Observable Strategies

Karl Wärneryd*

April 14, 2009

Abstract

The idea that reciprocal cooperation can be obtained in a one-shot game if the players observe each other's strategies before taking action has attracted many authors. From a strictly logical perspective, however, there cannot be such games. Nevertheless, there are games in which each player can observe which class, out of a collection of classes smaller than the number of strategies, the opponent's strategy belongs to. For any underlying 2-player, finite, normal-form game there is a game extended with such coarsely observable strategies that has equilibria with payoffs arbitrarily close to any feasible, individually rational payoff profile.

1 Introduction

Suppose that before playing, say, the Prisoners' Dilemma, a player could observe his opponent's strategy, and the opponent his. A common intuition suggests that there is a Nash equilibrium in this situation such that each player chooses to cooperate if the opponent is observed also to be a conditional cooperator, and defects otherwise. That is, the claim is that there is an equilibrium

*Department of Economics, Stockholm School of Economics, Box 6501, SE-113 83 Stockholm, Sweden, and CESifo. Email: Karl.Warneryd@hhs.se. I thank Ken Binmore, Hannu Salonen, and Robert Sugden for helpful remarks on earlier, considerably different, versions. This is still a preliminary version.

in which each player cooperates if the opponent is observed to be playing the *same* strategy, and defects otherwise.

A Nash equilibrium is a set of strategies, one from each player, such that each player's strategy is a best reply to those of the others. One reason the argument given above seems compelling is because it appears to allow us not even to have to think very carefully about which alternative strategies might be available—the proposed equilibrium strategy defects against *all* of them, inexorably leaving the opponent worse off than if he had also played the conditionally cooperative strategy.

But the argument is misleading. Since an equilibrium is a strategy profile such that each player's strategy is an optimal choice from the set of strategies he has available, we must at a minimum specify what the strategy set of a player is. As we shall see, there are no games where strategies are perfectly observable, for perfect observability implies that such a "game" has no strategy sets. Anytime we thought we had such a set, we would be able to find a new way of conditioning behavior on the strategies in the set, one that was not already in the set. While this problem has, to one extent or another, been noted before,[1] we then go on to show how games with *coarsely* observable strategies may be constructed, and how such games may have exactly the properties desired.

The idea of strategic observability or transparency allowing for reciprocal cooperation has a long pedigree and turns up in many different contexts. Inspired by von Neumann and Morgenstern's [26] discussion of "majorant" and "minorant" games, Howard [16] considered "meta-games" in which a player is allowed to make his action choice contingent on the strategy of his opponent, and suggested that this approach solved some problems with the game-theoretical notion of rationality. Danielson [6] does computer simulations based on this idea, studying a population of programs that have different levels of meta-knowledge about their opponents, an approach also related to the theory of level-k reasoning of Stahl [24] and Crawford [5]. In his influential work on moral philosophy, Gauthier [12] suggests that "true" rationality in a Prisoners'

[1] See, e.g., Binmore [3] and Rubinstein [23].

Dilemma must involve the desire to be a conditional cooperator and to make this disposition publicly observable.

In economics, Frank [10, 11] argues at length that a form of transparency of human agents should be expected to be a factor in social interaction, since evolutionary forces would have favored the development of physical characteristics that reliably signal a person's disposition or strategy. Actual evidence of this being the case includes the fact, reported by Ekman [8], that it is difficult to lie without giving it away through facial expressions that are beyond conscious control. Ockenfels and Selten [21], Fehr and Fischbacher [9], and Manzini *et al* [19] are examples of studies of the transparency notion in a behavioral economics context.

McAfee [20], Binmore [2], Anderlini [1], and Canning [4] introduced the idea of studying games played by Turing machines that input each other's descriptions before play. Building on this idea, Howard [15], Vulkan [27], and Tennenholtz [25] argue that it is possible to write programs that recognize copies of themselves and suggest that this allows such programs to be conditional cooperators. (See also Rubinstein [23].) Under this interpretation, the approach has important applications to automated trade and other transactions performed by computers. Kalai *et al* [17] study very general commitment or delegation devices and prove a "Folk Theorem"-like result. Levine and Pesendorfer [18] consider games where the players observe a signal about each other's strategies before play. The approach in the present paper is considerably simpler than these contributions, but produces similar results.

