

**THE REDUCTION OF DECISION COMPLEXITY:  
NORMATIVE POLICIES AND THE ROLE OF INFORMATION\***

by Luigi Mittone

---

\* This paper has benefited from the suggestions and comments made by the participants at a seminar held in the Economics Department of the University of Trento, to whom I owe my sincere thanks. I am especially indebted to Massimo Egidi and to Paolo Patelli, who were of great help in development of the dynamic model.

## 1. Introduction: hockey players and externalities

The principal aim of this work<sup>1</sup> is to interpret the function of the state as a *collective mechanism for reducing the complexity* of the environment within which economic agents must decide their behaviour. The source of complexity is assumed to be a phenomenon of diffuse externalities. The agents' decision problem takes the form of a binary choice - whether or not to buy a given a good, whether or not to drive a car, and so on - the extended consequences of which on their welfare can't be cognitively controlled by the agents themselves because they depend on a complex network of interrelated actions (mainly taken by other agents) external to the binary choice context strictly considered by them. The term "cognitively uncontrollable" is equivalent to Simon's (1955) concept of *bounded rationality*, and it is used here to denote the inability of decision-takers to use the *correct* number and to choose the *right* pieces of information necessary to accomplish the rational process that should induce them to choose the action(s) that maximises their utility (welfare)<sup>2</sup>. Under this definition, I shall use the game theoretical term "common knowledge" as synonymous with "extended (or perfect) rationality".

There exist a considerable number of possible economic spheres in which externalities occur which are not reducible to a simple model of dichotomous interaction (agent which causes the externality/agent which subject to the externality) but which are nonetheless produced as a consequence of a binary choice: these include, for example, the decision whether or not to move to another town, whether or not to be vaccinated, whether to go to work by bicycle, by public transport or by car. All these are examples of dichotomous decisions which exert an influence, sometimes even a crucial one, on the agent's interest and on the welfare of other individuals. In the case of the decision to change one's place of residence, for instance, this may increase, albeit to a minimum extent, the level of urban pollution of the new area of residence, while pollution should decrease in the old area. Likewise, places are made available in public structures (hospital, schools, etc.) but an equivalent number are occupied in the town to which the family has moved.

An interesting example of a quite peculiar kind of externality produced as a consequence of a dichotomous choice is suggested by Thomas Schelling (1973), who reports a comment by *Newsweek* (1969) concerning the severe injury caused to the hockey player Teddy Green by his refusal to wear a helmet:

Players will not adopt helmets by individual choice for several reasons. Chicago star Bobby Hull cites the simplest factor: 'Vanity'. But many players honestly believe that helmets will cut their efficiency and put them at a disadvantage, and others fear the ridicule of opponents. The use of helmets will spread only through fear caused by injuries like Green's - or through a rule making them mandatory . . . One player summed up the feelings of many: 'It is foolish not to wear an helmet. But I don't - because the other guys

---

<sup>1</sup> The literature on this subject is so vast that I have insufficient space to review it here. The reader is referred to the bibliography and to Egidi (1991), who discusses most of the works referred to here.

<sup>2</sup> On the definition of bounded rationality we shall come back more extensively in the next paragraphs.

don't. I know that's silly, but most of the players feel the same way. If the league made us do it, though, we'd all wear them and nobody would mind.' (p. 38).

The example reported by Schelling highlights the fact that the decision problem confronting hockey players was conceptually similar to a dynamic multi-agents Prisoner's Dilemma game (henceforth P-D). Each hockey player can choose whether or not to wear a helmet, and his payoff is influenced by the actions chosen by the others. A simplified version of the hockey players' dilemma can be represented by the classic static P-D with only two players (fig.1).

The payoff structure of the game (see fig. 1) is based on the assumption that the value attributed by hockey players possible injury due to their refusal to wear an helmet is lower than the psychological cost (shame at being ridiculed by the other players) that they would suffer if they decided to wear the helmet and the others did not. Note that the payoff structure shown in fig. 1 attributes a positive payoff if a player does not wear a helmet but his opponent does. This characteristic of payoffs is due to the special assumption that there is some sort of psychological symmetry between shame at being considered ridiculous by others and the psychological reinforcement felt in the reverse situation (feeling of strength and courage when the others wear helmets and one does not).

**Fig. 1 The hockey players' dilemma**

		Player $\alpha$	
		Don't Wear	Wear
Player $\beta$	Don't Wear	0, 0	2, -1
	Wear	-1, 2	1, 1

Obviously, the payoffs structure chosen is arbitrary because it incorporates a set of assumptions about the psychological value attributed by the players to the choice wear not to wear a helmet. On the other hand the assumptions made seem very realistic, at least on the bases of the declaration made by the hockey player interviewed by *Newsweek*.

Admitting then that the hockey players' dilemma can be modelled by the game of fig. 1 the "don't wear don't wear" solution is the only existing Nash equilibrium. Therefore the refusal to wear a helmet seems to be the only possible rational choice. Do we reach the same conclusion if we move to the real context that the hockey players must cope with, i.e. if we consider a multi-agent repeated game? Consideration of this more complex environment yields some interesting insights into the role played by an external supervisor

(the hockey league in the example considered here, or the state in society as a whole), especially if we introduce the assumption that agents are unable cognitively to govern the entire landscape of actions and consequences, i.e. in the presence of bounded rationality.

## **2. Some reasons for the failure of a spontaneous emerging of Pareto-efficient strategies**

The example of the hockey players can provide a precise model of the phenomenon with which I began, that is, the existence of external effects due to the inability to govern agent interactions rationally. The bounded rationality consequence in the example considered is the absence of a spontaneous convergence towards a Pareto-efficient solution of the game (a solution that can be assumed to be the state of the world without uncontrolled effects produced by agent interactions) ; effects which on the contrary should automatically arise if the players can cognitively govern the entire space of payoffs and strategies.

The possibility of the spontaneous emergence of a Pareto optimal outcome in a two-player repeated P-D context has been empirically demonstrated by Robert Axelrod (1984)<sup>3</sup> who has shown the existence of a dominant strategy that enables the agents to spontaneously converge towards the Pareto optimum. Axelrod's experiment consisted in a tournament where many strategies, developed by game theorists and computer scientists, had to compete each against all the others. One of the simplest strategies, known as "tit for tat", which consisted in making the first move co-operatively and then copying the moves of the adversary, was able to defeat much more elaborate and sophisticated strategies, while also producing in most cases a Pareto convergence.

Axelrod's experiment is interesting in two respects. First, it demonstrates that the spontaneous emergence of a Pareto-efficient strategy in a world of egoists - this being the subject of Axelrod's study - depends closely on the operation of a cognitive constraint: the agents' inability to govern the information structure of the decisional problem. Second, it shows that this strategy displays all the basic characteristics of a norm subjectively applied and reinforced. These two conclusions are closely connected, and we must begin with the latter in order to explain the former. The tit-for-tat strategy is based on imitative behaviour, delayed by one move, that is to say, it is based on the application, initially by only one player, and subsequently by both if the second agent recognises the strategy, of a system of rewards (co-operative choices) and punishments (non co-operative choices). In this sense the tit-for-tat strategy resembles a norm whose extremely simple system of enforcement is implemented by the players themselves without the intervention of an external judge. The simplicity of the enforcement system incorporated in the spontaneous

---

<sup>3</sup> Long before Axelrod, however, it had been demonstrated that dynamic games possess a large set of equilibria, including co-operation, see e.g. Friedman (1971).

Pareto-efficient strategy, identified by Axelrod, reduces the negative effects produced by the agents' cognitive limitations.

Given that in Axelrod's experiment each strategy was compared with all other strategies - that is, with the strategic plans of several agents of differing skills - an essential condition for the efficient functioning of the system of rewards and punishments that ensured the success of the tit-for-tat strategy was its universality: that is, it had to be understood by all the other strategies, i.e., by all the players involved. The empirically observed success of a simple enforcement system apparently showed that it is conceptually impossible, or cognitively inefficient, to design a sophisticated mechanism of behavioural response when the number of interacting agents exceeds even a few dozen. In fact, in order to design a system of rewards and punishments able to respond to all possible strategies one must know them all, that is, one must be able to predict them. And plotting the map of all possible strategies and the interactions among them requires enormous computational resources. The success of the tit-for-tat strategy can therefore be interpreted as proof of the existence of some form of cognitive constraint closely connected with the degree of complexity of the environment (Heiner, 1983). It is important to stress that most of the players invited by Axelrod were experts, and therefore that the constraint of the system of rewards and punishments depended not on the inability of the agents to understand a more sophisticated system, but on the impossibility of designing a response model sufficiently complex to take account of all the possible strategies followed by the various players. This consideration is particularly true of the probabilistic-based strategies defeated by tit-for-tat (see chapter 2 Axelrod 1984).

