chose GO at every opportunity they got, whereas less than 1 per cent chose STOP at every opportunity.

## 5. Common knowledge and rationality (CKR)

In order to analyse the logic of the backward induction argument, it is necessary to clarify the standard knowledge and rationality assumptions of game theory. These assumptions are collectively called *common knowledge and (or of) rationality* or CKR (see, e.g., Colman, 1997; Colman & Bacharach, 1997; Hollis & Sugden, 1993; Sugden, 1991, 1992). They are as follows:

- CKR1. The specification of the game and the players' preferences among the outcomes, together with everything that can be logically deduced about the game, are common knowledge among the players.
- CKR2. The players are rational in the sense that they always seek to maximize their own expected utilities, and this is common knowledge among the players.

The form of rationality specified in CKR2 is usually interpreted as referring to the axioms of rational choice under uncertainty as formalized within a Bayesian framework of subjective probabilities by Savage (1954). The term *common knowledge*, which was introduced into game theory by David Lewis (1969, pp. 52-68) and later formalized by Aumann (1976), needs further explanation.

A proposition is common knowledge among a set of players if each player knows it to be true, knows that every other player knows it to be true, knows that every other player knows that every other player knows it to be true, and so on. It is important to distinguish this from the situation of *general* knowledge in which all members of the group merely know the proposition to be true. An example that shows the distinction up especially clearly is the muddy children problem, which is a variant of the unfaithful wives or cheating husbands problem (Fagin, Halpern, Moses, & Vardi, 1995, pp. 3-7, 24-30, 248-250, 397-402). A group of children are playing together, and some of them have mud on their foreheads. All the children can see the mud on the foreheads of other children but cannot see mud on their own foreheads. A parent arrives and announces to the group, 'At least one of you has a muddy forehead'. The parent then asks: 'Does any of you know whether you have a muddy forehead? Put your hand up now if you know.' The parent then repeats this question over and over, pausing for a response each time. Assuming that the children are all intelligent, perceptive, and truthful, and that  $m$  of them have muddy foreheads, it is easy to prove that no children will put their hands up the first  $m - 1$  times the question is asked, but that all the muddy children will put their hands up when the question is asked for the  $m$ th time.

**THEOREM 2.** *If a group of  $n$  intelligent, perceptive, and truthful children includes  $m$  who are muddy ( $m \leq n$ ), and all children can see whether any other children are muddy but not whether they themselves are muddy, and if a public announcement is made to all the children that  $m \geq 1$ , and they are then asked repeatedly to respond iff they know that they themselves are muddy, then none of the children will respond to the first  $m - 1$  iterations of the question, but all of the muddy children will respond when the question is asked for the  $m$ th time.*

*Proof.* The proof proceeds by induction on  $m$ . If  $m = 1$ , then the single muddy child, seeing that none of the other children are muddy but having been told that at least one child is muddy, will respond as soon as the parent asks the question for the first time. Assume that the theorem holds for  $m = k$ . The proof will establish that if  $m = k + 1$ , then the  $k + 1$  muddy children will respond when the question was asked for the  $(k + 1)$ st time. With  $k + 1$  muddy children, each muddy child sees  $k$  muddy children. Because of the inductive hypothesis that the theorem holds for  $m = k$ , the muddy children know that if there were just  $k$  muddy children, then those  $k$  children would have responded when the question was asked for the  $k$ th

time. Knowing that they did not, the  $k + 1$  muddy children deduce that they are themselves muddy, and they all respond when the question is asked for the  $(k + 1)$ st time. The theorem follows by induction.  $\square$

The problem is interesting if  $m \geq 2$ . If  $m = 2$ , then when the parent asks the question for the first time, neither of the muddy children will respond, because they have been told that there is at least one muddy child in the group and will both assume that there could be only one. But when the question is repeated a second time, they will both respond, because they will both realize that if there were only one muddy child, then that child would have responded when the question was asked for the first time. For  $m = 3$ , when the parent asks the question for the third time, the three muddy children will deduce that they are muddy from the failure of the two other muddy children to respond when the parent asked the question for the second time. The argument generalizes to  $m = 4, 5$ , and so on indefinitely.