As we have argued, and shall show formally next, it is not possible to let decision rules be completely observable and simultaneously allow all logically possible decision rules. Approaches such as those of the Turing-machine school or Howard, Vulkan, and Tennenholtz get around this problem by considering only strategies that can be written down as computer programs, and that of Kalai *et al* by otherwise restricting the way in which a player may condition his action choice on what he observes about other players. In this paper we instead consider restricting the *information* available to a player about his opponent's

		Player 2	
		c	d
Player 1	c	$2,2$	$0,3$
	d	$3,0$	$1,1$

Table 1: The Prisoners' Dilemma.

strategy, while allowing players to condition their choices in any way they like on the information they do have. If a player can observe which class, out of a collection of classes that coarsely partitions the set of strategies, the opponent's strategy belongs to, then for any underlying 2-player, finite, normal-form game there is a game extended with such coarsely observable strategies that has equilibria with payoffs arbitrarily close to any feasible, individually rational payoff profile.

2 The Impossibility of Perfect Observability

Consider the familiar Prisoners' Dilemma (PD) game of Table 1. This game has the property that playing action d is a strictly dominant strategy for each player, in that it alone yields a player his highest payoff no matter what the other player does, while both players would be better off if they both played c. The game is quite trivial from a game-theoretic point of view, of course, being a rare example of a game that has a unique solution in dominant strategies. Nevertheless it has troubled many moral philosophers and political scientists, who see in it a stylized version of the quintessential problem of social interaction—a conflict between individual rationality and the common good.

Gauthier [12] suggests that the prospects for voluntary cooperation in one-shot interactions may not be as bleak as all that. For, he argues, it would be in the interest of the truly rational individual to develop a *disposition* to cooperate, conditional on the opponent having the same disposition, and otherwise play d, and furthermore to make this disposition public. Such an individual

		Player 2	
		CM	SM
Player 1	CM	$2,2$	$1,1$
	SM	$1,1$	$1,1$

Table 2: Gauthier's disposition game.

Gauthier calls a *constrained maximizer* (CM).

Gauthier feels we should really consider the new game, given in Table 2, constructed from the example PD by having it played by individuals who can choose between the CM disposition, which makes its behavior contingent on the disposition of the opponent, and the old, *straightforward maximizer* (SM) disposition that always plays d.

This game has a Nash equilibrium where both players adopt the CM disposition, inducing cooperation. For if one player were to deviate, the constrained maximizer would see this, dispositions being assumed public, and retaliate.

This account of conditional cooperation begs numerous questions. We shall focus on the following one: Where are all the other possible dispositions in Gauthier's game?

For clearly there must be more ways of conditioning behavior on the opponent's disposition than just CM and SM. We can immediately think of two more, one which plays c regardless of whether the opponent is CM or SM, and one which plays d if the opponent is CM and c if the opponent is SM. This makes four possible dispositions so far. But then CM and SM are incompletely specified dispositions, since they do not specify what to do if the opponent has one of the two new dispositions. And so on.

Gauthier's dispositions are not *strategies* of the transparency game in the orthodox sense of *complete contingent plans of action* that specify what to do in every situation that could arise in the game. But before we can tell what the consequences of this omission of possible behaviors from the game are for the possibility of conditional cooperation, we must ask what the *complete* strategy set might look like.

Consider a normal form game G with player set $N := \{1,2,\dots,n\}$ and finite action sets A_i. Is it possible to extend this game into one where prior to taking actions in G, the players observe each other's strategies? That is, is the sentence

> "Prior to taking actions in G, the players observe each other's strategies."

meaningful?

The answer to this question is no. This is easiest to see in a symmetric 2-player game with common action set A. Assume, by way of contradiction, that such an extended game exists, and let I be a player's set of information sets, or information partition, of that game. A pure strategy for a player is a mapping from his information partition into A, so his set of pure strategies is $S := \{s | s : I \to A\}$. Assuming a player observes the strategy choice of his opponent, we must have $I = S$. Hence S is the set of all mappings from S itself into A. The pure strategy set of the hypothetical extended game is therefore self-referentially defined by the equation

$$S = \{s | s : S \to A\}. \tag{1}$$

Does such a fixpoint set S exist? In the trivial case where A has a single element a, it does. Then the unique solution to equation (1) is the set $S = \{s\}$ where $s(s) = a$. In general, however, there is no solution, as can easily be proved using a Cantorian diagonalization argument.