Given the *naiveness* of tit-for-tat, why do hockey players not apply it? On the basis of the empirical evidence offered by Axelrod, the most plausible reasons are two: the first is that in the hockey players example the punishment mechanism is probably too weak to work effectively, the second one is that the number of simultaneously interacting players is greater than two (unlike in the Axelrod's experiment) and most of all they change with every match. These two explanations are obviously closely related, but the most interesting for us is the second, which once again related to bounded rationality.

Correct application of the tit-for-tat strategy to the hockey players dilemma, requires us to imagine a more complex decisional landscape, one in which each player builds his strategy by monitoring the actions of all the other players, and in which he must wait for the return match to apply the punishment scheme. This situation clearly weakens the punishment system itself because the action-reaction mechanism is delayed and therefore requires the players to have enough *memory* to remember what it has happened in previous matches. Indeed, they must be able to distinguish among the players of the teams that they have encountered over at least a year. Furthermore, it entails development of a multi-interactive punishment device (each player must be able to punish more than one antagonist simultaneously) which cannot be mechanically reduced to the original tit-for-tat structure. Both these complications of the decisional problem require the agents to manage

more cognitive resources (both memory and the ability to control the whole network interactions).

To model the hockey players dilemma we therefore need a different landscape which somehow takes account of the main ingredients of the problem considered here, i.e. the dynamic of the effects produced by the interactions of many cognitively constrained agents.

### **3. A game theoretical treatment of the hockey players' dilemma**

Before presenting in the next paragraph the game used here to model the phenomenon of unexpected effects external to the decisional frame and deriving from a cognitive constraint, it is advisable to mention some differences between the traditional notion of *incomplete information* and the concept of bounded rationality used here. It is well known in game theory that an incomplete information game (also called a *Bayesian game*) involves a situation where the players' payoff functions are not common knowledge. More precisely, dynamic games of incomplete information - that is, the family of games that move closely approximate the problem analysed here - are designed to model a world in which each agent possesses some form of private information and tries to learn the items of information possessed by her/his opponent(s) by observing their actions during repetitions of the game itself. The aim of the theory is to describe under which conditions it is possible to reach a so called *perfect Bayesian equilibrium*. Note (Gibbons, 1992) that perfect Bayesian equilibrium is equivalent to Bayesian Nash equilibrium in static games of incomplete information ; it is equivalent to subgame-perfect Nash equilibrium in dynamic games of complete and perfect information ; and it is equivalent to Nash equilibrium in static games of complete information. This means that overall outcome of the game theory approach to incomplete information in repeated games is the re-establishment of a decisional environment that is conceptually coherent (equivalent) with the original (and ideal) conditions of common knowledge that characterise the simplest family of games, i.e. the one comprising the static two-player P-D game. From an heuristic perspective, a hypothetical agent who behaves in accordance with the game theory's paradigm should pass through a sequential rational process which seeks to transform uncertainty about the possible states of the world into a problem of choice among homogeneous outcomes, each of which is weighted by a probability index, and all of which can be compared and selected by applying the traditional optimal choice rule.

When the game is sufficiently complex (many players and many repetitions), the conditions for a perfect Bayesian equilibrium require the agents should be able to build at each node of the game a complex landscape consisting of her/his beliefs (each of which is a point in the space of probability distributions over the states of the game) and by the

other players' strategies. Furthermore the building of beliefs requires solution of a quite sophisticated formula (the Bayes rule) of conditional probability.

This cognitive process is characterised by three features. The first is that it incorporates the incomplete information problem into a deterministic frame, because there is only one possible optimal choice for each node of the game, obviously given the informational constraints. The second is that to perform this task the agents require a potentially enormous amount of mental resources, and they must all be capable of performing the same calculus, by applying the same rules. The third is that this deterministic outcome is a Nash equilibrium, and under some further conditions it may also converge towards a Pareto optimum.

Can the hockey players' dilemma be realistically analysed and interpreted within a perfect Bayesian equilibrium context? First of all note that the hockey players example does not necessarily require any assumption of the absence of information about the opponents payoffs, which can, on the contrary, be quite realistically assumed as common knowledge. Furthermore, the concept of bounded rationality adopted here is not correctly interpreted by the traditional game-theory definition of incomplete information because it regards a computational limitation on the human mind that cannot simply be overcome by substituting unknown events with probabilities.

To answer the original question we must therefore take a step back in the degree of complexity of the problem described here, we must remove the assumption of incomplete information and resort to the so called Folk Theorem for infinitely repeated games. This operation is useful even if we admit that the hockey players problem is an incomplete information problem, since achievement of a perfect Bayesian equilibrium requires that the conditions for all the previous equilibria had been satisfied. Therefore if for some reason it is impossible to find an equilibrium in a full information world, this is even more the case in an incomplete information context. It is also worth stressing that the decision to examine the infinitely repeated games version of the Folk Theorem - instead of the finitely repeated one - is due to the fact that no hockey player knows with absolute certainty when he will play his last match.

The Folk Theorem has been known to game theorists since the 1950s and every textbook includes it. I recall here Fudenberg and Tirole's (1995) brief definition:

The "folk theorems" for repeated games assert that if the players are sufficiently patient then any feasible, individually rational payoffs can be enforced by an equilibrium. Thus, in the limit of extreme patience, repeated play allows virtually any payoffs to be an equilibrium outcome. (Fudenberg, Tirole, 1995, p. 150)

The Folk Theorem<sup>4</sup> assumes that players use a *trigger strategy* to reply to the moves made by their opponent(s). It thus bears some similarity to the tit-for-tat strategy discussed in the previous section. The main difference with respect to the tit-for-tat strategy is that

---

<sup>4</sup> For a formal representation of the Folk Theorem see Fudenberg and Tirole (1995) p. 146 ff.

the players can also use a sort of delayed punishment scheme which admits the possibility that the other player has not immediately recognised the logic of the trigger strategy applied and needs more than one observed round of the game to converge to some form of collusion. This willingness of players is usually defined as “patience” and it is measured by a discount factor which enters in the players’ utility functions.

It follows that when we leave the empirical terrain for theoretical analysis, we may once more expect the hockey players’ dilemma to converge on some collusive equilibrium. However the reason now why convergence to Pareto efficiency does not spontaneously emerge cannot be mechanically derived from the arguments in the previous section. The first of them, i.e. that in a hockey match there is a simultaneous interaction among more than two players, is immediately eliminated by the fact that the Folk Theorem is usually applied to a n-players game. The second reason - the time delay in the action-reaction mechanism, which means that with each match the agents encounter new opponents (at least until the first return match) - requires discussion in the light of the literature on repeated games with varying opponents.

Following the standard presentation (and using the same formal notations) illustrated in the handbook by Fudenberg and Tirole (1995, page 146-171) we assume that each stage of the game is a finite I-player simultaneous-move game with finite action spaces  $A_i$  and stage-game<sup>5</sup> payoffs functions ( $g_i : A \rightarrow \mathcal{R}$ ), where ( $A = \times_{i \in S} A_i$ ). Let  $\pi_i$  be the space of probability distribution over  $A_i$ . Using a classic approach to the varying opponents repeated games (e.g. Dybvig and Spatt, 1980; Klein and Leffler, 1981; Shapiro, 1982; Kreps, 1986), I assume that there are two kinds of player: the long-run player and the short-run player. The short-run player plays only once and represents the “varying” opponent, the one that long-run players meet in the various matches. The players are labelled so that long-run players are ( $i = 1, \dots, \ell$ ) and ( $j = \ell + 1, \dots, I$ ) represent sequences of short-run players. The long-run players choose their moves by following the ordinary repeated games rule, that is, they maximise the following normalised sum:

$$u_i = E_\sigma(1 - \delta) \sum_{t=0}^{\infty} \delta^t g_i[\sigma^t(h^t)] \quad [3.1]$$

where :

$u_i$  = payoff function of player i;

$\delta$  = discount factor;

$E_\sigma$  = expectation with respect to the distribution over infinite histories generated by strategy profile  $\sigma_i$  ( $\sigma$  denotes the mixed strategies of the overall game);

$(1 - \delta)$  = normalisation factor necessary to measure the stage-game and repeated-game payoffs in the same units;

---

<sup>5</sup> A stage-game is the basic structure of the game i.e. the game that will be repeated.

Given that at least one agent wishes to converge to a Pareto optimum (as in the hockey players example), in a world where there are only long-run players, the Folk Theorem demonstrates that with sufficiently great  $\delta$  there is a Nash equilibrium  $G(\delta)$  that coincides with the Pareto optimum<sup>6</sup>. Can then we conclude that the reason why the hockey players do not spontaneously converge to the Pareto optimum is because they are “impatient”? To verify the matter, we must introduce short-run players as well.

