A Decision-Theoretic Approach to Coordinating Multiagent Interactions

Piotr J. Gmytrasiewicz*, Edmund H. Durfee*, and David K. Wehe
*Department of Nuclear Engineering
tD epartment of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, Michigan 48109

Abstract

We describe a decision-theoretic method that an autonomous agent can use to model multiagent situations and behave rationally based on its model. Our approach, which we call the Recursive Modeling Method, explicitly accounts for the recursive nature of multiagent reasoning. Our method lets an agent recursively model another agent's decisions based on probabilistic views of how that agent perceives the multiagent situation, which in turn are derived from hypothesizing how that other agent perceives the initial agent's possible decisions, and so on. Further, we show how the possibility of multiple interactions can affect the decisions of agents, allowing cooperative behavior to emerge as a rational choice of selfish agents that otherwise might behave uncooperatively.

Introduction

A central issue in distributed artificial intelligence (DAI) is how to get autonomous intelligent agents, each of whom has its own goals and preferences, to model each other and coordinate their activities for their mutual benefit. This paper describes a recursive method that agents can use to model each other in order to estimate expected utilities more completely in multiagent situations, and thus to make rational and coordinated decisions. Our method works by letting agents explicitly reason about how the collective actions of agents can affect the utilities of individual actions. Thus, to choose an action that maximizes its individual utility, an agent should predict the actions of others. The fact that other agents are likely to take the same approach gives rise to the recursive nesting of models.

Our Recursive Modeling Method (RMM) represents this recursion explicitly to allow an agent to arrive, within the bounds of its processing, on the most rational decision in the multiagent environment. RMM considers all of the available information an agent might have about others and summarizes the possible uncertainties as a set of probability distributions. This representation can reflect the uncertainty as to the other agents' intentions, abilities, long-term goals, and sensing capabilities. Furthermore, on a deeper level of the recursion, the agents may have information on how other agents are likely to view them, how they themselves think they are viewed, and so on.

Our work, thus, extends other work [Rosenschein and Breese, 1989] that uses a game theoretic approach to coordinating interactions without communication. That work unrealistically assumes that agents have full information about each other's choices, preferences, and perceptions. Other research efforts in DAI use similar formalisms to our work, but avoid the recursive issues that we are studying by allowing agents to communicate about their beliefs, goals, and preferences, in order to make explicit deals [Werner, 1989; Zlotkin and Rosenschein, 1989; Zlotkin and Rosenschein, 1990].

Research in cooperation indicates that agents can converge on cooperative strategies during repeated interactions without ever explicitly communicating [Axelrod, 1984]. The most well-known example is the Prisoner's Dilemma game, where a rational "one-shot" strategy is to defect, but where a "Tit-for-Tat" strategy is best for repeated interactions. Following the methodology of metagames [Howard, 1966; Reagade, 1987], the goal of our work is to develop a formal, algorithmic model that captures how cooperative strategies can be derived by self-interested, rational agents.

In the remainder of this paper, we begin by outlining the basic concept of a payoff matrix from decision and game theories, and then we define the RMM and illustrate it with an example. Subsequently, we show how the possibility of multiple interactions changes the character of games, and illustrate this using the Prisoner's Dilemma problem. We revisit the earlier example and apply the multiple interactions concept within RMM. We conclude by summarizing our results and current research directions.

Establishing Payoffs

A decision-theoretic approach to multiagent interaction requires that an agent view its encounters with other agents in terms of possible joint actions and their utilities, usually assembled in the form of a payoff matrix. We have developed a system, called the Rational Rea-

°This research was supported, in part, by the Department of Energy under contract DG-FG-86NE37969, and by the NSF under Coordination Theory and Collaboration Technology Initiative grant IRI-9015423.