Observation 1 *Suppose we have $|A| \geq 2$. Then there is no set S satisfying the equation $S = \{s | s : S \to A\}$.*

Proof. Suppose there was a fixpoint set S. Consider a new mapping $s' : S \to A$ such that $s'(s) \neq s(s)$ for all $s \in S$. Since A has at least two elements, this construction is always possible. The mapping s' does not belong to S, since it differs from every $s \in S$ at at least one point. So we have a contradiction. Therefore S cannot be a solution to (1). □

			Player 2			
			P_1		P_2	
			$\{cd\}$	$\{cc$	dc	$dd\}$
Player 1	P_1	$\{cd\}$	$2,2$	$3,0$	$1,1$	$1,1$
		$\{cc$	$0,3$	$2,2$	$2,2$	$0,3$
	P_2	dc	$1,1$	$2,2$	$2,2$	$0,3$
		$dd\}$	$1,1$	$3,0$	$3,0$	$1,1$

Table 3: The Prisoners' Dilemma with coarsely observable strategies.

3 Coarse Observability: Examples

Suppose instead that a player can only observe which *class*, out of some collection of classes, his opponent's strategy belongs to. Consider again the Prisoners' Dilemma of Table 1.

Suppose a player's strategy can belong to one of two classes, P_1 and P_2. A pure strategy is now a mapping $s: \{P_1, P_2\} \to \{c, d\}$. Hence each player has four pure strategies. Writing a pure strategy as a string xy, with $x, y \in \{c, d\}$, we may take the first element of the string to be the action planned when the opponent's strategy belongs to the P_1 class, and the second element the action planned when the opponent's strategy belongs to the P_2 class. Suppose further that we have $P_1 = \{cd\}$ and $P_2 = \{cc, dc, dd\}$. We then get the new game of Table 3.

In this particular game extended with coarsely observable strategies, (cd, cd) is an equilibrium. The strategy cd is here, in effect, a conditional cooperator that cooperates against itself and defects against all others.

In this example, the conditionally cooperating strategy cd is, of course, perfectly identified when it is played, as it is unique in its class. This is not necessary, however, in order to generate cooperation in equilibrium. The only essential feature is that the conditional cooperator should belong to a class where all other strategies in the class respond with c against other members of the same class. Consider, for instance, the two-class game where we have $P_1 = \{cc, cd\}$

			Player 2			
			P_1		P_2	
			$\{cc$	$cd\}$	$\{dc$	$dd\}$
Player 1	P_1	$\{cc$	2,2	2,2	0,3	0,3
		$cd\}$	2,2	2,2	1,1	1,1
	P_2	$\{dc$	3,0	1,1	2,2	0,3
		$dd\}$	3,0	1,1	3,0	1,1

Table 4: The PD with different coarsely observable strategies.

		Player 2	
		a_1	a_2
Player 1	a_1	0,0	1,2
	a_2	2,1	0,0

Table 5: The Battle of the Sexes.

and $P_2 = \{dc, dd\}$, as depicted in Table 4. Here (cd, cd) is again an equilibrium. Hence it is not necessary that a strategy be able to identify identical copies of itself—which is the focus of, e.g., Howard [15], Tennenholtz [25], and Levine and Pesendorfer [18]—in order for conditional cooperation to be sustainable. What *is* needed is information about what the opponent plans to do against a cooperator.

Coarse observability also makes correlated equilibrium payoffs attainable. Consider the Battle of the Sexes game of Table 5. This game has two asymmetric pure-strategy equilibria, which the players rank differently, and a symmetric mixed-strategy equilibrium in which each player plays his a_1 action with probability 1/3. In the mixed-strategy equilibrium, each player has an expected payoff of 2/3.

