Implementing a Decision-Theoretic Approach to Game Theory for Socially Competent Agents

Piotr J. Gmytrasiewicz

Department of Computer Science and Engineering
University of Texas at Arlington, Arlington, TX 76019-0015
piotr@cse.uta.edu

Abstract. We present a summary of our work that applies decision-theoretic paradigm of expected utility maximization to design rational socially competent agents. Our framework is aimed at explicitly representing what the agent knows about the world, what it knows about the other agents, what it knows about what they know, and so on. We concentrate on realistic cases of finite knowledge states, and, for these states, we define a solution method for our representation, i.e., a process by which the agent can process the representation of its state of knowledge and arrive at an assignment of expected utilities to its available physical and communicative actions. The agent's executing actions that maximize its expected utility lead to it being rational while coordinating and communicating with the other agents.

1 Introduction

The goal of our research is to implement artificial agents that exhibit social competence. We define social competence as an agent's ability to coordinate its actions with actions of other agents sharing its environment, and its ability to effectively communicate with them. We define coordination as a choice of action that takes into account the anticipated actions of the other agents. By effective communication we understand the agent's ability to choose, among possibly large number of alternatives, the proper communicative action – one with the content best suited to the situation at hand, and one that is transmitted using the most appropriate language encoding and communication medium. The artificial agents we are interested in should be capable of competent social interactions not only within a homogeneous group of agents of same design; they should be able to interact with artificial agents of radically different designs and with human beings. Further, they should not need to rely on the assumption that the other agents are cooperatively predisposed; they should cooperate with others when it is appropriate, but be able to act in their own best interest if the situation demands it.

We believe that the paradigm of decision theory [8, 9, 11, 13] is able to provide the set of principles guiding the development of such socially competent agents.

[0] This research is supported, in part, by the Office of Naval Research grant N00014-95-1-0775, and by the National Science Foundation CAREER award IRI-9702132.

Our reason for turning toward decision theory is that it offers an answer to one of the central issues in designing intelligent agents, as articulated by Newell [28] within his knowledge-level view: How should an agent represent knowledge, and how should it operate on it to arrive at purposeful, rational action? Decision theory provides an answer by postulating that probability (intuitively, what an agent knows) [7, 32] and utility (i.e., what an agent prefers) [14, 39] be combined to define rationality as the agent's behavior that maximizes its expected utility [34, 36]. It can be shown (see [34] and [36]) that, if the agent's beliefs are consistent with the axioms of probability theory and its preferences are consistent with the axioms of utility theory, then the agent should make its decisions so as to maximize its expected utility.

The issue of a rational agent interacting with other rational agents is one that is central to game theory. While the main stream of research in game theory assumes, and attempts to refine, the central game-theoretic solution concept of Nash equilibrium, we are taking the knowledge level perspective, understood in decision-theoretic terms, which leads toward a less known area of game theory, called a decision-theoretic approach to game theory [22, 33]. The difference between the two is illustrated by the following quote, due to Myerson [27]:

> The decision-analytic approach to player i's decision problem is first to assess some subjective probability distribution to summarize i's beliefs about what strategies will be used by the other players and then to select a strategy for i that maximizes his expected payoff with respect to these beliefs. [...] In contrast, the usual game-theoretic approach is to analyze and solve the decision problems of all players together, like a system of simultaneous equations in several unknowns.

Clearly, the strategy of first assessing a probability of other agents' alternative actions, and then selecting the agent's own best strategy can be applied recursively to the other agents as well since their own actions depend on what they, in turn, expect of others, including the original agent. While, as Myerson points out, it seems that this recursion leads to circularity, we do not believe this is the case. The reason is that the solution of the agent's own decision problem, and what the agent thinks another agent thinks is solution to this agent's decision problem, are arrived at based on different bodies of information. The first problem is solved based on the information the original agent has about the environment and about other agents, i.e., based on its own decision-making situation. The second problem, on the other hand, is solved based on the information the original agent has about what the other agent knows about the decision-making situation of the original agent. Clearly, these two do not have to coincide. In particular, it may be that the original agent, while knowing its own decision-making situation quite well, has less information about what another agent knows about the original agent's decision-making situation. The postulate that the agents do run out of information at some level of nesting is the basis of our implementation of the solution using dynamic programming, as we describe briefly below.