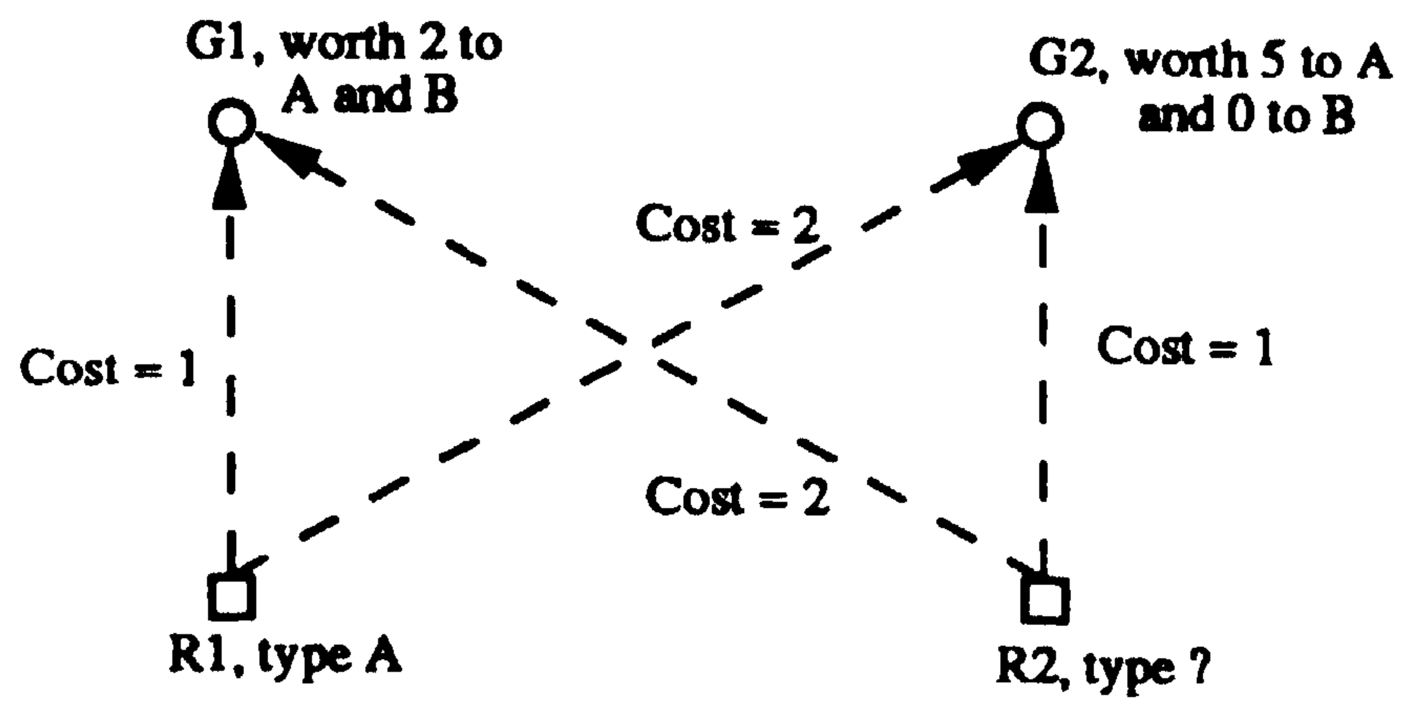

Figure 1: Example Scenario of Interacting Agents

soning System (RRS) [Gmytrasiewicz *et* al., 1991a] that determines plans' utilities [Jackobs and Kiefer, 1973] to automatically generate the information for a payoff mar trix. For brevity, we will not describe the details of RRS, beyond saying that it combines decision-theoretic techniques with hierarchical planning to generate alternative decisions (plans of action), and uses time-dependent calculations of utility to generate the expected payoffs. These calculations involve formal notions of agents' preferences and the ways specific tasks in the environment impact these preferences.

The Recursive Modeling Method

The Recursive Modeling Method (RMM) seeks to include all the information an agent might have about the other agents and can be viewed essentially as an extension of case analysis [Luce and Raiffa, 1957] to situations in which other players' payoffs and alternatives are uncertain. This technique attempts to put oneself in the shoes of the other player and to guess what he will prefer to do. Our approach thus follows, and extends to deeper levels of recursion, the main idea of a hypergame method [Bennett and Huxham, 1982; Vane and Lehner, 1990]. Our principal contribution is a complete and rigorous formalism that, unlike similar work [Cohen and Levesque, 1990; Rosenschein, 1988], directly relates the recursive levels of the agents' knowledge to the utilities of their actions and thus to the intentions of rational agents [Dennett, 1986].

Example Multigent Interaction

To put our description in concrete terms, consider the example scenario of two interacting agents (Figure 1). The environment has agents of type A and B. Agents of type A can perceive all of the goals in the environment, while B agents can only see goals of type GI. Moreover, A agents are aware of the two types of agents, A and B, while B agents are aware only of their type. Further, A agents can perform all of the types of goals, while B agents are equipped only for GI goals. The utility of a GI goal is equal to 2 for both types of agents, while the utility of a G2 goal is 5 for A agents and 0 for B agents. The cost of an attempt by either agent to achieve the farther goal is 2, and the cost of attempting the closer goal is 1. For simplicity, we assume that the agents can achieve only one goal.

Let us focus our attention on RI, which is of type A. We see R1 as having three options: pursue G1, pursue G2, or do something (including nothing) else (G1, G2 and S, for short). Using the above information, R1's payoffs are computed as the difference between the total worth of all the performed goals and the cost of achieving its own goal. For example, if R1 pursues G2 and R2 pursues G1, then the payoff for R1 is the total worth of the achieved goals minus its own cost: (2+5) - 2 = 5. These payoffs are assembled in the following payoff matrix:

		R2		
		G1	G2	S
	G1	1	6	1
R1	G2	5	3	3
	S	2	5	0

This matrix represents R1's utilities for its possible decisions (G1, G2, or S) depending on R2's decisions (G1, G2, or S). These utilities are described as the payoffs of R1 given the joint moves of both agents.