Suppose now that this game is extended with coarsely observable strategies that have three classes for each player. Let $P_1^1 = \{a_1 a_2 a_2\}$, $P_2^1 = \{a_2 a_1 a_2\}$, and let P_3^1 contain all other strategies of Player 1. Let $P_1^2 = \{a_2 a_1 a_2\}$, $P_2^2 = \{a_1 a_2 a_2\}$,

			Player 2		
			P^2_1	P^2_2	P^2_3
			$a_2 a_1 a_2$	$a_1 a_2 a_2$	$\cdots$
Player 1	P^1_1	$a_1 a_2 a_2$	$1,2$	$2,1$	$\cdots$
	P^1_2	$a_2 a_1 a_2$	$2,1$	$1,2$	$\cdots$
	P^1_3	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Table 6: Part of the extended Battle of the Sexes.

and let P^2_3 contain all other strategies of Player 2. Table 6 shows part of the payoff matrix of this extended game.

Now let Player 1 play his strategies in P^1_1 and P^1_2 with probability $1/2$ each, and let Player 2 play his strategies in P^2_1 and P^2_2 with probability $1/2$ each. This is an equilibrium since each player is indifferent between the strategies he assigns positive probability, which each yield an expected payoff of 1.5, and any other strategy yields an expected payoff of at most 1, since the opponent plays a_2 against all other strategies.

In the following section we show how this type of construction allows for the arbitrarily close approximation of any payoff profile in the convex hull of payoff profiles of the underlying game.

4 Coarse Observability: The General Case

Let G be a finite, 2-player, normal form game with action sets A_i for $i \in \{1,2\}$ and payoff functions $u_i\colon A_1 \times A_2 \to \mathbb{R}$. Let $\mathscr{A}_i$ be the set of mixed actions of player i, and in standard fashion extend u_i also to mixed actions.

$G^\star$ is the game formed by extending G with coarsely observable strategies. $G^\star$ associates with each player a finite set $P_i := \{P^i_1, P^i_2, \ldots, P^i_{m_i}\}$, the set of classes of player i's strategies. Before taking action in G, each player observes which class the strategy of the opponent belongs to. A pure strategy of player i who faces opponent $j \neq i$ is therefore a mapping $s_i\colon P_j \to A_i$. Since P_j and A_i are both finite, the set S_i of all pure strategies of player i is well defined and has $|A_i|^{|P_j|}$ elements. P_i partitions S_i into m_i classes. As there are necessarily always more strategies than classes—as long as each player has more than one action available—there is always at least one class that contains more than one strategy, justifying the use of the expression "coarsely observable strategies."

Define

$$\hat{u}_i := \min_{\alpha_j \in \mathscr{A}_j} \max_{\alpha_i \in \mathscr{A}_i} u_i(\alpha_i, \alpha_j),$$

player i's *minmax* payoff. Let V be the convex hull of payoff profiles of G, and let $\bar{V} := \{v \in V | v_i > \hat{u}_i \text{ for all } i\}$ be the set of *feasible, individually rational* payoff profiles of G.

We first show that any feasible, individually rational payoff profile of the underlying game can be approximated arbitrarily closely in an equilibrium of *some* extended game.

Lemma 1 *Let $\bar{u} \in \bar{V}$ be a feasible, individually rational payoff profile of G, and let $B^{\varepsilon}(\bar{u}) \subset \mathbb{R}^2$ be a ball of radius ε at $\bar{u}$. Then for any $\varepsilon > 0$ there is an extended game $G^{\star}$ with an equilibrium with payoffs u such that $u \in B^{\varepsilon}(\bar{u})$.*

Proof. Let μ, μ_1^d, and μ_2^d be such that

1. μ is a probability distribution with full support on $A := A_1 \times A_2$ such that for each $a \in A$, we have $\mu(a) = k(a)/m$, with $k(a)$ and m positive integers,
2. μ_i^d is a probability distribution on A_i such that for each $a_i \in A_i$, we have $\mu_i(a_i) = k_i^d(a_i)/m$, with the $k_i^d(a_i)$ non-negative integers, and
3. for each $i \in \{1,2\}$, $j \neq i$, it holds that

$$\sum_{a \in A} \mu(a) u_i(a) \geq \max_{a_i \in A_i} \sum_{a_j \in A_j} \mu_j^d(a_j) u_i(a_i, a_j).$$

