November 1984

Report No. STAN-CS-84-1032
Also numbered: HPP-84-41

Solving the Prisoner's Dilemma

by

Michael R. Genesereth
Matthew L. Ginsberg
Jeffrey S. Rosenschein

Department of Computer Science

Stanford University
Stanford, CA 94305

Stanford Heuristic Programming Project
Report No. HPP-84-41

November 1984

Solving the Prisoner's Dilemma

Michael R. Genesereth
Matthew L. Ginsberg
Jeffrey S. Rosenschein

COMPUTER SCIENCE DEPARTMENT
Stanford University
Stanford, California 94305

Solving the Prisoner's Dilemma

Abstract

A framework is proposed for analyzing various types of rational interaction. We consider a variety of restrictions on participants' moves; each leads to a different characterization of rational behavior. Under an assumption of "common rationality," it is proven that participants will cooperate, rather than defect, in the Prisoner's Dilemma.

§1. Definitions of rationality

We will follow the usual convention of representing a game as a payoff matrix. Here is an example of the notation we will use for two-player games:

	A	B
A	3/1	2
B	2/5	0/1

The first player, to whom we will generally refer as I, selects a move labeling one of the two rows and the second, to be referred to as II, selects one of the two columns. A single number indicates an identical payoff for both players, while '3/1', for example, represents a payoff of 3 for the first player and 1 for the second. Thus if I selects move B in the above game, and II selects move A, their payoffs are 2 and 5 respectively.

Before proceeding, however, we need to define some more formal notation as well: to a game corresponds a set P of players and, for each player $i \in P$, a set M_i of possible moves for i. For $S \subset P$, we denote $P - S$ by $\bar{S}$; we will also write i instead of $\{i\}$ where no confusion is possible. Thus $\bar{\imath} = P - \{i\}$. We also write M_S for $\prod_{i \in S} M_i$.

We will denote by m_S an element of M_S; this is a collective move for the players in S. To $m_S \in M_S$ and $m_{\bar{S}} \in M_{\bar{S}}$ correspond an element $\vec{m}$ of M_P. The payoff function for a game is a function

$$p : P \times M_P \to \mathbb{R}$$

whose value at $(i, \vec{m})$ is the payoff for player i if move $\vec{m}$ is made. For a fixed m_S, we also define the ***restricted game*** given by

$$p\big|_{m_S} : P \times M_{\bar{S}} \to \mathbb{R} \qquad p\big|_{m_S}(i, m_{\bar{S}}) = p(i, \vec{m}). \tag{1}$$

Intuitively, this is the game where the players in S are assumed to make the move M_S.

We also define a secondary payoff function $pay(i, m_i)$, the set of possible payoffs to i of making move m_i:

$$pay(i, m_i) = \{p(i, \vec{m}) : m_{\bar{i}} \in allowed(i, m_i)\}. \quad (2)$$

$allowed(i, m_i)$ is the set of responses considered possible to i's move m_i. For nonempty sets $\{\alpha_i\}$ and $\{\beta_j\}$, we write $\{\alpha_i\} < \{\beta_j\}$ if $\alpha_i < \beta_j$ for all i, j.

Of interest to us will be the rational moves available either to a single player or to a group; for S $\subset$ P, we will denote by $R(p, S)$ the rational moves for the group S in the game p. The crucial characterization of rationality for an individual player is given by:

$$pay(i, d_i) < pay(i, c_i) \Rightarrow d_i \notin R(p, i). \quad (3)$$

In other words, if every possible payoff to i of making move d_i is less than every possible payoff to i of making move c_i, then d_i is irrational for i. Note that this does not imply that c_i is rational, since there may be still better moves available.

It would be possible to define $R(p, i)$ to be the maximal subset of M_i satisfying (3), but we have not done this because we are willing to allow the eventual introduction of additional constraints on the rationality of individuals if desired.

For a subset S of P, we assume:

$$R(p, \emptyset) = \{\emptyset\}$$

$$R(p, S) \subset R(p, S') \times M_{S-S'} \qquad \text{for } S' \subset S.$$

The sccorrd of these states that no rational move for a set can require irrationality on the part of a subset.

Lemma 1. $R(p, S) \subset \prod_{i \in S} R(p, i)$. ***A* move that is rational *for* a group *of* players is rational *for* each *player in the group.***

Proof. For any $i \in$ S, $R(p, S) \subset R(p, i) \times M_{S-i}$. The conclusion follows. □

Lemma *2.* $R(p, S) \subset M_S$. ***Rational moves are legal.***

Proof. $R(p, S) \subset R(p, 0) \times M_S = M_S$. ▫

Once again, it would have been possible to define $R(p, S) = \prod_{i \in S} R(p, i)$, but we have not done so because lemmas **1** and **2** are strong enough for our purposes.