R1 may reasonably assume that R2 is trying to maximize its own payoff (see also [Dennett, 1986]), but the difficulty in predicting R2's actions is that R1 is uncertain as to R2's type. R1 thus does not know whether R2 will see, value, or pursue G2. In RMM, these uncertainties are represented in terms of probability distributions. Furthermore, R2's actions are likely to depend on how R2 views R1. R1 will thus form alternative models of itself corresponding to how it thinks R2 might perceive it.

We will now proceed to introduce the general form of RMM that captures the intuitions mentioned above. We will come back to the example scenario depicted in Figure 1 and solve it subsequently.

The General Form of RMM

To generalize over an arbitrary number of agents, let us assume that R1 is dealing with (N-1) other agents, R2, ..., RN. The utility of the m-th element of R1's set of alternative actions can be evaluated as:

$$u_m^{R1} = \sum_k \cdots \sum_l \{p_{R2-k}^{R1} \cdots p_{RN-l}^{R1} u_{m,k,\ldots,l}^{R1}\} \qquad (1)$$

where p_{Ri-k}^{R1} represents the probability R1 assigns to Ri's choosing to act on the k-th element of Ri's set of options, which we will refer to as an intentional probability. $u_{m,k,\ldots,l}^{R1}$ is R1's payoff (utility) as an element of the N-dimensional game matrix.

What makes the situation recursive is the fact that R1 may attempt to determine the intentional probabilities p_{Ri-k}^{R1} by guessing how the game looks from Ri's point of view. R1 models agent Ri using probability distributions p_{Pi}^{R1}, p_{Ai}^{R1}, and p_{Wi}^{R1}, which we will call modeling probabilities. p_{Pi}^{R1} summarizes R1's knowledge about Ri's preferences (the goals it will value). p_{Ai}^{R1} summarizes R1's knowledge about Ri's abilities, given its preferences. p_{Wi}^{R1} summarizes R1's knowledge about how Ri sees the world (content of Ri's world model), given its abilities. In every case of Ri having various preferences, abilities and world models, R1 can assume that Ri is rational (assumption of intentionality, see also [Dennett, 1986; Rosenschein, 1988]) and consider the probability that the k-th element of Ri's set of options is of the highest utility to Ri. R1 can then use the modeling probabilities to compute p^{R1}_{Ri-k} as the following probabilistic mixture:

$$p^{R1}_{Ri-k} = \sum_{Pi} \sum_{Ai} \sum_{Wi} \{p^{R1}_{Pi} p^{R1}_{Ai} p^{R1}_{Wi} \times Prob(Max_{k'}(u^{R1,Ri}_{k'}) = u^{R1,Ri}_{k})\} \quad (2)$$

where $u^{R1,Ri}_{k'}$ is the utility R1 estimates that Ri will assign to its option k', and is computed as

$$u^{R1,Ri}_{k'} = \sum_{r} \cdots \sum_{s} \{p^{R1,Ri}_{R1-r} \cdots p^{R1,Ri}_{RN-s} u^{R1,Ri}_{k',r,\ldots,s}\} \quad (3)$$

The $u^{R1,Ri}_{k',r,\ldots,s}$ is how R1 sees Ri's payoffs in the N-dimensional game matrix. The probabilities R1 thinks Ri assigns to agent Rn acting on its o-th option $p^{R1,Ri}_{Rn-o}$, can in turn be expressed in terms of $p^{R1,Ri,Rn}_{Rv-w}$ and $u^{R1,Ri,Rn}_{o',w,\ldots}$ and so on.

Solving the Example Interaction

Given the payoff matrix computed before, R1 can compute the expected utilities of each of its alternatives based on equation (1) as follows:

$$\begin{aligned} u^{R1}_{G1} &= p^{R1}_{R2-G1} + 6p^{R1}_{R2-G2} + p^{R1}_{R2-S} \\ u^{R1}_{G2} &= 5p^{R1}_{R2-G1} + 3p^{R1}_{R2-G2} + 3p^{R1}_{R2-S} \\ u^{R1}_{S} &= 2p^{R1}_{R2-G1} + 5p^{R1}_{R2-G2} \end{aligned} \quad (4)$$

where p^{R1}_{R2-k} denotes the probability that R1 assigns to R2's intending to act on its k-th option. R1 can stop the computation without recursion, in which case it assumes an equiprobable intentional probability distribution $p^{R1}_{R2} = (p^{R1}_{R2-G1}, p^{R1}_{R2-G2}, p^{R1}_{R2-S}) = (1/3, 1/3, 1/3)$ (following the entropy maximization principle [Neapolitan, 1990]). Based on this distribution with zero levels of recursion, R1 would choose G2, which we can summarize as:

Decision 0: $R1^0 \rightarrow G2$.