From another perspective, the difference between the traditional game theory and decision-theoretic approach to multiagent interactions stems from the different questions being asked. As an informal example, consider the task of mechanism design, for which traditional game theory is perfectly suited. The mechanism (for example an auction, or a negotiation) is to be applied by individuals while interacting with each other, but these individuals are not being designed or otherwise controlled by the game theorists (usually, they are humans). A mechanism is judged as good, typically, when it it is expected to result in desirable properties of the multiagent system as a whole, when, and if, the system reaches equilibrium (see [15] Chapter 7). Here, the equilibrium is defined without any reference to the state of knowledge of the agents involved. In the decision-theoretic approach, on the other hand, the central question is framed from the individual agent's point of view: Given my state of knowledge of the multiagent environment, what is my rational decision? Thus, while the AI knowledge-level problem of how an agent is to represent and process is available knowledge is peripheral in traditional game theory, it is central to the decision-theoretic approach.

The interest in the decision-theoretic approach is justified not only for those taking the AI knowledge-level perspective, however. Over the years if has become increasingly clear that the assumption of common knowledge, which underlies the concept of Nash equilibrium, is problematic [1–3, 6, 16, 22, 35, 38]. Similar results have been arrived at in the field of distributed systems. Halpern and Moses [20] have proven that arriving at the state of common knowledge in a realistic distributed environment is impossible. Take a simple realistic example situation of agents that have interacted with other agents for a finite amount of time, and therefore could have exchanged only a finite number, say K, of messages. Clearly, the information they have about each other cannot possibly be nested deeper then K, unless they jumped into unjustified conclusions (see the proof in [20] for details). This reasoning supports our postulate that, in such realistic situations, there will be a level of nesting of the agents' beliefs below which the agents will have *no information* about each other.

Finally, it could be postulated that while it is not strictly justifiable for the agents to assume common knowledge and use the game-theoretic equilibrium solution concept, it may be convenient for them to do so and rely on the rich body of research in traditional game theory. While it may be the case in some domains, we found that this approach encounters problems with multiple equilibria. In the domains we studied, for example in the pursuit game, the cooperative assembly, and in the air defense domain, there are usually many equilibria and no clear way for the agents to choose among them. In particular, in the air defense domain there are as many equilibria as there are alternative targets that can be intercepted, rendering this approach useless as the strategy to select the best target to shoot at next.

Below, we briefly summarize our work that takes the decision-theoretic perspective and implements it as our Recursive Modeling Method (RMM). We mention the issues of knowledge representation, and motivate the dynamic programming solution of the agent's decision problem. We mention results of coordination with and without communication among our agents, and compare them to performance of coordinating human teams.

2 Representation

Our design of the knowledge base (KB) is based on work on frame-based [5, 23] and object-oriented [31, 40] knowledge representation formalisms (see also [4] and references therein). These formalisms postulate that the KB be organized as a set of interrelated frames representing classes, i.e., sets of entities, and instances, i.e., the individual entities themselves. The frames representing the classes form a superclass/subclass hierarchy (ontology) allowing for usual inheritance of properties, while the leaves of the hierarchy are occupied by instances of classes identified in the agent's environment. An example of such KB for the air defense domain is depicted in Figure 1.