This puzzle exposes the essential nature of common knowledge. If the proposition that ‘at least one member of the group has a muddy forehead’ is denoted by  $P$ , then for any  $m \geq 2$ , all members of the group already know  $P$  before the parent announces it. It may therefore seem that the parent’s announcement that ‘At least one of you has a muddy forehead’ tells them nothing that they do not know already. But although they all know  $P$  already, the public announcement of  $P$  must none the less convey some additional information, because without the announcement none of the children will ever respond, no matter how often the parent asks the question. Without the announcement, if  $m = 1$ , when the parent asks the question for the first time, the muddy child will reason that there may be no muddy children in the group, will not respond, and will not deduce that he or she is muddy no matter how often the question is repeated. If  $m = 2$ , the children all know from the evidence of their own eyes that there is at least one muddy child in the group, even without any public announcement of this fact, but neither will respond when the parent asks the question for the first time, because they will both assume that the muddy child that they can see may be the only one in the group, and neither will respond when the question is repeated for a second time, because each will know that the failure of the other muddy child to respond when the question was asked for the first time could be explained by that child assuming that there are no muddy children in the group. No matter how often the question is repeated, the muddy children will never own up to being muddy. The point is that although the children all know that there is at least one muddy child in the group, without a public announcement they do not know that they all know this. It follows that the parent’s public announcement of  $P$ , a proposition that all members of the group already know to be true, does indeed convey some new information. It changes  $P$  from a proposition that every member of the group merely knows to be true into a proposition that is common knowledge in the group. After  $P$  has been announced publicly, not only does every child know  $P$ , a state of knowledge that we may symbolize by  $KP$ , but in addition to this  $P$  becomes common knowledge in the group, which we may write  $CKP$ :

$$CKP := KP \wedge KKP \wedge KKKP \wedge \dots \quad (1)$$

The backward induction argument rests on the assumptions CK1 and CK2 mentioned earlier. These are the assumptions that the specification of the game, the players’ preference functions, and the fact that the players are all rational are common knowledge in the game, in the sense of (1).

The leading explanation for empirically reported deviations from the backward induction path focuses on reputation effects and requires a partial relaxation of the CKR assumptions.

Kreps, Milgrom, Roberts, and Wilson (1982), often referred to in game theory circles as the Gang of Four, showed that two perfectly rational PDG players who each believe that there is a small probability that the other player is irrational will deviate from the prescriptions of backward induction in an attempt to influence the other player by building their reputations for cooperativeness (see also Kreps & Wilson, 1982a, 1982b). Note that this involves an element of incomplete information rather than incomplete rationality on the part of the players: the players are assumed to be perfectly rational, and it is the information rather than the rationality assumptions of CKR that are relaxed. Experiments designed specifically to test this theory in the finitely iterated PDG (Andreoni & Miller, 1993; Cooper, DeJong, & Forsythe, 1996) and the Centipede game (McKelvey & Palfrey, 1992) have tended to show that this explanation in terms of incomplete information and reputation-building cannot explain all violations of backward induction. The results suggest that a small proportion of people behave cooperatively or altruistically irrespective of any assumptions that they may hold about their co-players' rationality or any attempt to bolster their own reputations.

To summarize: The standard knowledge and rationality assumptions of game theory include the assumption that players choose their moves or strategies rationally in the sense of expected utility theory and that the fact that they do this is common knowledge in the sense that every player knows it, every player knows that every player knows it, and so on. The backward induction paradox, which can be traced to an analysis by Luce and Raiffa (1957) of the finitely iterated Prisoner's Dilemma Game, appears to show that CKR implies absurd or self-defeating behaviour in certain mutually interdependent decision situations involving sequential choices. Empirical evidence suggests that human decision makers do not always follow the backward induction path even when they are capable of understanding the logic of the argument.

## 6. Nonmonotonic reasoning

The backward induction argument apparently proves that Player I will stop the Centipede game on the first move. Bearing in mind the CKR2 rationality assumption that players always seek to maximize their own expected utilities, it follows that the backward induction argument must implicitly have proved that, for Player I, the expected utility from choosing STOP on the first move is greater than the expected utility from choosing GO. Backward induction therefore suggests that Player I must have believed that a choice of GO on the first move would result in Player II responding with STOP. If Player I believed that Player II would respond by also choosing GO, then Player I's expected utility would be higher from choosing GO than from choosing STOP, and so Player I would choose GO.