The power of the characterization (3) depends on the function ***allowed in*** **(2).** Here are some possibilities:

1. **Minimal rationality:** $allowed(i, m_i) = M_{\bar{i}}$. The effect of this is for each player to assume that the others may well be moving randomly.
2. **Separate rationality:** $allowed(i, m_i) \subset R(p, \bar{i})$. Each player assumes that the others are moving rationally.
3. **Unique rationality:** $allowed(i, m_i) = allowed(i, m_i')$ and $|allowed(i, m_i)| = 1$ **for all** m_i and m_i'. Each player assumes that the others' moves are fixed n advanced. This may be combined with separate rationality.
4. **Informed rationality:** $allowed(i, m_i) = R(p|_{m_i}, \bar{i})$. Each player assumes that all others will respond rationally to whatever move he makes.

It is clear that minimal rationality is entailed by any of the other conditions, in the sense that a move which can be proven irrational under the assumption of minimal rationality will be similarly irrational under the others.

§2. Case analysis

In this section we will investigate some of the consequences of our definition of ratio-, nality. None of the results is new [6] ; our intention is simply to recast them in terms of the formalism introduced in the last section.

Consider the following game:

	A	B
A	7/1	5/2
B	4/3	0/5

In analyzing this game, the first player notices that every payoff to him if he chooses move **A** is greater than any of the payoffs available if hc chooses move B. Note that this is ***not*** true for the second player. If the second player chooses move B, his payoff will be 2 if the first chooses A; if the second player chooses move A, his payoff will be 3 if the first selects B.

Even if the second player is both totally omniscient and purely malicious, the first should still choose move A; the second player has no such analysis available. We will refer to the first player's analysis in this game as ***restricted*** case ***analysis.*** Here is the general result:

Lemma 3 (Restricted case analysis). Assuming ***minimal rationality, if for some*** c_i ***and*** d_i***, for all*** $c_{\bar{i}}$ ***and*** $d_{\bar{i}}$,

$$p(i, \vec{d}) < p(i, \vec{c}),$$

then $d_i \notin R(p, i)$.

Proof. $pay(i, d_i) = \{p(i, \vec{d}) : d_{\bar{i}} \in M_{\bar{i}}\}$, so the definition can be applied directly. □

An iterated version of this result applies for separate rationality. In the above game, for example, the second player, realizing that the first will make move A, will himself make move B.

Assuming both omniscience and maliciousness on the part of the other players may be overly pessimistic. The example we used to introduce our notation is a bit more **usual:**

	A	B
A	3/1	2
B	2/5	0/1

Provided that the second player's choice is fixed in advance, the first will be better off if he makes move A: his payoff will be 3 as opposed to 2 if the other player chooses move **A,** and 2 as opposed to **0** if B is chosen. We will refer to this decision procedure as ***case analysis*** :

Lemma 4 (Case analysis). ***Assuming unique rationality, if for some*** c_i ***and*** d_i***, for all*** $c_{\bar{i}}$ ***and*** $d_{\bar{i}}$ **with** $c_{\bar{i}} = d_{\bar{i}}$,

$$p(i, \vec{d}) < p(i, \vec{c}),$$

then $d_i \notin R(p, i)$.

Proof. Let $m_{\bar{i}}$ be the unique element of $allowed(i, c_i)$. Then if the hypothesis of the lemma is satisfied, taking $c_{\bar{i}} = d_{\bar{i}} = m_{\bar{i}}$ allows us to apply (3). □

Again, an interated result can be applied for unique separate rationality. Thus the second player chooses move B in the above example, relying on the first to choose **A, and** receiving a payoff of 2 instead of 1.

The final result of this section is similar to the previous two, but deals with the case where a group of players has to coordinate its actions in order to make the analysis effective. Consider the following game:

	A	B
A	7	4
B	5	6

Here, both players can ensure the maximum payoff of 7 by selecting move A. It is possible, in games with three or more players, for a "clique" of players to obtain maximal results by coordinating their moves in this fashion and then using case analysis to investigate the possible moves of the remaining players.

Even separate rationality is not enough to generate this result, however. The basic reason is that in order for the two players to cooperate in the game given above, it is not enough for each to know that the other is rational. They have to know that the other knows *they* arc rational, that the other knows they know they are rational, and so on [3].