Alternatively, R1 could determine the values of p^{R1}_{R2-k} using equations (2) and (3). Plunging deeper into the recursion, R1 has to look at the game from R2's point of view. R2 can be either type A or B, and the corresponding payoff matrices are:

		R1		
		G1	G2	S
R2(A)	G1	0	5	0
	G2	6	4	4
	S	2	5	0

		R1	
		G1	S
R2(B)	G1	0	0
	S	2	0

The utilities of R2's options, if it is of type A, are:

$$\begin{aligned} u^{R1,R2}_{G1} &= 5p^{R1,R2}_{R1-G2} \\ u^{R1,R2}_{G2} &= 6p^{R1,R2}_{R1-G1} + 4p^{R1,R2}_{R1-G2} + 4p^{R1,R2}_{R1-S} \\ u^{R1,R2}_{S} &= 2p^{R1,R2}_{R1-G1} + 5p^{R1,R2}_{R1-G2} \end{aligned} \quad (5)$$

If R2 is of type B, the utilities are:

$$\begin{aligned} u^{R1,R2}_{G1} &= 0 \\ u^{R1,R2}_{S} &= 2p^{R1,R2}_{R1-G1} \end{aligned} \quad (6)$$

where $p^{R1,R2}_{R1-n}$ is the probability that R1 thinks R2 is assigning to R1's acting on the n-th of its alternatives. If R1's analysis of the situation were to stop at this point, R1 would assume an equiprobability distribution over its own options, as seen by R2, and conclude that R2 will pursue G2 if it is A, and S if it is B. Let us assume that R1 does not know whether R2 is A or B and thus assigns the probability $p^{R1}_{R2(A)} = 0.5$ to the possibility that R2 is of type A. Here, the probability $p^{R1}_{R2(A)}$ encapsulates all of the modeling probabilities regarding R2's preferences, abilities and world model. R1 can now use equation (2) to estimate the intentional probability distribution over R2's options G1, G2 and S, as $p^{R1}_{R2} = 0.5(0,1,0) + 0.5(0,0,1) = (0, .5, .5)$. In this case, R1 would estimate the expected utility of its own choices from equation (4) as:

$$\begin{aligned} u^{R1}_{G1} &= 3.5, \\ u^{R1}_{G2} &= 3 \\ u^{R1}_{S} &= 2.5. \end{aligned} \quad (7)$$

We see that R1 will opt for G1 at this first stage of the recursion. Let us summarize this as:

Decision 1: $R1^1 \rightarrow G1$.

We now go on to the next step of the recursion and consider how R1 would see itself through R2's eyes, and how that would influence R1's decision.

If R2 is A, it will form two possible views of the game from R1's perspective, corresponding to R1 being A or B, respectively:

		R2		
		G1	G2	S
R1(A)	G1	1	6	1
	G2	5	3	3
	S	2	5	0

		R2	
		G1	S
R1(B)	G1	1	1
	S	2	0

At this point, using an equiprobability distribution over R2's options, R1 would think that R2 would conclude that R1 will pursue G2 if it is A, and equally likely G1 or S if it is B. Assuming that R2 would treat R1's being A or B as equiprobable ($p^{R1,R2}_{R1(A)} = 0.5$), the resulting intentional distribution over R1's options G1, G2 and S, would be $p^{R1,R2}_{R1} = 0.5(0,1,0) + 0.5(.5,0,.5) = (.25, .5, .25)$. The expected payoffs of R2, as seen by R1, can then be computed from equation (5) as:

$$\begin{aligned} u^{R1,R2}_{G1} &= 5p^{R1,R2}_{R1-G2} = 2.5 \\ u^{R1,R2}_{G2} &= 6p^{R1,R2}_{R1-G1} + 4p^{R1,R2}_{R1-G2} + 4p^{R1,R2}_{R1-S} = 4.5 \\ u^{R1,R2}_{S} &= 2p^{R1,R2}_{R1-G1} + 5p^{R1,R2}_{R1-G2} = 3 \end{aligned} \quad (8)$$

If R2 is B, its view of R1 would consist of a single view, since it could only see R1 as type B. R1's payoff matrix in this case is:

		R2	
		G1	S
R1(B)	G1	1	1
	S	2	0

In this case, using an equiprobability distribution over R2's choices does not give R2 a clue about R1's choice. It will use the equiprobability distribution over R1's options in equation (6) and compute as follows:

$$\begin{aligned} u_{G1}^{R1,R2} &= 0 \\ u_S^{R1,R2} = 2p_{R1-G1}^{R1,R2} &= 1 \end{aligned} \tag{9}$$

Thus, R1 will find that R2 will choose G2, if it is A, and S, if it is B, to give $p_{R2}^{R1} = (0, .5, .5)$. Using equation (4) again, the expected utilities of R1 are:

$$\begin{aligned} u_{G1}^{R1} &= 3.5 \\ u_{G2}^{R1} &= 3 \\ u_S^{R1} &= 2.5 \end{aligned} \tag{10}$$

Thus, at the second stage of the recursion, R1's conclusion as to its best option is:

Decision 2: $R1^2 \rightarrow G1$.

Let us have a look at the conclusions R1 would reach at the stages of the recursion considered so far:

Decision 0: $R1^0 \rightarrow G2$.

Decision 1: $R1^1 \rightarrow G1$.

Decision 2: $R1^2 \rightarrow G1$.