There is something wrong, because Player I cannot have *any* belief about how Player II would respond to an opening choice of GO. If Player I were to choose GO on move 1, then a situation would exist that Player II would know to be strictly impossible under CKR2, because the logic of backward induction dictates that Player I will choose STOP on move 1 and Player II knows that Player I is rational. It is simply incoherent to ask what a theory predicts in a situation that is inconsistent with one of its assumptions. We cannot expect a theory to predict what would happen if the theory were false. So, as several game theorists have pointed out (Basu, 1990; Bicchieri, 1989; Binmore, 1987; Bonanno, 1991; Pettit & Sugden, 1989; Reny, 1986; Sugden, 1991, 1992), CKR2 implies the truth of an incoherent proposition, namely that Player I's expected utility must be greater from choosing STOP on the first move than from choosing GO. This means that CKR2 must itself be an incoherent assumption. On the one hand CKR2 requires a player to evaluate all available moves in order to maximize (subjective) expected utility, but on the other certain available moves cannot be evaluated or even defined without violating CKR2. Aumann (1995), who provided a rigorous

analysis of backward induction in sequential games, simply sidestepped this problem by explicitly restricting players' rationality at each node of the game tree to what happens from that point on, whether or not that node could have been reached by rational play, so that Aumann's players are rational when contemplating the future but both blind and amnesic to the past.

The CKR information and rationality assumptions can be made coherent by replacing them with assumptions that Monderer and Samet (1989) have called common beliefs and that Sugden (1992) has called entrenched common beliefs. The specification of the game, the players' payoff functions, and the players' rationality then become matters of common belief rather than common knowledge. A proposition is a matter of common belief if each player believes it to be true, believes that the other player(s) believe(s) it to be true, and so on, and if each player continues to believe it as long as this belief can be maintained without inconsistency. Sugden has shown that a replacement of CKR with entrenched common belief in rationality allows a player to evaluate the required expected utilities and eliminates the incoherence. But I shall argue below that this approach leads to something worse than incoherence, namely a straightforward contradiction, because it implies that a rational player will both choose and not choose the GO move in the Centipede game.

I suggest that greater clarity and precision may be achieved by reformulating the CKR2 assumption as a default inference rule within the framework of a system of nonmonotonic reasoning (Geffner, 1992; Ginsberg, 1987; Makinson, 1989; McCarthy, 1980, 1986; Reiter, 1987). Formalizations of nonmonotonic reasoning are designed to capture some of the features of common-sense reasoning. The canonical example consists of the two premises 'Birds fly' and 'Tweety is a bird' together with the conclusion 'Tweety can fly'. In this example, 'Birds fly' is interpreted to mean 'Normally, birds fly', 'Typically, birds fly', or 'If  $x$  is a bird, then assume by default that  $x$  flies'. In other words, the proposition that Tweety can fly is a default assumption that we adopt in the absence of information that might override it. If we introduce the additional item of information that Tweety is a penguin, or an ostrich, or an emu, or that Tweety is coated in heavy crude oil, or that Tweety is dead, then the conclusion no longer follows. It is in this sense that the reasoning is nonmonotonic: in classical first-order logic, if a set of premises  $P$  implies a conclusion  $q$ ,  $P \rightarrow q$ , then the addition of further premises from a larger set  $S \supset P$  cannot affect the truth of the conclusion  $q$ . This is called the extension theorem of semantic entailment in propositional calculus. But in nonmonotonic reasoning the extension theorem does not hold, and additional premises can negate conclusions derived from a smaller set of premises.

In McCarthy's (1980, 1986) system of circumscription, ' $x$  is a bird' may be symbolized by  $B(x)$ , ' $x$  can fly' by  $F(x)$ , and ' $x$  is abnormal' (in the sense that it is a non-flying bird) by  $Ab(x)$ . The default inference rule 'birds fly' then becomes

$$\forall x B(x) \wedge \neg Ab(x) \rightarrow F(x). \quad (2)$$

Reinterpreting (2) in terms of game theory, ' $x$  is a player' may be symbolized by  $P(x)$ , ' $x$  is rational in the sense of CKR2' by  $R(x)$ , and ' $x$  is abnormal' (in the sense of violating CKR2 in the game) by  $Ab(x)$ . The default inference rule 'the players in the game are rational in the sense of CKR2' then becomes

$$\forall x P(x) \wedge \neg Ab(x) \rightarrow R(x), \quad (3)$$

and this inference rule is common knowledge in the game. If Diana is a specific player, symbolized by  $d$ , then

$$(\forall x P(x) \wedge \neg Ab(x) \rightarrow R(x)) \wedge P(d) \rightarrow R(d). \quad (4)$$