Having defined the long-run players, let us now assume that the short-run players simply maximise each period’s payoff. We therefore have  $I$  players in the stage game, and the players  $(\ell + 1)$  through  $I$  will change at each period in the repeated game. Define:

$$B: \mathbf{A}_1 \times \cdots \times \mathbf{A}_\ell \rightarrow \mathbf{A}_{\ell+1} \times \cdots \times \mathbf{A}_I$$

as the relationship that ties any action profile  $(\alpha_1, \dots, \alpha_\ell)$  for the long-run players to the corresponding Nash-equilibrium actions for the short-run players. This means that for each action  $\alpha$  belonging to the graph (B) and for  $i \geq \ell + 1$ ,  $\alpha_i$  is the best response to  $\alpha_{-i}$ . For each long-run player  $i$  the minmax value  $v_i$  is:

$$\min_{\alpha \in \text{graph}(B)} \max_{\alpha \in \mathbf{A}_i} g_i(\alpha_i, \alpha_{-i})$$

Let  $U = \{v = (v_1, \dots, v_\ell) \in \mathfrak{R}^\ell \mid \exists \alpha \in \text{graph}(B) \text{ with } g_i(\alpha) = v_i \text{ for } (i = 1, \dots, \ell)\}$   
and set  $V = \text{convex hull}(U)$

This is the set of feasible payoffs given the structure of the game described. The main difference (Fudenberg, Kreps and Maskin, 1990) between the results obtained by applying the Folk Theorem respectively to the only long-run players game and to the varying opponents game is that the set of subgame-perfect equilibria in the presence of different typologies of players (long-run together with short-run) can be smaller, unless we introduce the assumption that each player’s choice of a mixed action in the stage game is publicly observable. This means that a spontaneous collusive result is still possible but it requires a public information assumption or an even greater  $\delta$  than that necessary for the long-run players game alone.

Reducing the hockey players’ dilemma to a problem of impatience alone is unsatisfactory, not only because this hypothesis implies that the agents suffer from a sort of psychological myopia but also because this myopia is attributed to very rational and sophisticated decision takers. As we have just seen, long-run players are assumed to be able to build expectations with respect to the distribution over infinite histories generated by each possible strategy profile, and to do so they must know the correspondence that

---

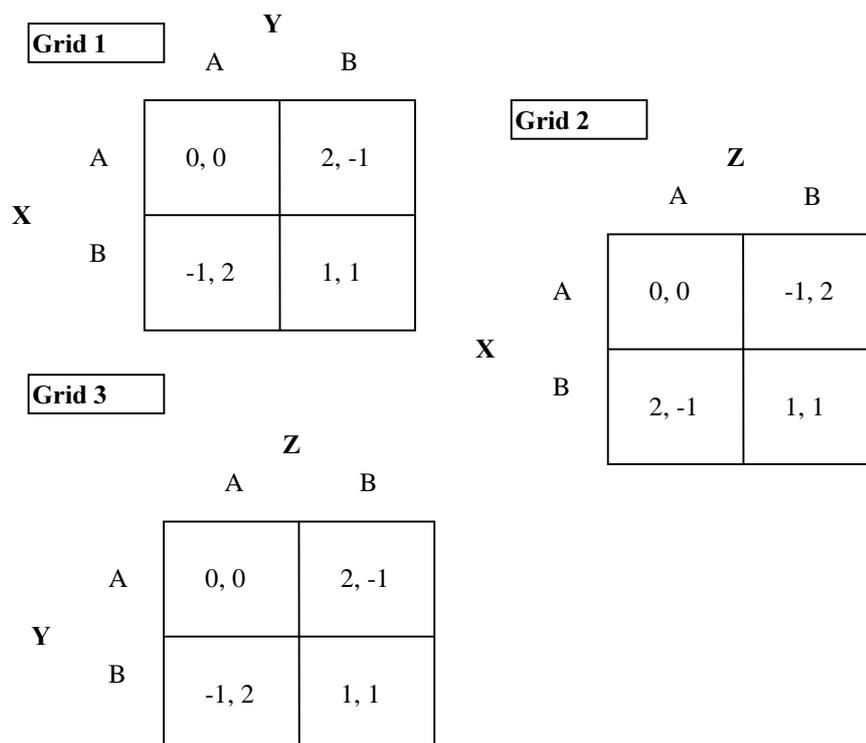
<sup>6</sup>The individual payoff that coincides with the Pareto optimum must obviously be comprised in the set of feasible payoffs that Pareto dominate the minmax payoffs of all players.

maps all their action profiles to the corresponding Nash-equilibrium actions for the short-run players. The existence of a cognitive constraint due to the incapacity of players to conceptually govern the complexity of the problem produces *wrong* expectations about the results generated by the different strategy profiles  $\sigma_i$ . In the hockey players example this constraint creates a sort of artificial compression in the number of strategies that collapses to the simplest one, i.e. to mechanically repetition of the action that corresponds to the Nash equilibrium for the static game.

#### 4. A three players static game with bounded rationality

One way to model the hockey players dilemma is to design a game with more than two players and which models a situation of bounded rationality.

**Fig. 3 Three players and three synchronous grids**



An attempt to model this context is reported in fig. 3, which shows three synchronous grids in which three agents X, Y and Z confront each other. Grids 1 and 3 simply replicate the P-D game structure while the payoffs in the second and in the third cell of grid 2 have been changed so that the Nash equilibrium strategy coincides with the Pareto optimum. This simple modification allows simulation of a situation in which players give an evaluation of the result of the game which depends on who their opponent is. In the

hockey players example this situation may arise when the players play a training match against their team mates or against a youth team and the friendly atmosphere prevents them from feeling be ridiculous when wearing an helmet<sup>7</sup>. This change reproduces an interesting cognitive complication, and combined with the assumption of synchronism, as we shall see, it allows simulation of a bounded rationality effect.

The hypothesis of synchronism of the three grids entails that if an agent decides to adopt strategy (A) on grid 1, where s/he is playing against her/his first opponent, s/he must also adopt strategy (A) on the grid where s/he confronts her/his second one (for example grid 2 for **X** against **Z**). Introducing the assumption of synchronism enables us to simulate the diffusion of extra-market effects produced by individual decisions in all the exchanges which occur in the system. Maintaining for the moment the usual common knowledge assumption, the choices of **X** and **Y** on grid 1 must also take account of their consequences in grids 2 and 3, where both agents must play against **Z**.

The payoff structure of the game is shown better in fig. 4.

**Fig. 4 Game extended form**

		$Z_A$	$Z_B$
$X_A$	$Y_A$	0 0 0	-1 2 1
	$Y_B$	2 -2 2	1 0 3
$X_B$	$Y_A$	1 2 -1	0 4 0
	$Y_B$	3 0 1	2 2 2

Fig. 4 shows that the decision problem of the three players is similar, even if not equivalent, to that in the two-agent P-D game. For players **X** and **Z** the min-max move is (B), while for player **Y** the min-max move is (A). By combining these strategies the Nash equilibrium is in cell (3,2) of fig. 4. Unlike the traditional P-D game, in the fig. 4 game there is no Pareto-optimum strategy because the ( $X_B, Y_B, Z_B$ ) choice represents a Pareto-improvement for **X** and **Z** but not for **Y**, who therefore has no incentive to modify her/his move.

<sup>7</sup>The game represented in grid 2 implicitly assumes a complete inversion in the psychological attitude towards wearing a helmet, which is always seen as preferable even when the opponent is wearing nothing.

This result does not enable us to clarify the main point at issue here: the role played by the cognitive uncontrollability of the outcomes of the game consequent on introducing a highly complex decisional landscape. Accordingly, in order to simulate the operation of a cognitive constraint we introduce two hypotheses:

H<sub>1</sub>) the players are unaware of the existence of more than one opponent (that is, they do not know how many grids they will have to play on);

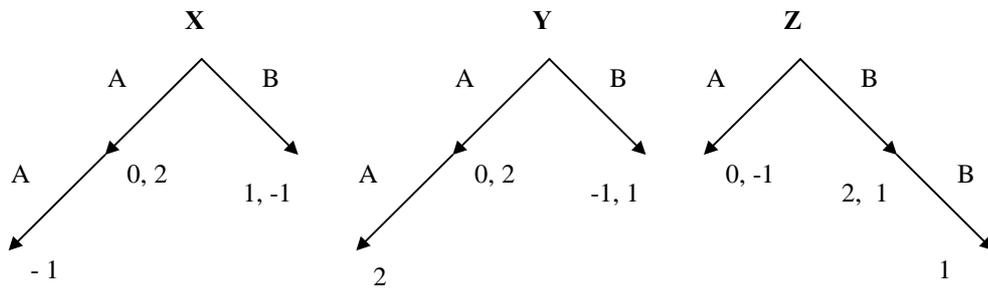
H<sub>2</sub>) the agents are unaware of the possibility of the existence of more than one opponent, *i. e.* they are unaware that more than one state of the world can exist.

The Bayesian approach is not applicable here because the agents behave *as if* they would do in conditions of perfect information. More precisely, their beliefs about the state of the world are totally wrong but they believe that they are perfectly informed, and they therefore apply a deterministic one-shot rationality to their strategies. Preserving for the moment the constraint of synchronism among moves, this simple modification highlights the importance of the assumption of common knowledge. One notes how crucial in this new informational configuration is the order in which the game is played. Different game sequences, in fact, can give rise to completely different outcomes as is clearly shown by the trees of the players' individual paths in fig. 5.