Thus, if R1 were to treat R2 as a complete unknown, as in the zeroth stage, it would decide for G2. Going deeper and considering how R2 may view the situation, R1's best option changes to G1. Going even deeper and seeing himself being analyzed by R2, R1's best option remains G1, despite R1 being aware that, if R2 were A, it would incorrectly think that, if R1 were A, R1 would pursue G2! This has a stabilizing effect on the deeper levels of recursion, and G1 remains R1's best option (assuming that R1 treats the modeling probabilities of the agents' types on deeper levels of the recursion as equiprobable as well).

The payoff matrices of both players considered in the above example can be depicted as in Figure 2. The recursion can, in principle, be continued indefinitely, but usually, for a given problem, going below certain level does not contribute anything new to the analysis. This level is the one at which the game starts looking the same from the perspective of the either player. Consider probabilities $p_{R2(A)}^{R1}$ and $p_{R2(A)}^{R1,R2,R1}$ in Figure 2. If they are equal (R1 sees R2 in the same way as R1 thinks R2 thinks it is being seen by R1), no new information would be contributed beyond the 4-th level of the recursion (assuming that $p_{R1(A)}^{R1,R2}$ and $p_{R1(A)}^{R1,R2,R1,R2}$ are in turn equal, and so on).

Non-Converging Example Interaction

The fact that the stages of the recursion start looking very similar from stage to stage does not mean that a unique probability distribution over the other agents' options can be reached in every case. A variation of the example interaction illustrates this. Imagine that R1, being A and quite sure that it looks like A, estimates the probability of R2's (if it is A) correctly identifying R1 as A as being high, say $p_{R1(A)}^{R1,R2} = 0.9$. The resulting intentional distribution over R1's options G1, G2 and S, as seen by R2 if it is A, would be $p_{R1}^{R1,R2} = (.05, .9, .05)$. The calculations in equations (8) and (9) then result in S being the best option of R2, if it is of either type ($p_{R2}^{R1} = (0, 0, 1)$). This, in turn, would result in the following expected utilities for R1:

$$\begin{aligned} u_{G1}^{R1} &= 1 \\ u_{G2}^{R1} &= 3 \\ u_S^{R1} &= 0 \end{aligned}$$

R1, in this case, would thus arrive at

Decision 2: $R1^2 \rightarrow G2$.

Were R1 to consider the next level, it would come up with the probability distribution over R2's options $p_{R2}^{R1} = (0, .5, .5)$ and

Decision 3: $R1^3 \rightarrow G1$.

This instability continues at the deeper levels of recursion with RMM unable to decide whether R2 makes its choices according to probability distribution (0, .5, .5) or (0, 0, 1). Our approach to resolving this situation is to apply the principle of indifference to the probability distributions over R2's choices. In the above case, the distributions (0, .5, .5) and (0, 0, 1) are merged with 0.5 weighting factor each, resulting in $p_{R2}^{R1} = (0, .25, .75)$. The calculation of the expected utilities of R1's options using this merged distribution over R2's options results in G2 being R1's best option, which we regard as a solution in this case. Let us note, that R1's decision to pursue G2 can also be called an uncooperative option, in a sense that it treats R2 as if it were not there, and that R2 would prefer that R1 choose G1, if it were either type A or B. We will come back to this point in the next section.

Our preliminary investigation of the general convergence properties of RMM suggests that, if the method does not yield a unique intentional probability distribution over the other agents' options (as in the previous example), it will converge on a *finite* set of alternative distributions, which can then be combined. The finiteness of this set is essentially due to the finite knowledge base of the agents, and, in particular, to the fact that they cannot have explicitly given knowledge about what other agents think about others thinking about others thinking ..., to infinite levels. It is, in fact, likely that the recursive hierarchy of payoff matrices will start becoming uniform down from level 4 or 5. That, in turn, permits agents to determine the finite set of intentional probability distributions of the others at relatively high levels of the hierarchy.