The conclusion on the right-hand side of (4) is defeasible, however. If the game is Centipede, for example, and Diana is in the role of Player I and chooses GO at the first decision node, then Player II may conclude that  $\neg Ab(x)$ . This is equivalent to  $Ab(x)$ ; it blocks the default inference rule (3) and prevents the conclusion from being drawn in (4).

It is possible to replace the standard information and rationality assumptions CKR with a revised set of assumptions CKR'. In CKR', CKR2 is replaced with the default inference rule (3), which I shall label *W*. Thus CKR2 is replaced by the following pair of assumptions:

CKR2'.  $W := \forall x P(x) \wedge \neg Ab(x) \rightarrow R(x)$ .

CKR3'.  $KW \wedge KKW \wedge KKKW \wedge \dots$

This is the default inference rule plus the stipulation that it is common knowledge in the game. The players are assumed to be rational as before, but their assumptions about their co-players' rationality are now defeasible, and these assumptions will be abandoned as premises for reasoning about the game if evidence comes to light showing players to be irrational. This brings CKR within the ambit of common-sense reasoning, and it may provide a formal framework for what Monderer and Samet (1989) and Sugden (1992) had in mind when they developed notions of common belief and entrenched common belief in rationality.

This modification of the CKR2 assumption does not prevent the backward induction argument from going through with the finitely iterated Prisoner's Dilemma, Chain-store, or Centipede games. In all cases, the players, who are still assumedly rational, will choose rationally, and so the default inference rule will not be blocked, therefore the conclusion will be the same as before, and the paradox will remain. The crucial difference is that it is no longer incoherent for a player to ask how a co-player would respond to a deviation from the backward induction path. In the Centipede game, for example, if Player I were to choose GO on Move 1, then Player II would simply cease to believe that Player I was rational, perhaps attributing this to Player I's 'trembling hand' (Selten, 1975, Myerson, 1978), and would choose the best reply in the light of this new information.

Now Player I can ask without incoherence: How would Player II respond if Player I were to choose GO at the initial decision node? What is important is Player II's assessment of Player I's likely response at the following (final) decision node if Player II were to choose GO at the second decision node. If Player II thinks that there is a probability of 1/2, say, that Player I would respond by choosing GO once again, then Player II, by choosing GO, would receive a payoff of 19 with probability 1/2 (if Player I responds with GO) and a payoff of 9 with probability 1/2 (if Player I responds with STOP). Thus, in deciding whether to choose STOP or GO at the second decision node, Player II's expected payoffs are as follows:

1. Player II's expected payoff for choosing STOP: 10 units with certainty.
2. Player II's expected payoff for choosing GO:  $19/2 + 9/2 = 14$  units.

The rational response for Player II, under the revised information and rationality assumptions, would be GO, because GO would maximize Player II's expected utility. In fact, if Player II had a subjective probability greater than 1/10 that Player I would respond to a GO choice by also choosing GO, then Player II's rational strategy, according to the revised version of CKR2, would be GO, because the expected payoff of choosing GO given a 1/10 subjective probability of a GO reply would be  $(1/10)(19) + (9/10)(9) = 10$ , which would be equal to the payoff of choosing STOP, and the expected payoff of choosing GO would be higher if the subjective probability of a GO reply were higher.

In the short tripede in Figure 3, of course, Player II may not expect Player I to repeat the GO strategy on the next move, because the next move is the last. But this reasoning applies quite generally, even in the middle of a long Centipede game---a millipede, perhaps---with many moves and fabulous riches dangling from its head. Suppose that Player II has observed Player I's behaviour on many previous moves, and Player I has invariably chosen GO. In that case, Player II may choose GO, confidently expecting a GO reply from Player I. But according to the backward induction argument, on the first move of the game Player I will none the less expect a lower payoff from choosing GO than STOP, which means that Player I attaches a low subjective probability to the prospect of Player II choosing GO on the second move.