Let there be $m+1$ classes of each player's strategies. Consider now a subset of strategies $\bar{S} \subset S$ such that $(\bar{s}_1^1(P_i^2), \bar{s}_i^2(P_1^1))$ for $i = 1 \ldots m$ is a vector of action profiles such that for all $a \in A$, a occurs exactly $k(a)$ times and in direct succession. Then recursively let

$$(\bar{s}^1_j(P^2_i), \bar{s}^2_i(P^1_j)) = \begin{cases} (\bar{s}^1_{j-1}(P^2_m), \bar{s}^2_m(P^1_{j-1})) & \text{for } i = 1 \\ (\bar{s}^1_{j-1}(P^2_{i-1}), \bar{s}^2_{i-1}(P^1_{j-1})) & \text{for } i = 2 \dots m. \end{cases}$$

Let $\bar{s}^i_j(P^\ell_{m+1})$, for $\ell \neq i$ and $j = 1 \dots m$ be such that for each $a_i \in A_i$, play of a_i is specified exactly $k^d_i(a_i)$ times. Let the P_i be such that for each $i \in \{1,2\}$, we have $P^i_j = \{\bar{s}^i_j\}$ for $j = 1 \dots m$, and P^i_{m+1} contains all player i's strategies not in $\bar{S}_i$. Finally, let each player play a mixed strategy that puts positive and equal probability on strategies in $\bar{S}_i$ and zero probability on strategies not in $\bar{S}_i$.

Given that the opponent plays the specified strategy, each player is indifferent between his strategies in $\bar{S}_i$, which each yield an expected payoff of

$$\sum_{a \in A} \mu(a) u_i(a).$$

If he plays a strategy not in $\bar{S}_i$, his expected payoff is at most

$$\max_{a_i \in A_i} \sum_{a_j \in A_j} \mu^d_j(a_j) u_i(a_i, a_j),$$

which by construction is less than or equal to $\sum_{a \in A} \mu(a) u_i(a)$. Hence we have an equilibrium. Clearly, by picking a large enough m, equilibrium expected payoffs can be made to approximate any profile in $\bar{V}$ arbitrarily closely, and the deviation payoffs be made to approximate each player's minmax payoff arbitrarily closely. □

We next show that any underlying game can be extended into a game with coarsely observable strategies where *every* feasible, individually rational payoff profile belonging to a finite subset is approximated arbitrarily closely in some equilibrium.

Proposition 1 *Let $\bar{U} \subset \bar{V}$ be a finite set of feasible, individually rational payoff profiles of G. Then for any $\varepsilon > 0$ there is an extended game $G^\star$ such that for each $\bar{u} \in \bar{U}$, there is an equilibrium of $G^\star$ with payoffs $u \in B^\varepsilon(\bar{u})$.*

Sketch of proof. For each $\bar{u} \in \bar{U}$, proceed to construct an equilibrium subset of strategies as in the proof of Lemma 1, by adding the appropriate number of classes. For all classes not containing strategies associated with the current equilibrium, let the equilibrium strategies respond with the approximation of the minmax distribution. Again, for each player there will be a class containing all strategies that are not used in supporting some equilibrium payoff profile.□

For simplicity the proofs utilize a construction where the strategies played with positive probability in equilibrium are all unique in their respective classes, which may not seem very much in the spirit of coarseness of observation. It should be clear, however, that the same result may be replicated also with more coarseness, by adding more out-of-equilibrium classes.

It might also be thought disappointing that this construction cannot yield *exactly* any feasible, individually rational payoffs as the payoffs in some equilibrium. But it should be recalled that, in any case, there are no real-world randomization devices that can implement probability distributions where the probabilities are not rational numbers. Hence for practical purposes the distinction is irrelevant.

5 The Literal Interpretation of Observability

We have seen how the naïve notion of transparency and reciprocal cooperation can be rescued, at least from a logical viewpoint, by substituting coarse observability for perfect observability. It is easy to see that the construction can be interpreted as a model of delegation contracts, or games where the players can make observable commitments about how to act. But does the *literal* interpretation make sense for direct human interaction?