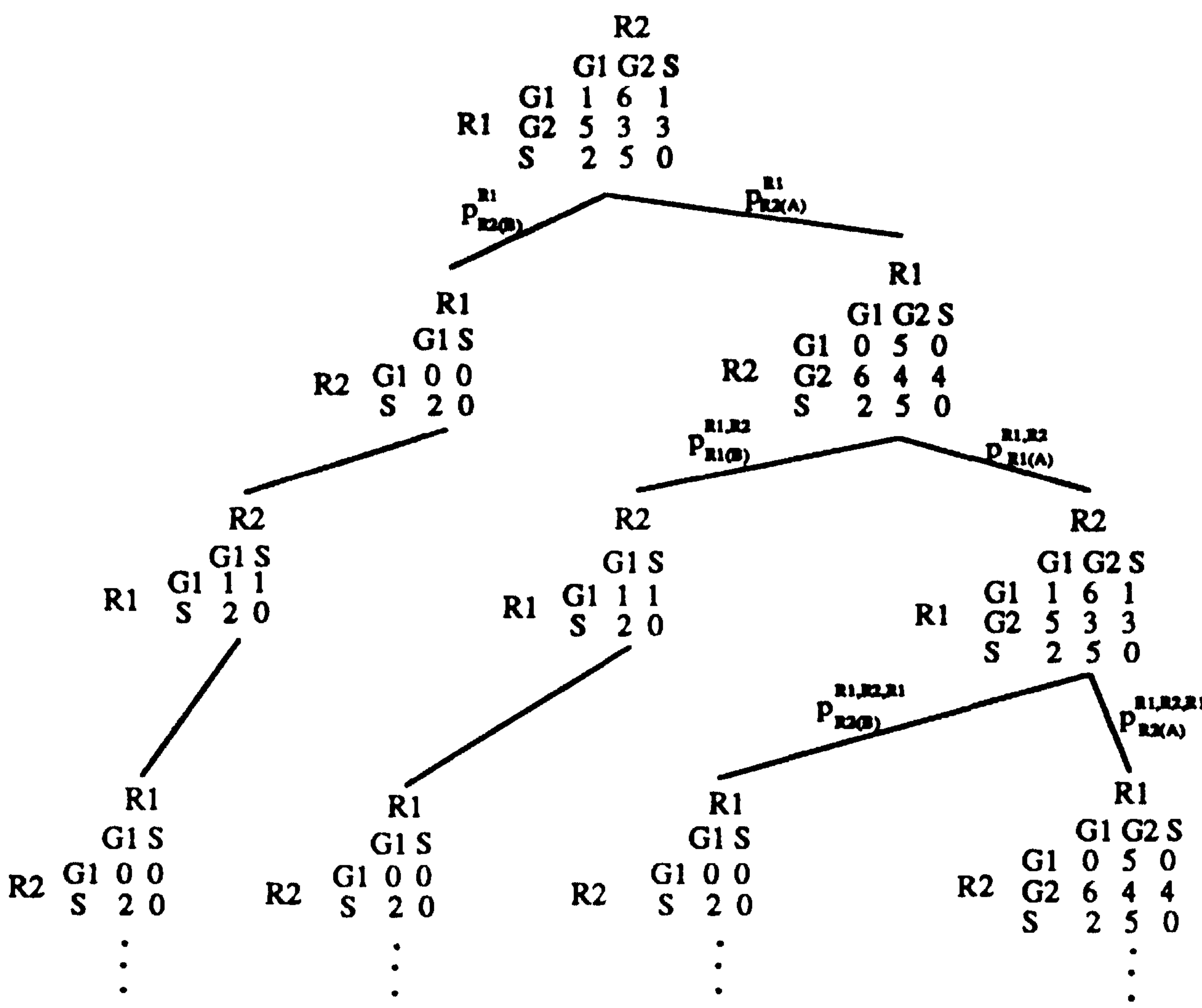

Figure 2: Example of the Recursive Hierarchy

The assumption of intentionality, as formalized in equation (2), can give rise to various probability distributions over agents' options, depending on how strongly they are assumed to be coupled to the probabilities of others' options. In the examples above, we followed the most straightforward, and possibly the most risky, path of assigning an equal, nonzero probability to the options with the highest expected payoff, and zero to all of the rest. Another cautious extreme would be to assign a probability of zero to all of the dominated options, and use the equiprobability assumption for the rest. The influence of these and other ways of interpreting equation (2) are under study.

Multiple Interactions

As shown in the previous section, selfish, rational agents employing RMM may fail to exhibit cooperative behaviors in one-time encounters. In this section, we present a methodology based on metagame analysis [Howard, 1966] that, when integrated with the RMM, makes agents more cooperative when they might interact repeatedly.

To introduce our methodology, we temporarily abandon our scenario of Figure 1. Instead, we use the Prisoner's Dilemma game, as a well known, and very simple example of a game in which repeated interactions lead to cooperation. We then revisit the example of Figure 1.

Repeated Prisoner's Dilemma

Most realistic problem domains involve a finite number of agents that periodically interact, so agents that have interacted in the past could encounter each other repeatedly. A strategy that is rational for one interaction might be counterproductive in repetitive situations where agents can consider their prior experiences in deciding on their actions.

The simplest illustration is the Prisoner's Dilemma (PD), with this payoff matrix:

		II	
		c	d
I	C	3\3	0\5
	D	5\0	1\1

If player I can be sure that he will never play with player II again, he would note that the payoffs of his D move dominate the payoffs of C. That is, no matter what player II does, player I is better off with D. Since the game is symmetric, both players choose to defect and a joint move D/d, with a payoff of 1 to each player, results. The paradox of PD is that, if both players were irrational, they could cooperate and each receive a payoff of 3. Thus, in a one-time interaction, a paradoxical, noncooperative solution results.

It has been previously demonstrated that, for repeated Prisoner's Dilemma, the one-time strategy is a poor choice [Axelrod, 1984; Smith, 1984]. In a population of alternative strategies that compete with each other over multiple generations, Axelrod experimentally discovered that a "Tit-for-Tat" strategy, in which a player is predisposed to cooperate but will defect against a player who defected in their previous encounter, is the "fittest" for PD.