This conclusion is surprising and paradoxical. The backward induction argument mandates a STOP choice from a player at any decision node that may be reached. This implies that, in evaluating the expected utility of choosing GO, the player must be assigning a small subjective probability to a GO reply from the co-player, because it is assumed that players always seek to maximize their own expected utilities. Given the payoffs shown in Figure 3, for example, the backward induction argument implies that Player II, evaluating the subjective utility of a GO choice at the second decision node, must assign a subjective probability of less than 1/10 to a GO reply from Player I. Furthermore, this low subjective probability of a GO reply appears to be insensitive to influence from relevant empirical evidence. It is not difficult to think of circumstances in which the subjective probability of a GO reply should increase. For example, in the middle of a long Centipede game, after Player I had chosen GO at all previous decision nodes, it would seem reasonable for Player II to choose GO and to assign a higher subjective probability to a further GO choice from Player I. Many would argue that there must come *some* point at which inductive reasoning---by which I mean empirical induction of the type conventionally used for deriving posterior probabilities in Bayesian analysis, not mathematical induction of the type used in backward induction---forces the subjective probability of a GO reply above the threshold that gives a Player II a higher expected utility for a GO than a STOP choice. This is essentially the argument put forward by Sugden (1992). But the backward induction argument implies that a player will invariably choose STOP at any decision node, irrespective of empirical evidence that a GO reply is probable. Thus we have two rational arguments derived from the same knowledge and rationality assumptions, one for choosing GO and one for choosing STOP. Perhaps this is merely an elaborate new refutation of empirical induction (cf. Popper, 1959), which lurks behind the usual Bayesian interpretation of expected utility embodied in CKR2. But it is the backward induction argument, based on mathematical induction, rather than the expected utility argument, based on empirical induction, that appears suspect, because the empirical evidence of past moves must surely be of some probative relevance.

## 7. Conclusions

The backward induction paradox has attracted attention from decision theorists ever since Selten (1978) highlighted it in relation to the Chain-store game and Rosenthal (1981) provided a particularly stark example of it in the Centipede game. I have not resolved the paradox; in fact, like an over-zealous physician, I have induced a more serious pathology than the one I set out to cure. The backward induction argument suggests that in the Centipede game Player I will choose STOP at the first decision node, that in the Chain-store game the Chain-store will always respond cooperatively to challenges, and that in the finitely iterated PDG both players will defect on every round. Under the revised CKR' assumptions, this implies that the expected utility of these strategies is less than the expected utility of their seemingly more plausible alternatives. This conclusion arises from the reasonable default

assumption that the co-players are rational.

But the backward induction argument rests on a hidden assumption about the likely responses to behaviour off the backward induction path, and this assumption cannot necessarily be maintained irrespective of the co-player's past behaviour. The standard CKR assumptions of game theory allow no prediction about how a co-player would respond to a deviation from the backward induction path, because they imply that this situation simply cannot arise. With the revised CKR' assumptions, in which CKR2' is a default inference rule in a system of nonmonotonic reasoning, subject to common knowledge, a player can at least contemplate the co-player deviating from the backward induction path. For example, in the Centipede game, the revised knowledge and rationality assumptions imply that Player II may respond to GO with GO. Player II will choose GO if the expected utility of doing so exceeds the expected utility of choosing STOP. Having seen Player I choosing GO on the first move, Player II will be forced to infer that Player I is not rational in the sense of CKR2 and will thus abandon the default assumption (4) above. Faced with an irrational co-player, Player II may therefore judge that Player I, having already chosen a GO move off the backward induction path, may choose further GO moves in future if given the opportunity to do so, and Player II may therefore be happy to provide Player I with an opportunity to do so, given that this would benefit both players. All this being common knowledge in the game, Player I may therefore assess the expected utility of choosing GO on the first move to be greater than the expected utility of choosing STOP. Yet according to the backward induction argument Player I will none the less choose STOP on the first move, because at that stage there is nothing to block the default assumption (4) that Player II is rational in the sense of CKR2'. One line of reasoning (backward induction) prescribes a STOP choice on the first move, and another (the principle of expected utility maximization) prescribes GO. This is a straightforward contradiction and a genuine paradox, and it shows that something is seriously wrong with the knowledge and rationality assumptions of game theory.