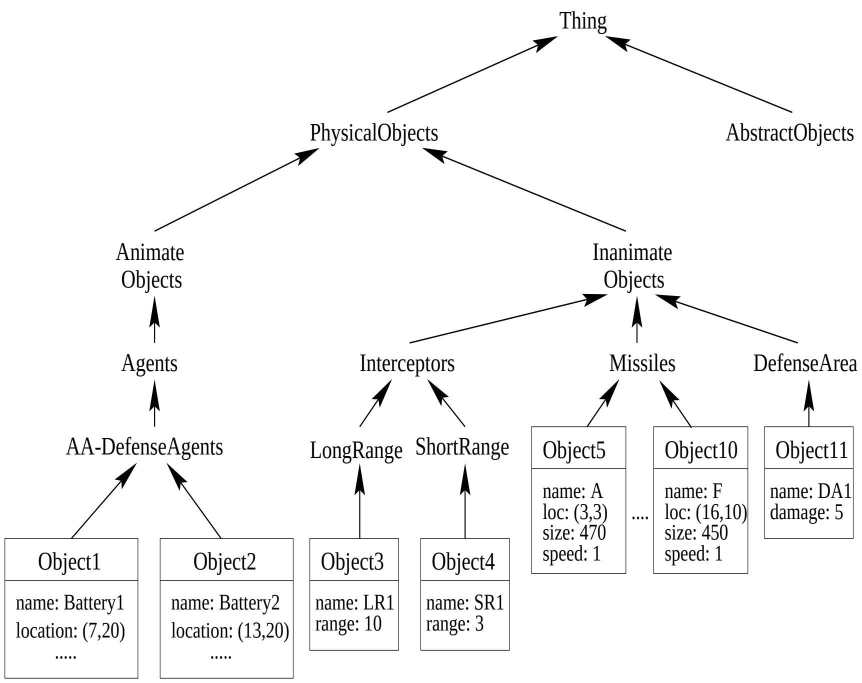

Fig. 1. A High-level View of a KB design for the Air Defense Domain Scenario.

A fundamental limitation of the frame formalisms is that they do not support uncertainty, which, as we mentioned, is ubiquitous and particularly important for socially competent agents. First, the properties of objects in the environment may be not known with certainty. Second, to effectively interact, the agents have to be able to represent and reason about limits of knowledge of other agents, uncertainty of others' knowledge, their own uncertainty as to others' states of knowledge, and so forth.

The framework widely used for uncertainty reasoning is that of Bayesian networks [32], which represent probabilistic relations among uncertain variables describing the domain at hand. Our KB design builds on an effort that overcomes some of the limitations of frame-based systems by combining them with Bayesian networks [24, 25]. The basic idea is to treat the slots of frames (or attributes of objects) describing the properties that may not be known with certainty as nodes of a Bayesian network.

For example, in KB used by the Battery1 in scenario depicted in Figure 2, the speed and the range of the missile objects may be uncertain, subject to inaccuracies of the radar systems. Further, an interceptor launched against a missile may destroy it only with some probability, which, in turn can cause the missile to penetrate the defenses, strike its intended target, and cause damages, magni-

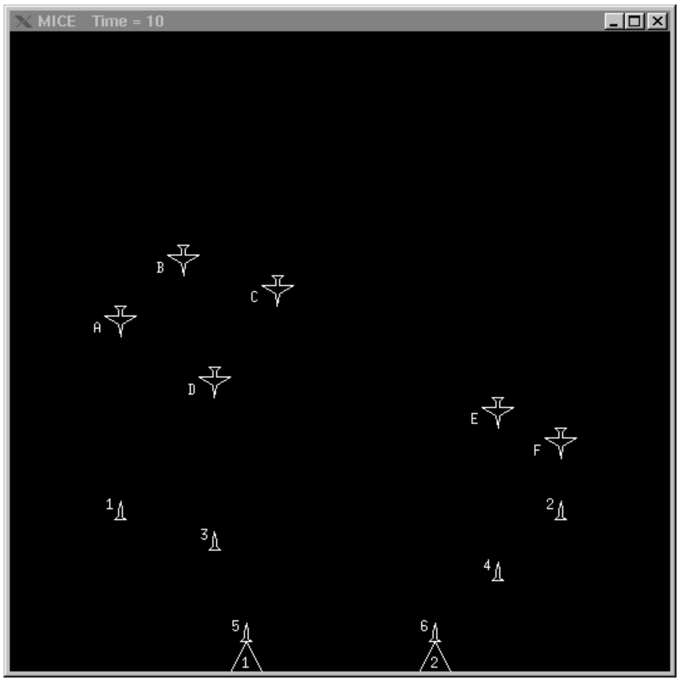

Fig. 2. A Simulation of the Air Defense Domain.

tude of which is also uncertain. The relations among these and other uncertain attributes are depicted in a network in Figure 3.

While it is clearly useful to combine the frame-based and Bayesian representations to overcome the limitation of traditional frames in representing uncertainty, the other important advantage is in making it possible to generate the Bayesian representation of decision-making situation on-the-fly. In our implementation, for example, as new objects are identified by the agent's sensory systems, they are automatically represented as objects belonging to appropriate classes in the frame-based KB, and automatically become part of the Bayesian representation of the agent's decision-making situation. For lack of space we refer the interested reader to papers available on our web site (http://dali.uta.edu) for further details. Now, we bring in results that allow us to convert the above representation to payoff matrices used in the Recursive Modeling Method.