In order to formalize this, we will need to extend the notion of payoff to sets of players. Let $p(S, \vec{m})$ to be the n-tuplc whose components are the $p(i, \vec{m})$'s for the i's in S, and write $p(S, \vec{d}) < p(S, \vec{c})$ if $p(i, \vec{d}) \leq p(i, \vec{c})$ for each $i \in$ S, with the inequality being strict in at least one case. (In other words, a move is an improvement for a group of players if no player in the group loses and at least one player benefits.) We now define

$$pay(S, m_S) = \{p(S, \vec{m}) : m_{\bar{S}} \in allowed(S, m_S)\},$$

this being the set of possible payoffs to a group S of players if they make a collective move m_S. The new function *allowed* might be defined as, for example; *allowed*$(S, m_S) = R(p|_{m_S}, \bar{S})$.

Theorem 5 (Coordination). Assuming informed rationality, *if there exists* a c_S such that *for all* $d_S \neq c_S$, *pay*$(S, d_S) <$ *pay*(S, c_S), *then* $R(p, S) \subset \{c_S\}$.

Proof. By lemma 1, it suffices to show that, for each $d_S \neq c_S$, there exists an $i \in S$ such that $d_i \notin R(p, i)$.

The proof proceeds by induction on the size of S. For S of size **1,** the theorem is an immediate consequence of the definition of rationality (3).

Suppose, then, that the theorem is true for all S of cardinality less than n, and that the hypotheses of the theorem are satisfied for some S of size n. With c_S and d_S as in the statement of the theorem, so that $d_i \neq c_i$ for some fixed $i \in S$, it follows from the inductive hypothesis that $R(p|_{c_i}, S - i) \subset \{c_{S-i}\}$, so that $R(p|_{c_i}, \bar{i}) \subset \{c_{S-i}\} \times R(p|_{c_S}, \bar{S})$, and $pay(i, c_i) \subset pay(i, c_S)$. We also have that $pay(i, d_i) \subset \bigcup_{d_{S-i}} pay(i, d_S)$ and therefore, since $pay(i, d_S) < pay(i, c_S)$ for any d_S, that $pay(i, d_i) < pay(i, c_S)$. Since $pay(i, c_i) \subset pay(i, c_S)$, we get $pay(i, d_i) < pay(i, c_i)$ and $d_i \notin R(p, i)$. □

Note that this theorem is much weaker than the earlier two, in the sense that it cannot declare a joint move to be irrational if there is another joint move that is better. To see why, consider the prisoner's dilemma [6,7]:

	C	D
C	3	0/5
D	5/0	1

Case analysis forces each player to choose move D (defection), although this is worse for both than the move (C,C) (joint cooperation). What the theorem does say is that if there is a single move that is preferable to *all* the others, it will be selected.

§3. Cooperation. and the prisoner's dilemma

The definition of informed rationality is our first attempt to understand the consideration one player may give to the analyses of the others. Informed rationality is, however, a bit too strong—it may not be to one player's advantage to make a move if his opponents know he will make it (witness the prisoner's dilemma, where informed rationality forces each player to defect).

What is required is some way to encode the concept of, "I would be willing to cooperate if you didn't make me regret it." In order to do this, we modify our notion of move to include a move which is provisional in the sense that the player considering it will retract it if it works out badly.

Let i be a player, and suppose that $S \subset \bar{\imath}$. For $r_S \subset m_S \times M_{\bar{S}}$, we will define a set $R(p, i, m_S, r_S)$, the set of rational moves for player i, under the assumption that group S is willing to make move m_S ***provided that the final outcome of the game is in*** r_S. The set $R(p, i, m_S, r_S)$ consists not only of moves m_i, but of pairs (m_i, r_i) where r_i is an additional restriction put by i on the eventual outcome of the game.

The expressions *(2)* and (3) now become:

$$pay(i, m_i, r_i, m_S, r_S) = \{p(i, \vec{m}) : m_{\bar{\imath}} \in allowed(i, m_i, r_i, m_S, r_S)\} \tag{4}$$

$$pay(i, d_i, r_i, m_S, r_S) < pay(i, c_i, r_i', m_S, r_S) \Rightarrow (d_i, r_i) \notin R(p, i, m_S, r_S). \tag{5}$$

The second of these says that a (move,deal) pair is irrational if ***either*** the move or the deal can be improved.