We derive this result more rigorously using metagame theory [Howard, 1966; Reagade, 1987]. Let us define a strategy, or a met amove, of player I as an initial move to be made complemented by a mapping from the set

		S1	S2	S3	S4	S5	S6	S7	S8
					II				
	S1	4\4	4\4	16\1	16\1	8\3	8\3	20\0	20\0
	S2	4\4	4\4	12\7	9\9	7\7	10\10	14\9	9\9
	S3	1\16	7\12	10\10	4\14	5\15	11\11	14\9	8\13
I	S4	1\16	9\9	14\4	8\8	6\11	9\9	20\0	20\0
	S5	3\8	7\7	15\5	11\6	6\6	10\5	18\3	14\4
	S6	3\8	10\10	11\11	9\9	5\10	12\12	12\12	9\9
	S7	0\20	9\14	9\14	0\20	3\18	12\12	12\12	3\18
	S8	0\20	9\9	13\8	0\20	4\14	9\9	18\3	8\8

Figure 3: Metagame Matrix for Prisoner's Dilemma

of moves of his opponent {c, d}, to the set of moves of player I {C, D}. A strategy will be understood as a response of player I to the previous move of player II, following the specified first move. The strategies of player I, S1 through S8, as generated by RRS, are:

S1: (D (D (c d)) C)

S2: (D (C (c)) (D (d)))

S3: (D (C (c d)) D)

S4: (D (D (c)) (C (d)))

S5: (C (D (c d)) C)

S6: (C (C (c)) (D (d)))

S7: (C (C (c d)) D)

S8: (C (D (c)) (C (d)))

The first strategy, for instance, instructs the player to start with D and then play D no matter what the opponent did in their most recent encounter. The sixth strategy, which is the "Tit-for-Tat" strategy, calls for an opening with C and responding with C to c and with D to d.

The strategies of the other player are defined symmetrically. The payoffs of the players exercising these strategies in the PD game played four times can be depicted in the meta-PD matrix shown in Figure 3. Using standard game-theoretic techniques it can be determined that strategy S6 ("Tit-for-Tat"), is a dominant, equilibrium strategy for each of the players in the meta-PD in this case.

The metagame approach can be applied in RMM by replacing the original payoff matrix by a metamatrix with strategies instead of individual moves, and with payoffs reflecting the accumulation of the outcomes over the expected number of interactions. For the case of the meta-PD game matrix above (and for other meta-PD games with the expected number of interactions over four), RMM chooses "Tit-for-Tat" strategy as a rational one.

Repetition in Example Interaction

Unlike the pure metagame approach, however, RMM can deal with more realistic situations in which options and payoffs of other players are uncertain. Returning to our robotic example of a one-time interaction (Figure 1), recall that we introduced a variation in which R1 thinks that R2 will correctly identify RI as type A with a probability 0.9. Unlike the equiprobable case where RI decides to pursue G1, this skewed probability leads it to choose G2. As we mentioned, it can also be called an uncooperative option, in the sense that it treats R2 as if it were not there, and that R2 would prefer that RI choose GI, if it were either type A or B. Of course, RI would then welcome reciprocation by R2's choice of G2, if it happens to be type A.

We have applied the metagame approach for the repeated case of the above interaction using RMM. The hierarchy of matrices depicting the accumulated payoffs for all of the possible strategies in the above example is too large to include here, so we just report our end result. The best strategy of RI, as RMM finds, is: (GI (G2 (GI S)) (GI (G2)) S). This strategy directs RI to choose G1 initially, rather than the uncooperative choice of G2. In the subsequent interactions RI will reciprocate with GI in response to G2, the cooperative alternative of R2. R2's uncooperative choices, GI and S, will cause RI to pursue G2. RI will never choose S.

RMM reaches this result as a stable outcome for an expected number of interactions over three. For 2 or 3 interactions, RMM stabilizes and gives equal preference to three strategies (GI (GI (GI G2)) (G2 (S)) S), (GI (G2 (GI S)) (GI (G2)) S), and (GI (S (GI)) (GI (G2)) (G2 (S))). For a one time interaction, the result mirrors the one obtained previously, directing RI to pursue G2.

The result of RMM applied to R2 in the above scenario also provides the cooperative strategy for it, if it is of type A. Its behavior for one-time interaction was to choose S. For multiple interactions, the choice strategy for R2 is a cooperative strategy of starting with G2 and always responding with G2 to R1's G1. That means that RI will not be disappointed when counting on R2 to pursue G2 and cooperation between the two agents will result.

Conclusions and Further Research

We have presented a powerful method, called the Recursive Modeling Method, that we believe rational, autonomous agents should use to interact with other agents. RMM uses all of the available information the agents may have about each other and themselves, modeling the uncertainties as probability distributions. It explicitly accounts for the recursive nesting of beliefs evident in agents' encounters, in which their decisions depend on what they expect others to do. RMM can also easily be extended to account for the possibility of repetitive interactions. We have shown how this fact influences the agents' willingness to exhibit cooperative behaviors toward each other.