The following theorem establishes a correspondence between influence diagrams, contained in the agent's probabilistic frame-based KB, and payoff matrices used for reasoning in multiagent situations:

Theorem 1: Each influence diagram has a corresponding and unique payoff matrix representing the same decision-making situation.

Proof (Outline): The proof proceeds by repeated absorption [37] of the nature nodes[1] other than the node(s) representing the actions of the other agent(s) in the influence diagram. The information contained in the absorbed nodes is compiled into the probabilities and values of the remaining nodes. When the only nodes remaining are the decision node, the utility node, and the nodes rep-

[1] The absorption operation is available in many of the Bayesian network tools, including Netica.

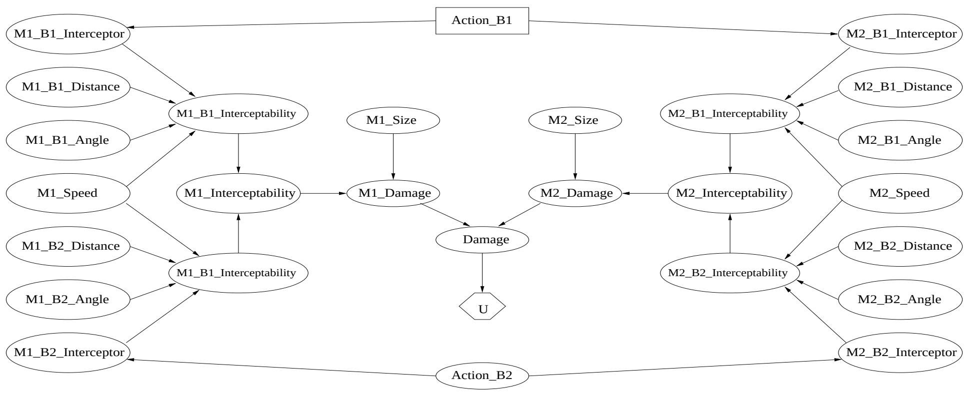

Fig. 3. Influence Diagram Spanning the KB of the Battery1.

resenting the actions of the other agent(s), it is straightforward to tabulate the values of the utility node depending on the values of the agents' action nodes and to form a payoff matrix.□

An example matrix for the air defense scenario in Figure 2 is depicted on top of Figure 5. The rows and columns are labeled with the alternative actions of the agents, and the entries in the matrix are payoffs (i.e., the expected utilities) resulting from the agents' executing their corresponding alternatives. For example, the entry 65.1 in the row labeled B and column labeled F means that the expected payoff (in terms of the threat reduction, as we explain in [29, 30], also on http://dali.uta.edu) to Battery1 resulting from Battery1's attempting to intercept missile B and Battery2's attempting to intercept missile F is 65.1.

As we mentioned, the agent's information about the other agent's decision-making situation is also represented in the KB. In the scenario above we assumed that Battery1 is uncertain whether Battery2 has short range interceptors, or whether it has been incapacitated. Thus, Battery1 has three alternative models of the state of knowledge of Battery2, depicted in Figure 4. The models representing different decision-making situations of Battery2 can also be compiled, and these, together with the Battery1's own payoff matrix for RMM's recursive model structure, depicted in Figure 5.

In Figure 5 the matrix corresponding to the possibility that Battery2 has both long and short range interceptors has seven rows, representing the Battery2's ability to intercept any of the six missiles visible, in addition to its possibly remaining still. The matrix representing the possibility that Battery2 does not have any short range interceptors has less rows since Battery2 cannot target the missiles that are too close to it, i.e., D, E and F in Figure 2. The third model on the second level in Figure 5 corresponds to the possibility that Battery2 in incapacitated. Since, in this case, Battery2 cannot attempt to intercept any of the missiles and has to remain stationary, this model is simply a probability distribution over Batter2's possible actions that assigns 100% probability that Battery2 will execute it's "s" option, i.e., stay put. The fourth model on the second level in Figure 5 is the *No-Info*[1] model. Since it represents the fact that Battery1 has no information it can use to model Battery2, this model assigns equal probability to each of the Battery2's seven alternative actions. The uniform