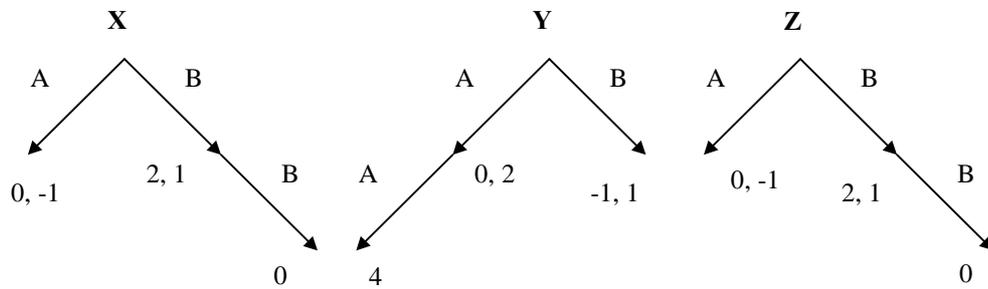
Fig. 5 shows that when the players are unaware of the possibility of the existence and nature of the three grids, and if the game follows the game sequence 1,2,3, agents **X** and **Y** will both choose to play A, the which, because of the constraint of synchronism among moves, they will also apply to grids 2 and 3. On grid 2, independently of her/his knowledge of the existence of grid 1, **Z** will choose her/his min-max strategy, that is (B). However, the structure of grid 2 is reversed with respect to grid 1, and **Z**'s dominant strategy is B. This, combined with the choice already made by X on the first grid, entails a loss of (-1) for **X** and a win of (2) for **Z**. On the other hand, on grid 3, again because of the constraint of synchrony among moves, **Z** will have a loss of (-1) and **Y** a win of (2). The cognitive constraint designed in the form of ignorance about the number of states of nature, the operation of the synchrony constraint, the number and the positions of the agents (in the rows or in the columns) and the starting grid determine the degree of success of the players' min-max strategies.

**Fig. 5 Individual paths**

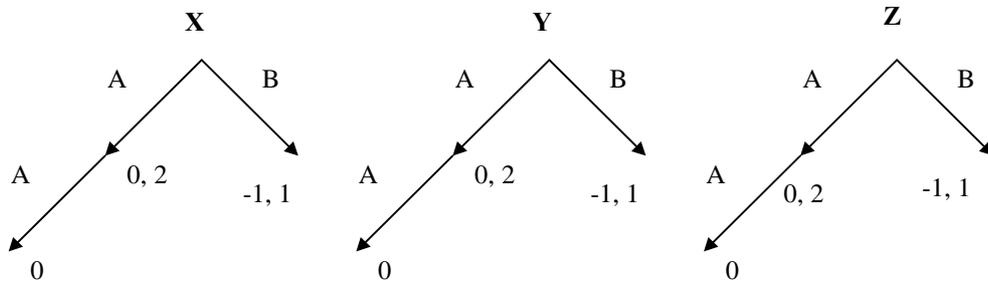
Sequence grids 1,2,3



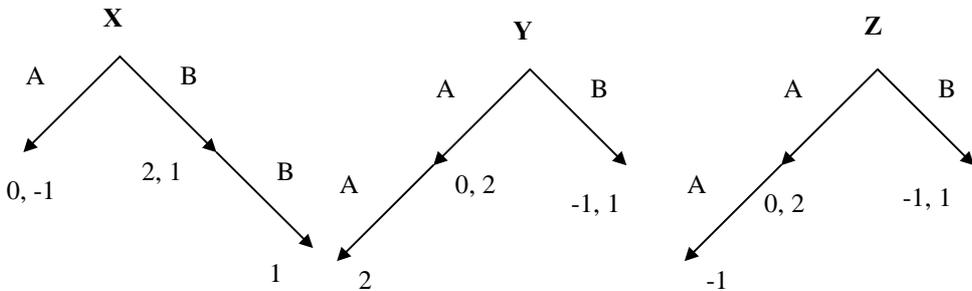
Sequence grids 2,1,3



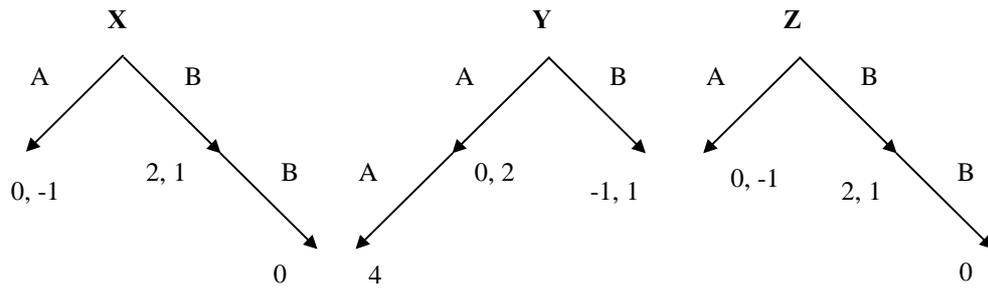
Sequence grids 3,1,2



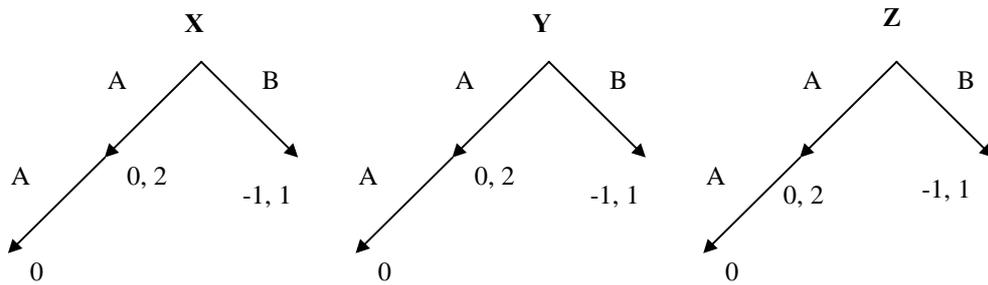
Sequence grids 3,2,1



Sequence grids 2,1,3



Sequence grids 1,3,2



Note that, given the structure of the game, **Y** is the player in the best position because s/he has a payoffs set equal to  $(2, 4, 0, 2, 4, 0)$ . **Z** and **X** follow with equal payoffs sets which are respectively:  $(-1, 0, 0, 1, 0, 0)$  and  $(1, 0, 0, -1, 0, 0)$ . As we have just seen when commenting on the extended form shown in fig. 4, the Nash equilibrium of the game with common knowledge gives a payoff of 4 for **Y** and 0 for both **X** and **Z** which obviously corresponds to one of the possible outcomes of the cognitively constrained game. Restoring the conditions of complete rationality can prove to be desirable for all three agents, given that in the presence of the cognitive constraint, as defined here, none of them could be certain of their own strategic position and of the sequence of the game. In the absence of information about the number and features of the grid, and if the three agents do not know their relative positions, it is therefore preferable for all of them to restore the common knowledge conditions.

Admitting that the probability of occupying **X**'s position is equal to the probability of occupying **Y** or **Z**'s positions, and that the sequences of the game are chosen randomly with equal probability, the expected payoff structure is given by the average of the payoffs of each sequence, as shown in table 1.

**Table 1**

Sequence	<b>X</b>	<b>Y</b>	<b>Z</b>
1,2,3	-1	2	1
2,1,3	0	4	0
3,1,2	0	0	0
3,2,1	1	2	-1
2,3,1	0	4	0
1,3,2	0	0	0
<b>Average</b>	<b>0</b>	<b>2</b>	<b>0</b>

The Nash equilibrium payoffs structure of the common knowledge game, (0, 4, 0), represents a Pareto improvement if compared with the expected payoffs structure, which is (0, 2, 0), calculated for the bounded rational version of the game. It follows that all the players will benefit from a correction of their informational constraint, given that none of them can be sure of her/his position (e.g. if they occupy the position of player **X** or **Y** or **Z**). Furthermore, the game sequence  $\mathbf{X}_B, \mathbf{Y}_B, \mathbf{Z}_B$ , which gives a payoffs structure equal to (2, 2, 2), this time represents a Pareto improvement when compared with the expected payoffs structure obtained by applying the bounded rationality min-max strategy. This result means that under conditions of bounded rationality all the players will benefit both from an intervention that simply eliminates their cognitive constraint and/or from the introduction of some form of exogenous rule which guides their behaviours towards a payoffs balanced (equitable) outcome.

We have so far considered a static, one-shot, game. What changes occur if we move to the situation illustrated by our example, i.e. the hockey players' dynamic context?