If S if empty, WC write $pay(i, m_i, r_i)$ instead of $pay(i, m_i, r_i, m_\phi, M_P)$. With r_i varying, the various sets $pay(i, m_i, r_i)$ are partially ordered under $<$; we denote the union of all of those that arc ***not*** dominated by some other $pay(i, m_i, r_i')$ by $pay(i, m_i)$:

$$r \equiv \{r_i : \nexists\, r_i' \,.\, pay(i, m_i, r_i) < pay(i, m_i, r_i')\}$$

$$pay(i, m_i) = \bigcup_{r_i \in r} pay(i, m_i, r_i).$$

Intuitively, $pay(i, m_i)$ is the set of possible payoffs to i if he makes both move m_i and any of.the best deals associated with it. The point of this definition is that rationality itself still corresponds to

$$pay(i, d_i) < pay(i, c_i) \Rightarrow d_i \notin R(p, i). \tag{6}$$

In other words, if i's "best deal" with move d_i is worse than his best deal with c_i, then d_i is irrational.

Finally, we need to define $allowed(i, m_i, r_i, m_S, r_S)$. There are two possiblities, depending upon whether or not i has broken the tacit "understanding" he might have had with the players in S. ***allowed*** is therefore defined as:

$$allowed(i, m_i, r_i, m_S, r_S) = \begin{cases} R(p, \bar{\imath} - S, m_i \times m_S, r_i \cap r_S) & \text{if } r_i \cap r_S \neq \emptyset; \\ R(p, \bar{\imath}, m_i, r_i) & \text{otherwise.} \end{cases} \quad (7)$$

The first case corresponds to i keeping the deal; the second to his breaking it. We will refer to this definition **as common rationality.**

Theorem 5a. ***The coordination theorem holds for common rationality.***

Proof. The proof is the same as that of theorem 5. The inductive step uses the fact that if $R(p, S - i) = \{c_{S-i}\}$, then $pay(i, c_i, r_i) = pay(S, c_S, r_i)$. □

The added power of common rationality can be seen in an example; we will use the prisoner's dilemma. Here is the payoff matrix again:

	C	D
C	3	0/5
D	5/0	1

Evaluating the various payoffs, it is fairly clear that $pay(\mathrm{I}, C, \{CC, CD\}) = 0$ **and** $pay(\mathrm{I}, D, \{DC, DD\}) = 1$, since II will surely defect if I is willing for him to do so. We also have $pay(\mathrm{I}, C, \{CD\}) = 0$ and $pay(\mathrm{I}, D, \{DD\}) = 1$, since II has still not been "asked" **to** cooperate. The other two cases are more interesting.

For $pay(\mathrm{I}, V, \{IX\})$, we need to evaluate $R(p, \mathrm{II}, D, \{DC\})$. Turning our attention to II, we obviously have

$$pay(\mathrm{II}, C, r_{\mathrm{II}}, D, \{DC\}) = 0$$

for all r_{II}, since II has gone along with I's request to cooperate while I **defects.** But

$$pay(\mathrm{II}, D, \{CD, DD\}, D, \{DC\}) > \{0\},$$

since II guarantees himself a payoff of at least 1 by defecting. Thus $R(p, \mathrm{II}, D, \{DC\}) = \{D\}$; in other words, II will respond to any deal involving I's defection by defecting himslf. It follows from this that $pay(\mathrm{I}, D, r_{\mathrm{I}}) = \{1\}$ for all r_{I}.

It remains to evaluate $pay(\mathrm{I}, C, \{CC\})$. Here I is willing to cooperate if II does. To evaluate $R(p, \mathrm{II}, C, \{CC\})$, we have

$$pay(\mathrm{II}, D, r_{\mathrm{II}}, C, \{CC\}) = \{1\}$$

for all r_{II}, since $pay(\mathrm{II}, D, r_{\mathrm{II}}, C, \{CC\}) = pay(\mathrm{II}, D, r_{\mathrm{II}})$ (II breaks the deal), and we evaluated this (albeit for I) in the last paragraph. We also have

$$pay(\mathrm{II}, C, r_{\mathrm{II}}, C, \{CC\}) = \{3\}.$$

It follows that $R(p, \mathrm{II}, C, \{CC\}) = \{C\}$, so that $pay(\mathrm{I}, C, \{CC\}) = 3$.

We therefore have $pay(\mathrm{I}, D) = \{1\}$ and $pay(\mathrm{I}, C) = \{3\}$, so that $D \notin R(p, \mathrm{I})$. $D \notin R(p, \mathrm{II})$ is of course similar and the players therefore cooperate.