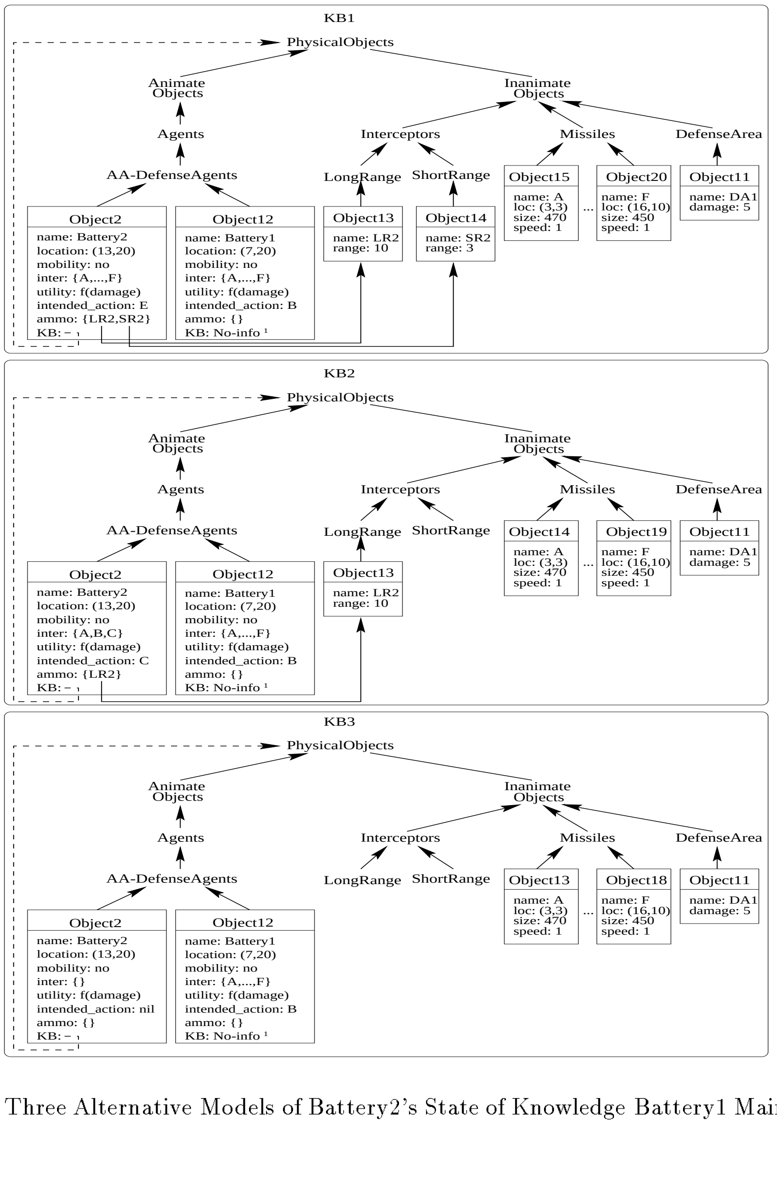

Fig. 4. Three Alternative Models of Battery2's State of Knowledge Battery1 Maintains.

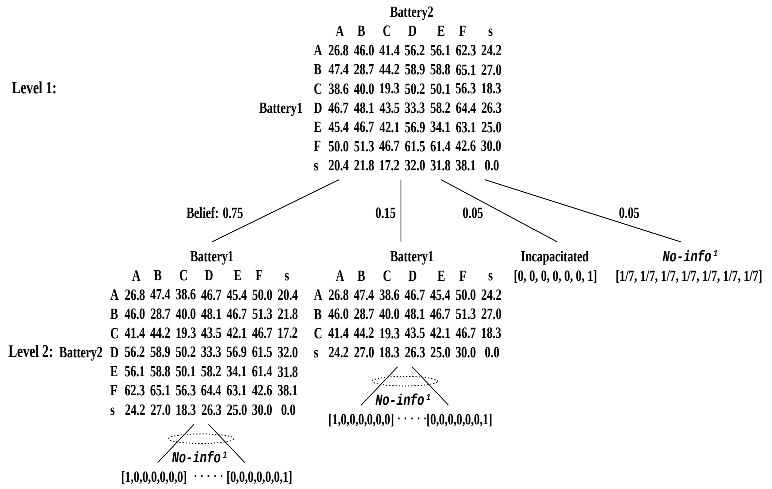

Fig. 5. Hierarchy of Payoff Matrices Corresponding to Nested Models of Agents.

probability distribution has a maximum entropy and contains zero information, correctly representing the state of Battery1's knowledge in this case.