### 5. The dynamic game (with finite repetitions)

The infinitely repeated version of the game sheds clearer light on the role played by the cognitive constraint. Returning to the common knowledge context and looking at fig. 4, it is evident that the mechanical repetition of the Nash equilibrium for the static game (i.e.  $\mathbf{X}_B, \mathbf{Y}_A, \mathbf{Z}_B$ ) does not necessarily coincide with the subgame-perfect equilibrium if the discount factor of **Z** and **X**'s objective functions are sufficiently close to 1. Both **X** and **Z** can in fact apply a classic trigger strategy designed to "punish" **Y** whenever s/he plays (A). If (A) is played by both **X** and **Z**, **Y** will receive a 0 payoff, and there is no change in the payoffs of both **X** and **Z**, which remain the same as the Nash equilibrium for the static game. The problem is that, to be effective, this punishing strategy requires co-ordination between **X** and **Z**, who must both play (A). On the other hand the same punishment

strategy can be applied by **X** against **Z** and *vice-versa*. The final expected result is therefore similar to that described in the introduction with respect to the Folk Theorem.

A similar result, but without it being necessary to assume anything about the players' patience, can be obtained by means of self-enforceable agreement within a coalition of players. Admitting that the players can communicate before the game it is possible to demonstrate the existence of an efficient self-enforcing agreement (Bernheim, Peleg, Whinston, 1987). Unfortunately also this self-enforcement agreement requires the assumption of unlimited pre-game communication, i.e. it presupposes a potentially very high cognitive ability to handle information.

To return to the bounded rationality context (that is, to maintain the general assumptions made in the static game) in a dynamic environment, one way to analyse the results that emerge from the game is to carry out a numeric simulation, using the following specific assumptions :

H<sub>3</sub>) the players can memorise the results from the previous games ;

H<sub>4</sub>) the players always follow a min-max strategy weighted by the past experiences.

We are interested in analysing the following topics:

- a) the final payoffs achieved (after a finite number of repetitions of the game);
- b) the path followed to achieve this result;
- c) the role played by memory depth.

The choice model for the generic player that we have used is the following :

$$\alpha_{A,B} = \max\left(\frac{\min \hat{\pi}_{A_t}}{t} + \bar{\pi}_A, \frac{\min \hat{\pi}_{B_t}}{t} + \bar{\pi}_B\right) \quad [4.1]$$

where :

$\alpha_{A,B}$  = actions

$\pi_{A_{t-1}}$  = actual payoff obtained from choice A at time t-1 (with t = 1,...,T)

$\pi_{B_{t-1}}$  = actual payoff obtained from choice B at time t-1

$\hat{\pi}_{A_t}$  = observed theoretical payoff from choice A at time t

$\hat{\pi}_{B_t}$  = observed theoretical payoff from choice B at time t

and,

$$\bar{\pi}_A = \frac{\sum_{i=1}^T \pi_{A_{i-1}}}{t};$$

$$\bar{\pi}_B = \frac{\sum_{i=1}^T \pi_{B_{i-1}}}{t}$$

The rationale for the model is the following: suppose that the player in the first round (time = 1) of the game chooses action A by applying the same cognitively constrained min-max logic described in the static game, then in the second round s/he knows that the payoff for choice A does not necessarily coincide with her/his observed theoretical payoff, i.e. with the payoff that s/he was expecting. Furthermore s/he knows that the reason for this violation of her/his expectations is due to her/his incapacity to govern the complexity of the decisional environment, because her/his opponent has played according to the traditional min-max strategy as foreseen by her/himself. From round 1 onwards the player must therefore keep account of the possibility of obtaining an unforeseen payoff. The simplest way to model the treatment of this kind of uncertainty is to make a correction of the observed theoretical payoffs structure by also computing the historic average payoff obtainable by playing the given choice.

Given the structure of model [4.1] we then assume [H<sub>5</sub>] that each past value of the given choice is valued by the player in the same manner, i.e. with the same weight. The main consequence of this hypothesis is that, as time passes, the relative importance of the observed theoretical payoff from the given choice at time t becomes increasingly weaker. This means that the agent applies a sort of simplified Bayesian rule by attributing a fixed additive weighted probability to each event, independently of its place in time. Obviously this assumption can be criticised from many perspectives, and in particular by arguing that human beings are more attentive to choices related to contemporary or recent consequences on their welfare than to decisions taken in a far past.

This objection can be discussed by referring to the ample literature on the experimental approach to learning processes. This, however, would go beyond the scope of this article. I therefore merely stress that the model can be easily adapted in order to check other behavioural hypotheses. The only adaptation of the model presented here is the one introduced to verify the role played by memory. For this purpose, I tested a model with limited memory, which entailed reducing the number of actual payoffs remembered by the players. The limited memory results were then compared with those obtained using agents with unlimited memory.

The numeric simulations were computed considering a random sample of all the possible structures of the game (changing the starting grid and the grids sequences) yielded by 50 different game stories for each player. I then analysed the decisional patterns for 50 repeated rounds. With reference only to agents **X** and **Y** (because of the structure of the grids the outcomes for **Z** are the perfect opposite of those obtained for **X**) the results obtained for agents with no memory at all, for those able to remember up to 5 rounds, and for those able to remember 25 rounds, are shown in figures 6, 7 and 8, where each line represents a single story.

Inspection of figs. 6 and 7 shows the cyclical dynamic of the behaviours of both **X** and **Y**, which never converge on a static outcome. A different conclusion is drawn from analysis of fig. 8, namely that between rounds 18 and 27 behaviours become steady. The

spontaneous emergence of a stable Nash equilibrium after a given number of repetitions of the game is simply confirmation of the theoretical premises of the model, and therefore comes as no surprise. The interesting insight yielded by the numerical simulation is that stability requires players able to memorise no fewer than 18 rounds. Furthermore, when the players' memories are refreshed (i.e. some items of information are replaced by new ones) the equilibrium once more becomes unstable.

Fig. 6 Payoffs dynamic (players without memory)

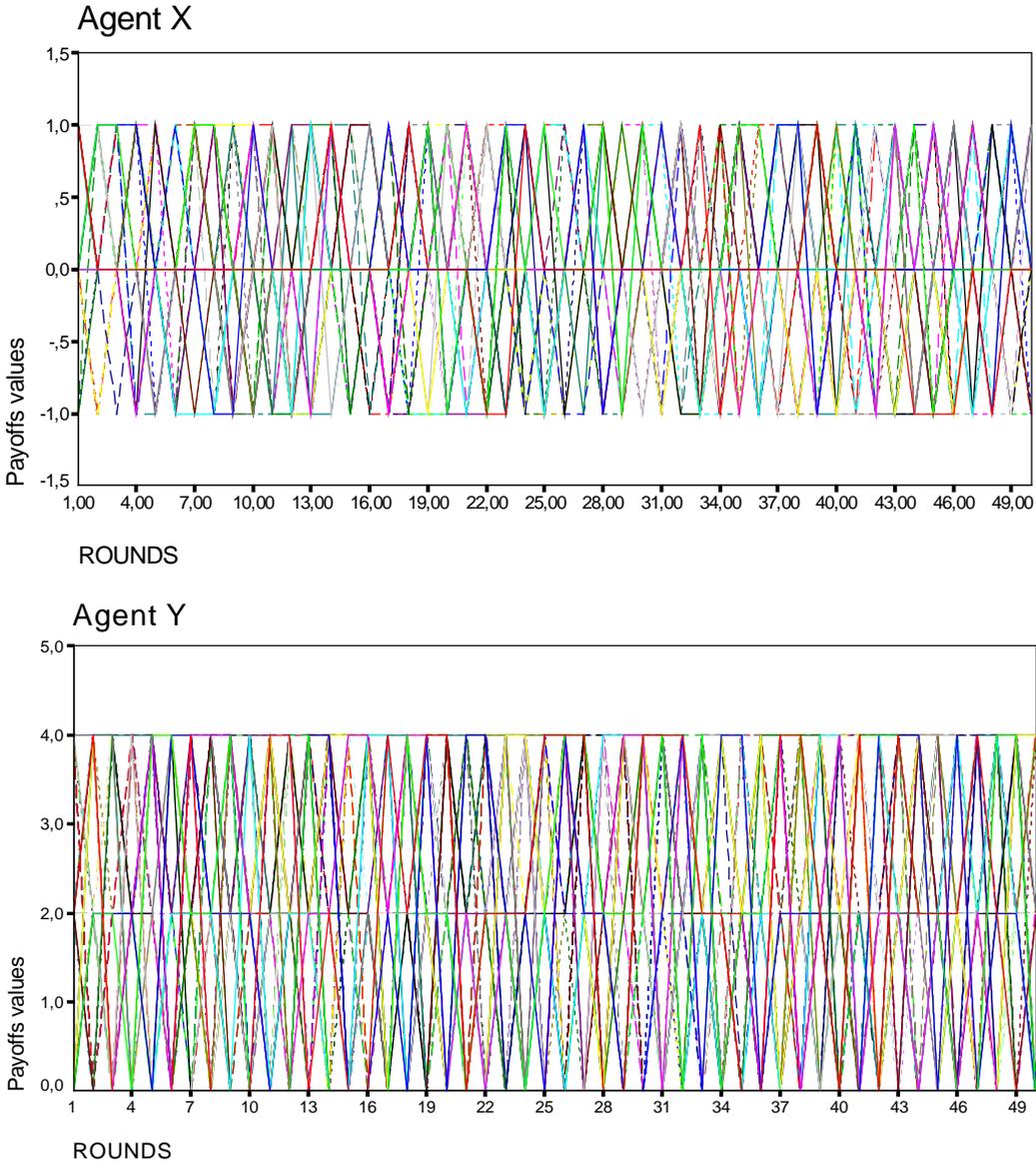
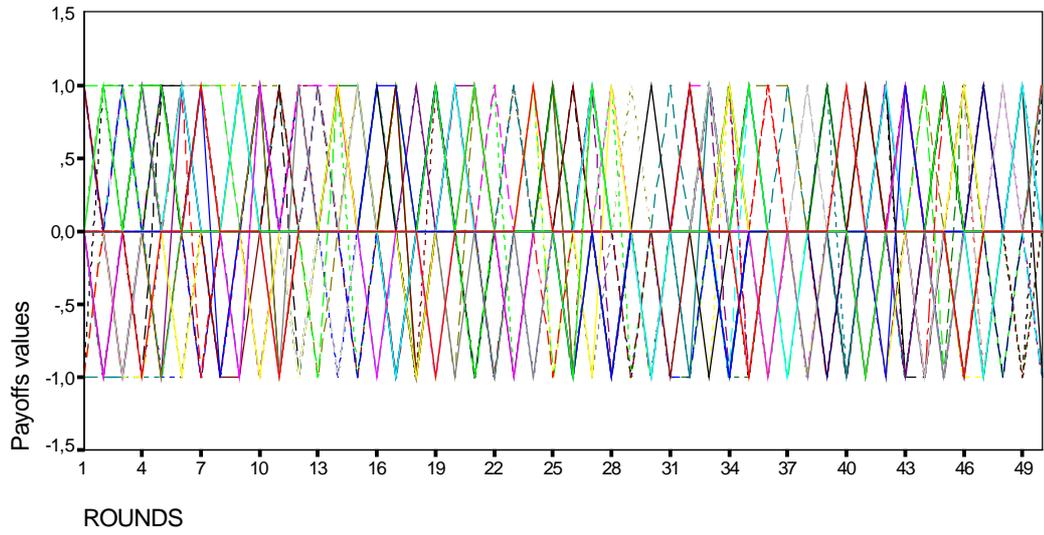