There are reasons to believe that it is unlikely that evolution would favor observable traits that reliably reveal something about their human carrier's decision procedures, even coarsely. The construction of the present paper relies crucially on *direct*, even if coarse, observability of strategies themselves. If, for instance, players instead reported their strategies themselves, there would be incentives to lie that would cause the construction not to work. Similarly, evolution would not favor traits that independently and truthfully signalled intentions in Prisoner's Dilemma-type interactions, since if it were ever the case that every conditional cooperator came with a "green beard," green-bearded unconditional defectors would have a selective advantage. (See, e.g., Hamilton [14] and Dawkins [7].) The literature on "cheap talk" and evolution in games (e.g., Robson [22], Wärneryd [28], and Kim and Sobel [13]) further supports this conclusion.

References

[1] Luca Anderlini. Some notes on Church's thesis and the theory of games. *Theory and Decision*, 29:19–52, 1990.

[2] Ken Binmore. Modeling rational players: Part I. *Economics and Philosophy*, 3:179–214, 1987.

[3] Ken Binmore. *Playing Fair: Game Theory and the Social Contract I.* The MIT Press, Cambridge, MA, 1994.

[4] David Canning. Rationality, computability, and Nash equilibrium. *Econometrica*, 60:877–888, 1992.

[5] Vincent P Crawford. Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions. *American Economic Review*, 93:133–149, 2003.

[6] Peter Danielson. *Artificial Morality: Virtuous Robots for Virtual Games.* Routledge, London, 1992.

[7] Richard Dawkins. *The Selfish Gene.* Oxford University Press, Oxford, 1976.

[8] Paul Ekman. *Telling Lies.* W W Norton & Company, New York, 1985.

[9] Ernst Fehr and Urs Fischbacher. Altruists with green beards. *Analyse & Kritik*, 27:73–84, 2005.

[10] Robert H Frank. If *homo economicus* could choose his own utility function, would he want one with a conscience? *American Economic Review*, 77:593–604, 1987.

[11] Robert H Frank. *Passions Within Reason: The Strategic Role of the Emotions*. W W Norton & Company, New York, 1988.

[12] David Gauthier. *Morals By Agreement*. Oxford University Press, Oxford, 1986.

[13] Yong Gwan Kim and Joel Sobel. An evolutionary approach to pre-play communication. *Econometrica*, 63:1181–1193, 1995.

[14] W D Hamilton. The genetical evolution of social behaviour I–II. *Journal of Theoretical Biology*, 7:1–52, 1964.

[15] J V Howard. Cooperation in the Prisoner's Dilemma. *Theory and Decision*, 24:203–213, 1988.

[16] Nigel Howard. *Paradoxes of Rationality: Theory of Metagames and Political Behavior*. The MIT Press, Cambridge, MA, 1971.

[17] Adam Tauman Kalai, Ehud Kalai, Ehud Lehrer, and Dov Samet. A commitment folk theorem. Working paper, 2007.

[18] David K Levine and Wolfgang Pesendorfer. The evolution of cooperation through imitation. *Games and Economic Behavior*, 58:293–315, 2007.

[19] Paola Manzini, Abdolkarim Sadrieh, and Nicolaas J Vriend. On smiles, winks, and handshakes as coordination devices. *Economic Journal*, 119:826–854, 2009.

[20] R Preston McAfee. Effective computability in economic decisions. Working paper, University of Western Ontario, 1984.

[21] Axel Ockenfels and Reinhard Selten. An experiment on the hypothesis of involuntary truth-signalling in bargaining. *Games and Economic Behavior*, 33:90–116, 2000.

[22] Arthur J Robson. Efficiency in evolutionary games: Darwin, Nash and the secret handshake. *Journal of Theoretical Biology*, 144:379–396, 1990.

[23] Ariel Rubinstein. *Modeling Bounded Rationality*. The MIT Press, Cambridge, Massachusetts, 1998.

[24] Dale O Stahl II. Evolution of smart$_n$ players. *Games and Economic Behavior*, 5:604–617, 1993.

[25] Moshe Tennenholtz. Program equilibrium. *Games and Economic Behavior*, 49:363–373, 2004.

[26] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1944.

[27] Nir Vulkan. Equilibria in automated interactions. *Games and Economic Behavior*, 35:339–348, 2001.

[28] Karl Wärneryd. Evolutionary stability in unanimity games with cheap talk. *Economics Letters*, 36:375–378, 1991.