3 RMM's Solution Method

We now briefly describe how dynamic programming method that can be used to solve recursive model structures, like the one in Figure 5; the reader is referred to [17] for details.

Dynamic programming is an appropriate implementation of decision-theoretic approach to games since the decision-making problem an interacting agent has to solve exhibits the properties of optimal substructure and overlapping subproblems, which are characteristic for dynamic programming [10]. Optimal substructure reveals itself when we notice that an agent making the optimal decision in presence of other agents has to solve their decision-making situations first. In other words, the optimal solution of the agent's decision-making problem contains, as subproblems, the optimal solutions to the decision-making problems the agent thinks the other agents are facing. The fact that other agents may have information about others, including the original agent, gives rise to recursive nesting of models. The dynamic programming solution explores this recursion and, after the recursion bottoms out, propagates the results of solving the models, such as the one in Figure 5, bottom-up. The property of overlapping subproblems reveals itself when we realize that some of the alternative models the agents have of other agents may be identical to the models other agents have.

The above procedure is applicable, of course, only if the recursive model structures, like the one in Figure 5, are finite and actually do bottom out. We argue that this is always the case. Since the agents have finite amount of information about others, and what others know about others, and so on, the nesting of models has to terminate when the agents run out of information. We represent it by using the *No-Info* models at the leaves of the recursive model structures.

The property of overlapping subproblems reveals itself when we realize that some of the alternative models the agents have of other agents may be identical to the models other agents have. While this is not the case in the structure in Figure 5, it is easy to see if we consider three interacting agents. Then, the first agent may think that its model of the second agent is the same as the third agent's model of the second agent.[2] Once the overlapping models are identified, they can be solved once and their solution can be cached and reused in other parts of the recursive model structure. In domains in which such overlap does exist, it can substantially reduce the complexity and run time.[3]

We now describe the solution of recursive model structures using dynamic programming using the structure in Figure 5 as an example. As we mentioned the solution proceeds bottom-up. Starting with the model of Battery2 on the left consisting of the 7×7 payoff matrix and a *No-Info*^{1} model below it, we use sampling over the space of the alternative distributions in the *No-Info*^{1} model, and, for each distribution compute the best action for Battery2.[4] For example, if the sampled distribution is $[0.9, 0.1, 0, 0, 0, 0, 0]$, then the expected utilities of Battery2's alternative actions A through F and s, are 28.72 ($= 26.8 \times 0.9 + 46.0 \times 0.1$), 45.53, 38.74, 46.84, 45.53, 50.13 and 20.27, respectively. This shows that is Battery2 thinks that Battery1 will execute its action A with probability 0.9 and its action B with probability 0.1, then the best action for Battery2 would be to intercept missile F, since it has the highest expected payoff in this case.

To summarize the results of processing all of the distributions in the *No-Info*^{1} model, our sampling algorithm counts the proportion of distributions that favor each particular action of Battery2. It turned out that 78% of them favored Battery2's action F (as the distribution above), 10% favored action E, and 12% favored D. This means that, given that it is unknown what Battery2 expects Battery1 to do, one can assign the following probability distribution to alternative behaviors of Battery2: $[0, 0, 0, 0.12, 0.10, 0.78, 0]$. This is the case in which Battery2 has both kinds of ammunition. Similarly, processing the next model of Battery2 in Figure 5 according to which it has no short range interceptor, results in the distribution $[0.37, 0.58, 0.05, 0, 0, 0, 0]$, which means that, if Battery2 has no short interceptors, we can expect it to shoot at missile A with probability 37%, at missile B with probability 58%, and at missile C with probability 5%.