Fig.7 Payoffs dynamic (players with 5 rounds of memory)

Player X



Player Y

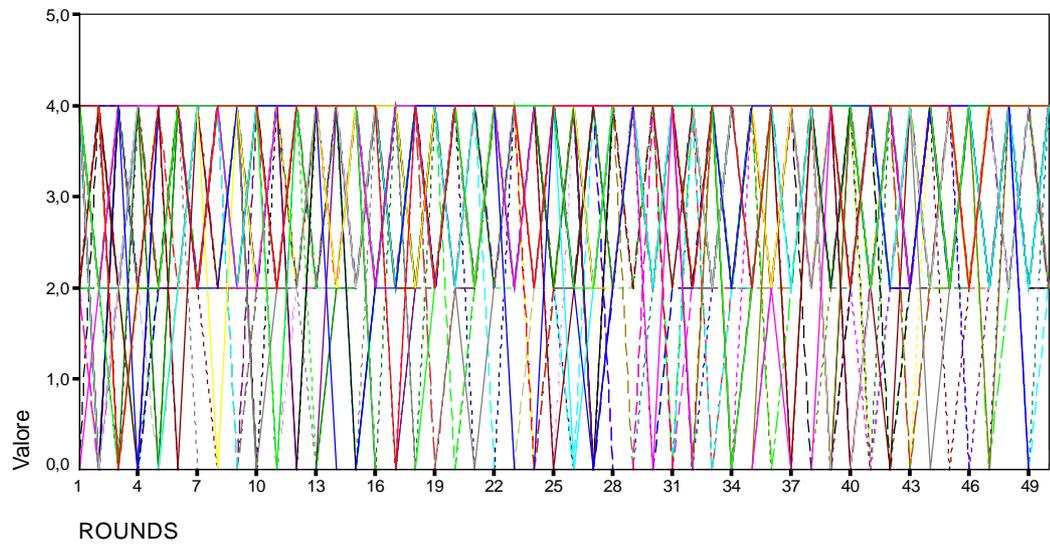
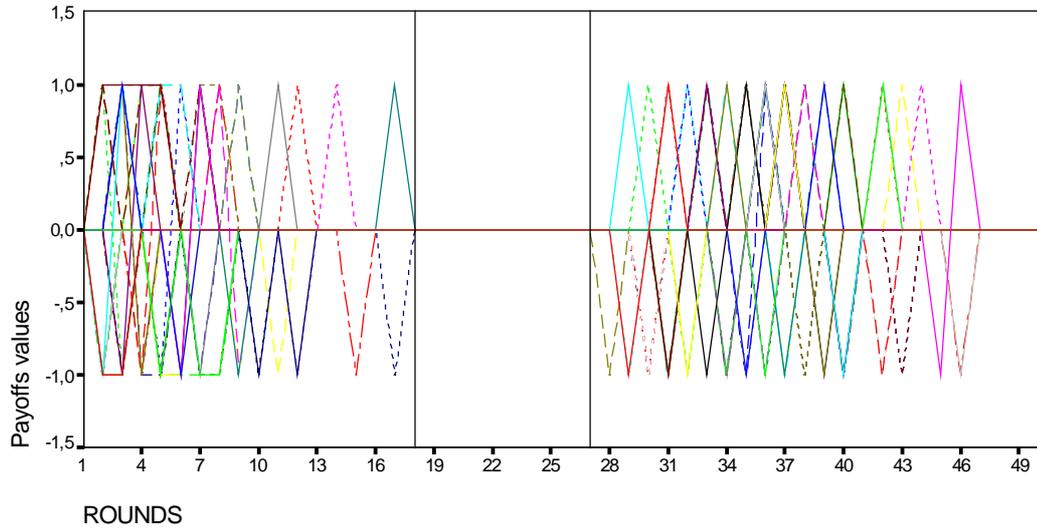
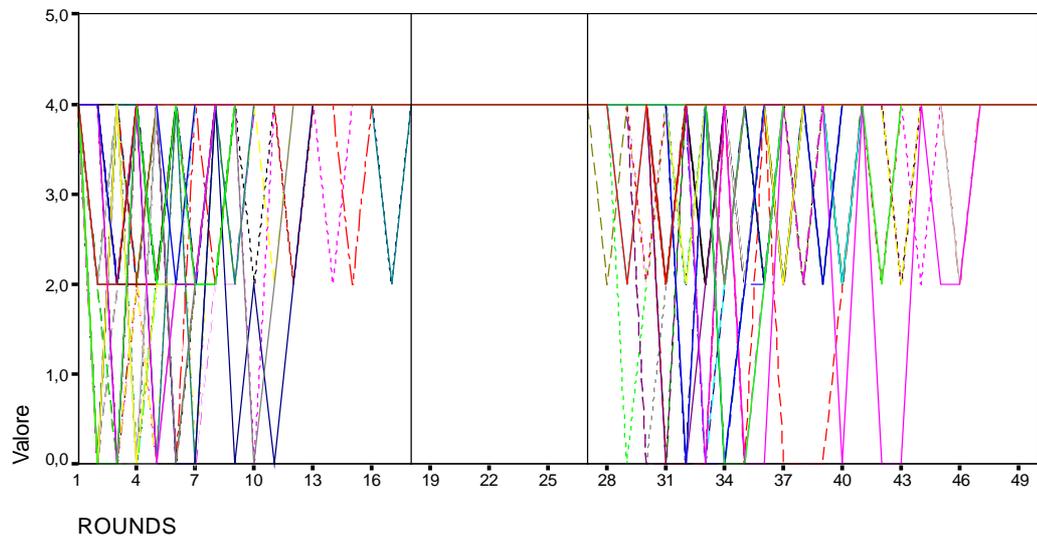


Fig.8 Payoffs dynamic (players with 25 rounds of memory)

Player X



Player Y



The results obtained from the unlimited memory players are shown in fig. 9 (note that the time series were interrupted at round 21 because, as we already know, the results become steady after round 18). Although the high number of stories reported in fig. 9 make it impossible to conduct visual analysis of individual trends, one nevertheless notes that many of the decisional patterns converge on the (0, 4, 0) static common knowledge result after 8-10 rounds (more precisely 38% of stories become steady between rounds 3 and 6. 24% between rounds 7 and 8, another 12% of stories converge on the common knowledge result between rounds 9 and 10, and the remaining 26% of stories require more than 11 rounds before they become steady ).