The Prisoner's Dilemma, Chain-store, and the Centipede games are well defined games that real people can and do play. It is reasonable to ask what are the rational ways of playing them. Backward induction seems to imply unwavering defection in the PDG, cooperative responses to challenges in the Chain-store game, and STOP choices in the Centipede game, and many writers have commented that these conclusions are highly counter-intuitive. A more careful analysis has shown that the conclusions are not merely counter-intuitive but may violate the principle of expected utility maximization built into the assumption of the players' rationality in CKR2. The standard information and rationality assumptions are incoherent. If we tidy them up by replacing CKR2 with a default inference rule regarding the players' rationality, which is assumed to be common knowledge in the game, then we can derive conclusions that are mutually contradictory. They suggest that in the Centipede game there are circumstances in which Player I both will and will not choose GO at the first decision node. This suggests that the knowledge and rationality assumptions, even as modified via nonmonotonic reasoning to allow reasoning about departures from the backward induction path, must be false or incoherent. In particular, the rationality assumption CKR2 or its nonmonotonic equivalent CKR2' may be suspect, but it is not clear where the problem lies. The expected utility argument relies on empirical induction, which has no adequate rational justification, but it is the conclusions of the backward induction argument that are hard to swallow.

## References

- Anderson, J. R. (1990). *The Adaptive Character of Thought*. Hillsdale, NJ: Erlbaum.  
 Anderson, J. R. (1991). Is human cognition adaptive? *Behavioral and Brain Sciences*, 14,

471-485.

- Andreoni, J. & Miller, J. H. (1993). Rational cooperation in the finitely repeated Prisoner's Dilemma: Experimental evidence. *The Economic Journal*, 103, 570-585.
- Aumann, R. J. (1976). Agreeing to disagree. *The Annals of Statistics*, 4, 1236-1239.
- Aumann, R. J. (1995). Backward induction and common knowledge of rationality. *Games and Economic Behavior*, 8, 6-19.
- Basu, K. (1990). On the non-existence of a rationality definition for extensive games. *International Journal of Game Theory*, 19, 33-44.
- Bicchieri, C. (1989). Self-refuting theories of strategic interaction: A paradox of common knowledge. *Erkenntnis*, 30, 69-85.
- Binmore, K. (1987). Modeling rational players: Part 1. *Economics and Philosophy*, 3, 179-214.
- Bonanno, G. (1991). The logic of rational play in games of perfect information. *Economics and Philosophy*, 7, 37-65.
- Colman, A. M. (1995). *Game Theory and its Applications in the Social and Biological Sciences* (2nd ed.). Oxford: Butterworth-Heinemann.
- Colman, A. M. (1997). Salience and focusing in pure coordination games. *Journal of Economic Methodology*, 4, 61-81.
- Colman, A. M. & Bacharach, M. (1997). Payoff dominance and the Stackelberg heuristic. *Theory and Decision*, 43, 1-19.
- Cooper, R., DeJong, D. V., & Forsythe, R. (1996). Cooperation without reputation: Experimental evidence from Prisoner's Dilemma games. *Games and Economic Behavior*, 12, 187-218.
- Dawkins, R. (1989). *The Selfish Gene* (2nd ed.). Oxford: Oxford University Press.
- El-Gamal, M. A., McKelvey, R. D., & Palfrey, T. R. (1993). A Bayesian sequential experimental study of learning in games. *Journal of the American Statistical Association*, 88(422), 428-435.
- Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. Y. (1995). *Reasoning about Knowledge*. Cambridge, MA: MIT Press.
- Geffner, H. (1992). *Default Reasoning: Causal and Conditional Theories*. Cambridge, MA: MIT Press.
- Ginsberg, M. L. (Ed.). (1987). *Readings in Nonmonotonic Reasoning*. Los Altos, CA: Morgan Kaufmann.
- Hollis, M. & Sugden, R. (1993). Rationality in action. *Mind*, 102, 1-35.
- Krebs, J. R. & Davies, N. B. (1991). *Behavioural Ecology: An Evolutionary Approach* (3rd ed.). Oxford: Basil Blackwell.
- Kreps, D. M. (1990). *A Course in Microeconomic Theory*. Princeton, NJ: Princeton University Press.
- Kreps, D. M., Milgrom, P., Roberts, J., & Wilson, R. (1982). Rational cooperation in the finitely repeated Prisoner's Dilemma. *Journal of Economic Theory*, 27, 245-252.
- Kreps, D. M. & Wilson, R. (1982a). Sequential equilibrium. *Econometrica*, 50, 863-894.
- Kreps, D. M. & Wilson, R. (1982b). Reputation and imperfect information. *Journal of Economic Theory*, 27, 253-279.
- Lazarus, J. (1982). Competition and conflict in animals. In A. M. Colman (Ed.), *Cooperation and Competition in Humans and Animals* (pp. 26-56). Wokingham: Van Nostrand Reinhold.
- Lazarus, J. (1987). The concepts of sociobiology. In H. Beloff & A. M. Colman (Eds), *Psychology Survey 6* (pp. 192-217). Leicester: The British Psychological Society.