Now, using the four alternative models of Battery2, Battery1 can form an overall probability distribution describing Battery2's likely actions. The four probability distributions, describing Battery2's expected behavior in each of four models, are combined with weights representing the likelihood that each of these models is correct, i.e., 0.75, 0.15, 0.05 and 0.05, respectively. This procedure is called Bayesian model averaging [21]. The result is $[0.063, 0.094, 0.015, 0.097, 0.082, 0.592, 0.057]$, which summarizes Battery1's overall expectation as to Battery2's alternative actions. Now, Battery1 can compute the expected utilities of its own alternative actions using its payoff matrix (on top on Figure 5) and the probabilities of Battery2's actions, as:

$$\begin{aligned} U_A = {} & 0.063 \times 26.8 + 0.094 \times 46.0 + 0.015 \times 41.4 + 0.097 \times 56.2 + \\ & 0.082 \times 56.1 + 0.592 \times 62.3 + 0.057 \times 24.2 = 54.9460 \end{aligned}$$

[2] See our implementation of the pursuit coordination game for an example of such overlap, at http://dali.uta.edu.

[3] In the pursuit game, taking advantage of overlapping subproblems allowed us to reduce the running time of decision-making from exponential to polynomial.

[4] As we mentioned, the alternative distributions represent different ways in which Battery2 may be modeling Battery1's intended action.

Using the same probability distribution we can similarly compute: $U_B = 56.9601$, $U_C = 49.8316, U_D = 55.7424, U_E = 54.9772, U_F = 46.6022$ and $U_s = 31.8592$.

The final result shows that Battery1's best action is to attempt to intercept missile B, and to expect the payoff of 56.96. This conclusion is derived from processing all of the information Battery1 has about the environment and about Battery2 contained in the knowledge base. Based on the information it has, Battery1 formed expectation as to Battery2's likely actions, and chose its own action appropriately. The resulting action is, therefore, *coordinated* with action expected of Battery2.

4 Rational Communication

We identify a communicative act with its *decision-theoretic (DT) pragmatics*, defined as the transformation of the state of knowledge about the decision-making situation (i.e., the recursive model structure described above) the act brings about [18].[5] The transformation in the agent's decision-making situation may change the expected utilities of alternative actions. It is natural to identify the change of the expected utility brought about by a communicative action as the expected utility of this action itself.[6]

Formally, the expected utility of the communicative act M, executed by an agent, is the difference between the payoff the agent expects before and after executing the act:

$$U(M) = U_{p^M}(Y) - U_p(X). \tag{1}$$

where $U_p(X)$ is the utility of the best action, labeled X, expected before sending the message (in the example solved in previous section the best action was B and its expected utility was 56.96), and $U_{p^M}(Y)$ is the utility of the best action, Y, expected if a message were to be sent.

In considering the set of messages available to an automated agent, we classify the messages into types: intentional messages, modeling messages, questions, imperatives, etc. [19]. Intentional messages describe the intended actions of the agent, and modeling messages describe the environment or the status of the agent relevant to the achievement of its goal. Both types of messages are intended to remove some of the uncertainty present in the absence of communication.

The language we use is similar to KQML, and performatives provided in KQML, which include *Attempt* and *Assert* [26]. To send an intentional message our agents use the *Attempt* performatives, and to send a modeling messages, they use *Assert*. For simplicity, we assume here that the agents can mutually understand communication primitives, but it turns out that this assumption can be relaxed.

[5] Since our formalism is probabilistic, it naturally handles cases when the meaning of a message is itself uncertain.

[6] The notion of the utility of a message we use here differs from the notion of the value of information considered in decision theory [32, 36]. The latter expresses the value of information to its recipient. We, on the other hand, consider the value of a message to its sender, since, of course, it is the sender that makes the decision of whether and what, to communicate. The two notions coincide in two special cases: when the preferences of the agents perfectly coincide, and when a speaker requests information from the hearer by asking a question.

5 Experiments in Air Defense Domain

In our implementation, the anti-air defense simulator with communication was written in Common LISP and built on top of the MICE simulator [12]. During the experiments two coordinating defense units were faced with an attack by seven incoming missiles. Therefore, the RMM agents used an 8×8 payoff matrix to represent the agents' decision-making situations. Each of the two defense units was assumed to be equipped with three interceptors each, if they were not incapacitated. Thus, they could launch one interceptor at a time, and did it three times during a course of one defense episode.