Fig. 9 Payoffs dynamic (players with unlimited memory)

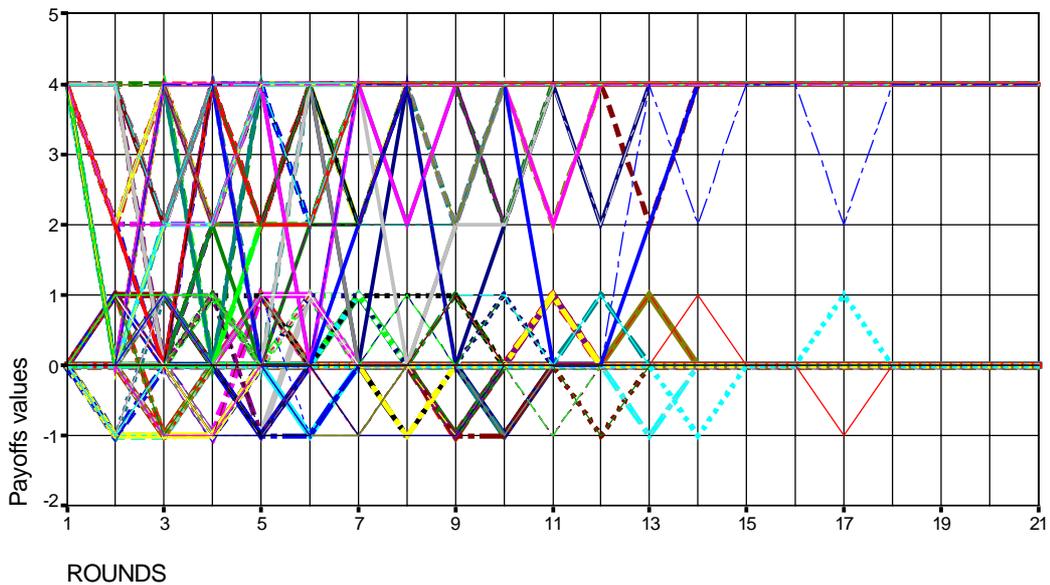
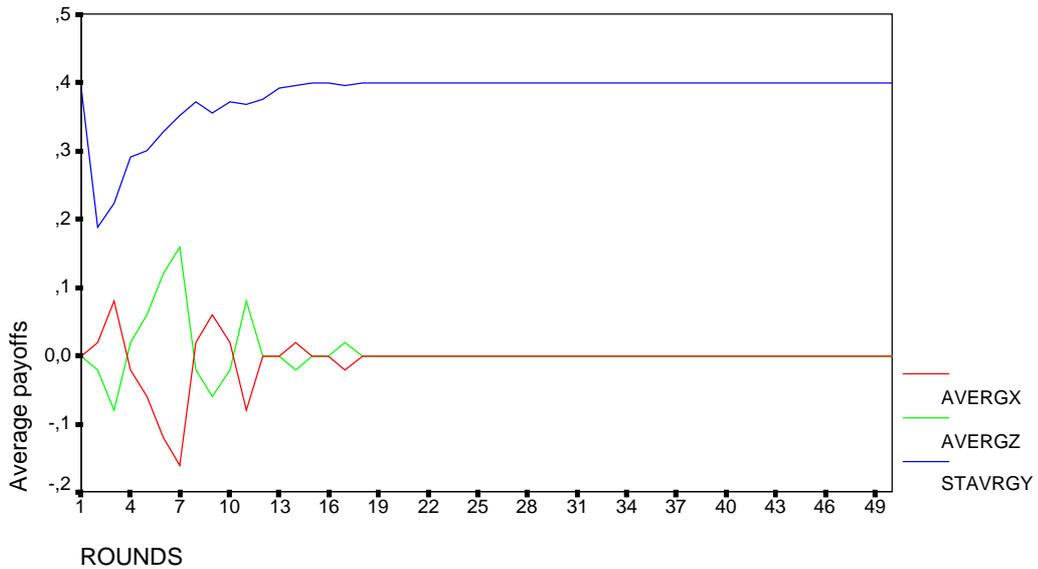


Fig. 10 Average payoffs (players with unlimited memory)



A clearer picture of this phenomenon is provided by fig. 10, which sets out the trends of the average individual payoffs (note that, for simplicity, the visual comparisons between **X**, **Z** and **Y** trends, **Y**'s values have been divided by 10). Analysis of this last figure makes it even clearer that, even when players have unlimited memory, the existence of a cognitive constraint is harmful to all players. Supposing that the game does not stop after 50 rounds but, for example, at round 12, the average payoffs for the three players were (0, 3.76, 0) which is a worse result than the common knowledge one (0, 4, 0). More precisely, if we were to compute the algebraic sum of the individual payoffs we would need to wait until round 15 before obtaining a result equal to the common knowledge one.

To sum up, the main conclusion to be drawn from both the static and the dynamic variant' of the classic P-D is fundamentally the same: even when players have unlimited memory, and admitting that they can repeat the game for more than one round, there is nevertheless room for a Pareto improvement, achievable simply by re-establishing the conditions of common knowledge.

On this basis we may now undertake brief comparative analysis of the instruments that an institutional, specialised agent delegated by the players can activate in all the economic contexts where conditions analogous to those described above arise. I shall identify such institutional agent in the state.

## **6. Intervention by the state: the issuing of norms versus the production of information as a public good**

Returning for a moment to the classic non-repeated P-D, the two mechanisms which ensure solution of the dilemma (that is, the devices which allow the players to approach the Pareto optimum) are the following:

- the production of information about the effective choices of the two agents;
- the drawing up of behaviour-regulating agreements (private contracts or laws)

The state is able to produce mechanisms of both kinds, whereas the private sector can, subject to certain restrictions, produce only devices belonging to the second category. I now briefly examine these various mechanisms.

The first group of tools includes every mechanism able to inform simultaneously, to the same extent, and invariably, both agents about each other's behaviour. In this sense information is a pure public good and its production is one of the tasks conventionally attributed to the state. An example of a device of this kind would be a loudspeaker which announces the moves of each player at the moment when they are made. On the assumption that both players are allowed to correct their choices in real time, this gives a result analogous to the one generated in the dynamic-repetitive context described by

Axelrod's experiment; that is, spontaneous convergence towards the Pareto optimum. Note that the production-diffusion of information must necessarily be a pure public good: if, instead, it were selective - that is, if it discriminated between the agents by placing one in a state of uncertainty about the choices of the other - it would become totally ineffective. In the context analysed here, in fact, the provision of a selective information service creates margins of non-perfect exclusion from the benefits of the service; margins which derive from the fact that the agent excluded from the information knows that his opponent has acquired (or may have acquired) the information service and is therefore aware of his moves. In this case the only possible strategy available to the player excluded from information is once again A, since this is the only strategy that insures him against the risk of receiving a -1 payoff. Of course, in the two-player context the production of non-differentiated information can be decided and managed by the agents themselves, although they still have to devise a mechanism which guarantees to both that their information is genuine. When the number of agents increases, the guarantee mechanism may take the form of interpersonal agreements, the application of which requires the creation of a collective institution of control similar to the state - in the sense of the term used here.

The second kind of instrument which induces convergence towards strategy B is the production of agreements, in the form of either norms (laws) or of private contracts between the agents involved in the dilemma. As is well known, crucial in the use of this kind of instrument is the choice of a system which adjusts the payoff structure (i.e. an enforcement system) so that strategy B predominates over strategy A (Ullman Margalit, 1977). Although this type of correction mechanism can be included in a private contract, it is very probably that the agents will also need some form of external control over compliance with the agreement, and this is a function which is normally performed by the state. This control function becomes well-nigh indispensable when we pass from the two-agent context to a situation with a large number of interacting agents; a situation in which, as we have seen, the nature of the information structure can easily change and thereby create the problem of cognitive limits described in the previous section.

Moving to the context of the three agents operating in conditions of bounded rationality, and remembering that we have assumed that these conditions are produced by incomplete *local* information or by a subjective inadequate ability to process information, the potential role of the state expands, and the two instruments just described take different forms with respect to the classic P-D situation. First, the production of information can be extended to the restoration of the players' conditions of common knowledge through the diffusion of data about the nature of the possible states of the world - that is, in our example, the number of players and the layout of the grids. This is possible to an external, specialised agent only if the source of bounded rationality is the narrowness of each agent's cognitive perspective, *i.e.* is a consequence of a subjective constraint. Negative externalities like pollution are a good example of this kind of cognitive distortion: the individual agent generally has only a very partial representation of

her/his own contribution to the total amount of pollution, but a specialised agent (like an environmental agency) can achieve much more precise understanding of the global phenomenon.

Given, however, that the realisation of conditions of common knowledge is associated, or may be associated, with a change in the players' potential payoffs, the degree of desirability of this kind of intervention will no longer be identical for all the agents - as it was in the case of the classic P-D - but will change as each agent's assessment of her/his position changes. It is evident, in fact, that if an agent believes that s/he has an advantage over her/his opponents because of the operation of the cognitive constraint, s/he will not at all be in favour of an intervention designed to remove the conditions that enable her/him to enjoy this advantage. On the other hand, if an agent believes her/himself to be damaged by the operation of the uncertainty constraint on the possible states of the world, s/he will seek to have it lifted. Before analysing the conditions under which the decision to restore extended rationality can be reached, it should be stressed that here too, as in the case of the diffusion of information about the players' moves in the classic P-D, the production of information should come about through a collective accord.