There are a number of issues regarding RMM and its extension to repetitive interactions that remain to be investigated. They include the choice of the level of the elaboration of plans that are to be included as options in the scheme and its cost and benefit characteristics. For repetitive interactions, the influence of previous encounters on predictions for the future has to be addressed more rigorously. The potential computational burden of examining all of the possible strategies, particularly as the number of agents grows larger, may become an obstacle in applying our method, and ways to remedy this problem that involve the concept of "bounded rationality" in the agents are being investigated. We believe that RMM also offers an excellent tool for studying communication [Gmytrasiewicz *et al,* 1991b]. Finally, our method uses an intentional approach normatively assuming that other agents will do what seems to be rational for them. To deal with realistic situations where agents can update models of each other through observation and plan recognition, we will complement RMM with an empirical method [Gmytrasiewicz, 1991].

References

[Axelrod, 1984] Robert Axelrod. *The Evolution of Cooperation.* Basic Books, 1984.

[Bennett and Huxham, 1982] P. G. Bennett and C. S. Huxham. Hypergames and what they do: A 'soft OR' approach. *Journal of Operational Research Society,* 33:41-50, 1982.

[Cohen and Levesque, 1990] P. R. Cohen and H. J. Levesque. Rational interaction as the basis for communication. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication.* MIT Press, 1990.

[Dennett, 1986] D. Dennett. Intentional systems. In D. Dennett, editor, *Brainstorms.* MIT Press, 1986.

[Gmytrasiewicz *et al.,* 1991a] Piotr J. Gmytrasiewicz, Edmund H. Durfee, and David K. Wehe. Combining decision theory and hierarchical planning for a time-dependent robotic application. In *Proceedings of the Seventh IEEE Conference on AI Applications,* pages 282-288, February 1991.

[Gmytrasiewicz *et al.,* 1991b] Piotr J. Gmytrasiewicz, Edmund H. Durfee, and David K. Wehe. The utility of communication in coordinating intelligent agents. In *Proceedings of the National Conference on Artificial Intelligence,* July 1991.

[Gmytrasiewicz, 199l] Piotr J. Gmytrasiewicz. Rational reasoning system: Application of decision theory in autonomous robotics. Technical report, University of Michigan, In preparation 1991.

[Howard, 1966] N. Howard. The theory of metagames. *General Systems,* 11:167-200, 1966.

[Jackobs and Kiefer, 1973] W. Jackobs and M. Kiefer. Robot decisions based on maximizing utility. In *Proceedings of the Third International Joint Conference on Artificial Intelligence,* pages 402-411, August 1973.

[Luce and Raiffa, 1957] R. D. Luce and H. Raiffa. *Games and Decisions.* John Wiley and Sons, 1957.

[Neapolitan, 1990] Richard E. Neapolitan. *Probabilistic Reasoning in Expert Systems.* John Wiley and Sons, 1990.

[Reagade, 1987] Rammohan K. Reagade. Metagames and metasystems. In John P. van Gigch, editor, *Decisionmaking about Decisionmaking.* Abacus Press. 1987.

[Rosenschein and Breese, 1989] Jeffrey S. Rosenschein and John S. Breese. Communication-free interactions among rational agents: A probablistic approach. In Les Gasser and Michael N. Huhns, editors, *Distributed Artificial Intelligence,* volume 2 of *Research Notes in Artificial Intelligence,* pages 99-118. Pitman, 1989.

[Rosenschein, 1988] Jeffrey S. Rosenschein. The role of knowledge in logic-based rational interactions. In *Proceedings of the Seventh Phoenix Conference on Computers and Communications,* pages 497-504, Scottsdale, AZ, February 1988.

[Smith, 1984] J. M. Smith. The evolution of animal intelligence. In C. Hookway, editor, *Minds, Machines and Evolution.* Cambridge University Press, 1984.

[Vane and Lehner, 1990] R. R. Vane and P. E. Lehner. Hypergames and AI in automated adversarial planning. In *Proceedings of the 1990 DARPA Planning Workshop,* pages 198-206, November 1990.

[Werner, 1989] Eric Werner. Cooperating agents: A unified theory of communication and social structure. In Les Gasser and Michael N. Huhns, editors, *Distributed Artificial Intelligence,* volume 2 of *Research Notes in Artificial Intelligence.* Pitman, 1989.

[Zlotkin and Rosenschein, 1989] Gilad Zlotkin and Jeffrey S. Rosenschein. Negotiation and task sharing among autonomous agents in cooperative domains. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence,* pages 912-917, August 1989.

[Zlotkin and Rosenschein, 1990] Gilad Zlotkin and Jeffrey S. Rosenschein. Negotiation and conflict resolution in non-cooperative domains. In *Proceedings of the National Conference on Artificial Intelligence,* pages 100-105, July 1990.