Theorem 6 (Cooperation). Assume common rationality, and let c_S be fixed. For any alternative d_S to c_S, let T be the set of all $i \in S$ with $d_i \neq c_i$. Assume that for all such d_S and T, either

$$S = T \text{ and } pay(S, d_S) < pay(S, c_S),$$

or there exists some $i \in S - T$ and $m_i \neq c_i$ such that

$$pay(i, m_i) > pay(i, d_{i \cup T}).$$

Then $R(p, S) \subset \{c_S\}$.

Before proving this result, let us consider its content. What it says is that any set of players will cooperate if the attempted defection of any subset forces the defection of at least one additional player outside the subset, and if the defection of the entire set damages every member of it.

A stronger result would be that

$$pay(S, d_S) < pay(S, c_S) \Rightarrow d_S \notin R(p, S),$$

but this cannot be obtained without some way for the players to communicate with each other. If we imagine the prisoner's dilemma to be changed so that the payoff for *both* players is 5 if either one cooperates while the other defects, there is no way to ensure this payoff without some form of communication on the part of the players.

This should not lead us to underestimate the power of theorem 6; it includes both the coordination theorem and the prisoner's dilemma as special cases. Turning to the proof:

Proof. We will show that for any $d_S \neq c_S$ and r_S,

$$pay(S, d_S, r_S) < pay(S, c_S). \tag{8}$$

After this, an inductive proof such as that used in the coordination theorem can be applied.

Suppose that (8) does not hold for all d_S. Then we can consider the collection of counterexamples to it for which $R(p, \bar{T}, d_T, r_T) \subset r_T$; in other words, the counterexamples for which the "deal" offered by the defectors is accepted by the other players. Let d_S be an element of this collection which is maximal in the sense that $|T|$ is as small as possible.

Clearly $|T| \neq 0$, since this would involve defection of all of S, and this is guaranteed not to be an improvement by the hypothcscs of the theorcm. But if $|T| \neq 0$, we know that there is an $i \in S - T$ and m_i with $pay(i, m_i) > pay(i, d_{i \cup T})$, *so* that i will defect as well, contradicting the assumption that d_S was maximal. □

§4. Conclusion

WC have prcscntcd a unificd framework for considcring various typcs of interactions that occur without communication. Using <assumptions about what typcs of movcs other agents will make, a participant is able to reason about what constitutes rational behavior on its own part. Several of the characterizations of rationality havc parallels in existing game theory litcrature, and lead to familiar rcsults such as case analysis and iterated case analysis.

The power of our approach is seen in the ("common rationality" assumption, which forces both participants to cooperate in the prisoner's dilemma (the formal result parallels previous informal arguments [2, 4]). Game theory and philosophy [8] have generally defined rationality in such a way as to require mutual defection in this game, and have looked to changes in the interaction (such as its iteration [1] or explicit agreements) to motivate cooperation.

While our approach does not require communication (or meta-game analysis [5]), there is a sort of deal-making that is implicit in our formalism and its results. The agents reason about the best deal they could extract from the other player, and come to a common conclusion as to what move is warranted by this deal. Reasoning usurps the role of communication. There is also no need for a "binding agreement" to be reached among the participants, since each is constrained by its own definition of rationality and these common constraints are appreciated by all.

. Acknowledgement

The authors have chosen to list themselves in alphabetical order. This research has been supported by the Office of Naval Research under grant number N00014-81-K-0004 and by DARPA under NAVELEX grant number N00039-83-C-0136.

References

[1] Axelrod, R., ***The Evolution of Cooperation,*** Basic Books, Inc., New York (1984).

[2] Davis, L., Prisoners, paradox and rationality, ***American Philosophical Quarterly*** **14** (1977).

[3] Halpern, J., and Moses, Y., Knowledge and common knowledge in a distributed environment, ***Proceedings of the Third Annual ACM Conference on Principles of Distributed Computing,*** Vancouver, British Columbia, Canada (1984).

[4] Hofstadter, D.R., Metamagical themas—the calculus of cooperation is tested through a lottery, ***Scientific American*** 248 (1983) 14-28.

[5] Howard, N., The theory of meta-games, ***Yearbook of the Society for General Systems Research*** **XI** (1966) 167-186.

[6] Luce, R.D., and Raiffa, H., ***Games. and Decisions, Introduction and Critical Survey,*** John Wiley and Sons, New York (1957).

[7] Rapoport, A., ed., ***Game Theory as a Theory of Conflict Resolution,*** D. Reidel Publishing Company, Boston (1974).

[8] Sowden, L., That there is a dilemma in the prisoners' dilemma, ***Synthese*** **55** (1983) **347-352.**