- Lazarus, J. (1994). Behavioural ecology and evolution. In A. M. Colman (Ed.), *Companion Encyclopedia of Psychology* (Vol. 1, pp. 69-87). London: Routledge.
- Lewis, D. K. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.
- Luce, R. D. & Raiffa, H. (1957). *Games and Decisions: Introduction and Critical Survey*. New York: Wiley.
- Makinson, D. (1989). General theory of cumulative inference. In M. Reifrank, J. de Kleer, M. L. Ginsberg, & E. Sandewall (Eds), *Lecture Notes in Artificial Intelligence: Subseries of Lecture Notes in Computer Science* (pp. 1-18). Berlin: Springer-Verlag.
- McCarthy, J. (1980). Circumscription: A form of non-monotonic reasoning. *Artificial Intelligence*, 13, 27-39.
- McCarthy, J. (1986). Applications of circumspection to formalizing common sense knowledge. *Journal of Artificial Intelligence*, 28, 89-116.
- McKelvey, R. D. & Palfrey, T. R. (1992). An experimental study of the Centipede game. *Econometrica*, 60, 803-836.
- Monderer, D. & Samet, D. (1989). Approximating common knowledge with common beliefs. *Games and Economic Behavior*, 1, 170-190.
- Myerson, R. (1978). Refinements of the Nash equilibrium concept. *International Journal of Game Theory*, 8, 73-80.
- Pettit, P. & Sugden, R. (1989). The backward induction paradox. *The Journal of Philosophy*, 86, 169-182.
- Popper, K. R. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.
- Rapoport, A. & Chammah, A. M. (1965). *Prisoner's Dilemma: A Study in Conflict and Cooperation*. Ann Arbor, MI: University of Michigan Press.
- Reiter, R. (1987). Nonmonotonic reasoning. *Annual Review of Computer Science*, 2, 147-186.
- Reny, P (1986). *Rationality, Common Knowledge and the Theory of Games*. Unpublished doctoral dissertation, Princeton University.
- Reny, P. (1992). Rationality in extensive form games. *Journal of Economic Perspectives*, 6, 103-118.
- Rosenthal, R. W. (1981). Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory*, 25, 92-100.
- Savage, L. (1954). *The Foundations of Statistics*. New York: Wiley.
- Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4, 25-55.
- Selten, R. (1978). The chain store paradox. *Theory and Decision*, 9, 127-159.
- Selten, R. & Stoecker, R. (1986). End behavior in sequences of finite Prisoner's Dilemma supergames. *Journal of Economic Behaviour and Organization*, 7, 47-70.
- Sugden, R. (1991). Rational choice: A survey of contributions from economics and philosophy. *The Economic Journal*, 101, 751-785.
- Sugden, R. (1992). Inductive reasoning in games. In R. Selten (Ed.), *Rational Interaction: Essays in Honor of John C. Harsanyi* (pp. 201-221). Berlin: Springer-Verlag.
- Zermelo, E. (1912). Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels. *Proceedings of the Fifth International Congress of Mathematicians*, Cambridge, 2, 501-10.

### Acknowledgements

The research reported in this article was supported by Grant No. L122251002 from the

Economic and Social Research Council as part of the research programme on Economic Beliefs and Behaviour. I am grateful for comments on an earlier version given by participants at the conference on Rational Models of Cognition at Warwick University, 22-24 July 1996, and to Kenneth Hughes for a helpful suggestion.