Let us begin by supposing that only one agent wishes to know the decision-making context and that s/he commissions a research company to find out the number and features of the grids and the order of play. Of course, the results of this research must remain confidential, as must the name of the player who commissioned the study.

Reintroducing, for the purpose of illustration, the example of the three - player game played on asymmetrical grids, let us assume that the research has been commissioned by **X**, and that the order of play is 1-2-3. It is immediately clear that the unilateral lifting of the cognitive constraint will, in this case, not be of any advantage to **X**, since s/he cannot modify her/his choice A. Choosing B will incur to a loss of -2 as the other players **Y** and **Z** will play respectively A and B on grid 1 and grid 2. The only way for **X** to take advantage of information purchasing is to inform the other two players as well, in order to achieve the common knowledge conditions. On the other hand if information is purchased by **Y** it will inform her/him on the advantage that s/he obtains from the incomplete information conditions of the other players.

Although, of course, it is possible to imagine other cases, the above two suffice to show that a selective distribution of information about the possible states of the world does not necessarily improve the decision scenario for all agents. Hence it follows that selective information does not always guarantee advantages for its purchasers unless those who buy this information are unable to divulge it to the other players - a situation which can only arise when the number of possible agents is such to ensure sufficient diffusion of information. In this case too, however, the information scenario may be so complex as to render the information useless, so that whoever decides to buy information may in any case incur a loss.

Returning, therefore, to the initial situation in which no player can be certain of the strategic strength of her/his position, and having established that there are situations in which the purchase of selective information does not produce guaranteed advantages, the only solution to the dilemma is the financing of a programme of public information diffusion which - bearing in mind the conclusions reached in the previous section - is able to guarantee an improvement in the initial game conditions of all the players. Accordingly, the production of information in order to remove the cognitive constraint leads to a Pareto-superior improvement in the decision scenario *ex-ante* individual choices. However, in order to convince all the players that this is to their advantage, there should also be a normative policy which ensures co-ordination among the players' moves such as to ensure a Pareto-efficient result also *ex-post* the restoration of the conditions of common knowledge. One notes, in fact, that the first of the two instruments described at the beginning of this section - i.e., the production of information about the players' moves - is not enough, in both the static and the dynamic contexts, to guarantee the outcome that was achieved in the two-player P-D. The reason for this is the asymmetry of the players' positions, which prevents the "automatic" equal-proportion distribution - as in the P-D with only two agents - of the costs presumably necessary to finance the two information programmes (the one intended to restore common knowledge, the other to broadcast information about the players' moves).

In verifying this statement, I begin with the static case. Here, as we saw, the passage from a context of bounded rationality to one of common knowledge created a situation where two players (**X** and **Z**) obtained a payoff of 0, whereas the third (**Y**) obtained a payoff of 4. Recall also that the only possible improvement in a Pareto sense left the situations of the first two agents unaltered and was only to the advantage of **Y**. Hence it follows that there is no inducement to **X** and **Z** to finance a further programme of information gathering (about moves), since this would be to their detriment if the costs of its implementation were borne equally by all three agents. On the other hand, the first programme presumably also involves costs which, if equally divided among the agents, could convert **Y**'s payoff into a loss if compared with the 2 average payoff obtainable "automatically" in a bounded information environment. In contexts analogous to the P-D proposed here, therefore, the information production instrument should be combined with a system of agreements on the distribution of costs, since these should not be borne by all the players indiscriminately but, as the occasion arises, only by agents who register positive payoffs.

Turning to the dynamic context, it will be remembered that the need to reach a agreement, that is, to introduce a norm governing inter-agent behaviour, becomes unavoidable once extended rationality conditions had been restored - given that **Y** had no incentive to change her/his strategy. Compared with the static context, therefore, in dynamic conditions it is even more necessary to flank the information instrument with a

norm or an agreement that guarantees that the players excluded from the benefits of co-ordination are not only exempt from the costs of the scheme but also receive incentives.

There now arises the problem of why a collective agreement, and not a private contract among the actors, should necessarily arise in order to implement the system of norms and incentives required to achieve a Pareto-efficient solution of the dilemma. Obviously, if the agents are only **X**, **Y** and **Z**, a private contract will be the ideal co-ordination instrument, because it can give efficient definition of the value of co-operation between **X** and **Z** through presumably straightforward and cheap negotiation (because it is restricted to only three agents). Once again, however, this conclusion does not apply when the number of agents increases; in other words when the cognitive constraint is the result of the multiplication of the number and the features of the game grids - a situation which coincides precisely with that described at the beginning of this work and of which the example with three players then proposed was an extreme simplification. When, in fact, conditions of bounded rationality are due to the complexity of the information structure of the decision problem, one certainly cannot be sure that it will be eliminated simply by the production of information and subsequent private negotiation. In all these cases, it becomes necessary to draw up collective agreements: that is, to adopt a system of norms able to bring subjective choices closer to Pareto-efficient situations even in the absence of the explicit (i.e. negotiated) involvement of all agents. Thus private contracts are preferable to norms in all cases where they can be substituted for them, because they produce a unanimous agreement whereas norms generally lead to majority decisions.

## **7. Conclusions**

The foregoing analysis has shown the existence of decision contexts in which the operation of a cognitive constraint - represented by ignorance of the possible states of the world - may generate subjectively and collectively inefficient choices. We have seen that the co-ordination of individual choices makes it possible to achieve a Pareto-superior situation, although this will require agents to reach agreements among themselves. Restoring the conditions of extended rationality by producing information about the states of the world is not, in fact, enough to secure spontaneous co-operation among agents unless the nature of the decision problem is such that all the subjects involved receive the same payoffs. Thus the rise of the state can be viewed as a collective response to the impossibility of spontaneously generating co-operative behaviour in situations of bounded rationality and of unequal individual endowments.

The necessity or otherwise of resorting to an interpersonal valid norm - as opposed to a private accord (that is, an agreement restricted to just a few agents) - depends on the diffusion and importance of the effects of a particular decision on society. The more widespread and the more momentous these effects, the more necessary it becomes to

introduce a norm that applies to all members of society without exception, and not to rely on a private agreement. This necessity arises from the fact that it is impossible to devise specific accords (that is, private contracts) when the number of interacting agents and the number of the effects they produce exceeds the capacity for rational analysis of the agents themselves. This situation, in fact, is analogous to that of the ignorance of the complete picture of the states of the world that requires, as I have stressed, the guided co-ordination of agents. It is a situation, therefore, which entails general agreements, or laws, which operate as approximations to private contracts when these prove impossible to draw up and apply in all circumstances.

## References

- Axelrod R (1984), *The evolution of co-operation*, Penguin Books, London.
- Bernheim, Peleg, Whinston (1987), "Coalition - Proof Nash Equilibria : Concepts", in *Journal of Economic - Theory*, 42(1) ; pp. 1-12.
- Bobbio N. - Matteucci N. (1976), *Dizionario di politica*, UTET, Torino.
- Coase R.H. (1937), "The nature of the firm", *Economica*, n. 4, pp. 386-405.
- Coase R.H. (1960), "The problem of social cost", *The Journal of Law and Economics*, vol. III, pp. 1-43.
- Egidi M. (1991), "Dalla razionalità all'apprendimento organizzativo: l'analisi del comportamento umano nelle organizzazioni economiche"
- Friedman B. "A Non-cooperative Equilibrium for Supergames ", *Review of Economic Studies* ; 38(113), pp. 1-12.
- Fudenberg D., - Tirole J. (1991), "Game Theory", Cambridge, Mass. ; London : Mit press
- Fudenberg D. Kreps, Maskin (1990), "Repeated Game with long - run and Short - run Players", in *Review of economic - studies*, 57(4), pp. 555-73.
- Gibbons R. (1992), "A primer in game theory", New York : Harvester - Wheatsheaf..
- Hayek F.A. (1937), "Economics and knowledge", *Economica*, n. 13, pp. 33-54.
- Hayek F.A. (1945), "The use of knowledge in society", *American Economic Review*, vol. 35, n. 4, pp. 519-530.
- Hayek F.A. (1949), *Individualism and economic order*, Routledge & Sons, London.
- Heiner R.A. (1983), "The origin of predictable behaviours", *American Economic Review*, vol. 73, n. 4, pp. 560-595.
- Schelling T.C. (1973), "Hockey helmets, concealed weapons, and daylight saving", *Journal of Conflict Resolution*, vol. XVII, n. 3, pp. 381-428.
- Schiera P. (1976), "Stato moderno", in N. Bobbio, N. Matteucci (1976), *Dizionario di politica*, op. cit..
- Simon H.A. (1955), "A behavioral model of rational choice", *Quarterly Journal of Economics*, vol.49, n.3.
- Ullman Margalit E. (1977), *The emergence of norms*, Oxford University Press, Oxford.