5.1 Coordination without Communication

Our experiments was aimed at determining the quality of modeling and coordination achieved by the RMM agents in a team, when paired with human agents, and when compared to other strategies. To evaluate the quality of the agents' performance, the results were expressed in terms of the total expected damage to friendly forces after all six interceptors were launched. The total expected damage is defined as a sum of the residual warhead sizes of the attacking missiles. Thus, if a missile was targeted for interception, then it contributed $(1 - P(H)) \times Size$ to the total damage. If a missile was not targeted, it contributed all of its warhead size value, T_j, to the damage.

The target selection strategies are as follows:

- Random: selection randomly generated.
- Independent, no modeling: selection of arg $\max_j\{P(H_{ij}) \times T_j\}$ for agent i.
- Human:[7] selection by human.
- RMM: selection by RMM.

The random agents were included to provide the worst-case base line of the system performance in our experiments. We included the independent agents to show what coordination can be expected when agents maximize but do not model each other in making their choices. It turned out that the ways that human agents choose a missile were different and sometimes quite arbitrary. For example, some of our human subjects shot only the 3 missiles coming to their own side by dividing the grid world into left and right sides. They sometimes had difficulties in splitting the screen when the missiles were clustered at the center area, which led to much duplicated effort. Others tended to choose missiles with the largest missile size. Still others tried to consider the multiplication of the missile size and the hit probability, but did not model the other agent appropriately.

We experimented with the above policies to understand the agent interactions in two groups: heterogeneous teams of agents with the same policy and the mixed agent teams with different policies.

As shown in Figure 6, we found that the all-RMM team outperformed the human and independent teams. The human performance was very similar to the performance of independent agents. The most obvious reason for this is that humans tend to depend on their intuitive strategies for coordination, and, in this

[7] We should remark that our human subjects were simply CSE and EE graduate students who were informed about the criteria for target selection. We would expect that anti-air specialists, equipped with a modern defense doctrine, could perform better than our subjects. However, the defense doctrine remains classified and was not available to us at this point.

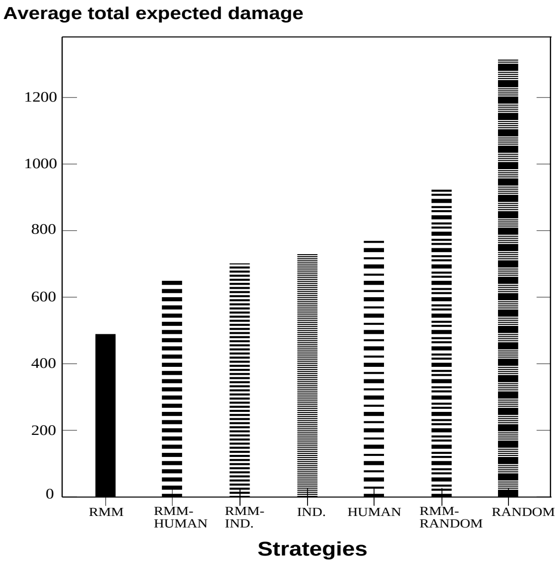

Fig. 6. Average total expected damage (over 100 runs).

case, found it hard to engage in deeper, normative, decision-theoretic reasoning. The common reason for the lower score of human teams was the choice of the greatest threat simultaneously–they made redundant efforts to attack the same missile. This, again, suggests that human agents attempted to minimize their expected damage but did not model the decision making of the other human very well, while RMM agents were rationally optimizing given what they expected of the other agent. The performance of the RMM team was not perfect, however, since the agents were equipped with limited and uncertain knowledge of each other.

The above results show that the decision-theoretic approach to multiagent interactions implemented in RMM allows the automated agents to achieve high quality coordination in this unpredictable environment without relying on predefined protocols. As we argued, methods using traditional game-theoretic equilibria would not be sufficient to coordinate agents in our domain. A particularly promising facet of our results is that the Recursive Modeling Method is a robust mechanism for modeling and coordination not only among RMM agents, but also with the human-controlled agents.

5.2 Coordination with Communication

In the experiments with communication we ran, two defense units were faced with an attack by seven incoming missiles. Therefore, the RMM agents used an 8×8 payoff matrix to represent the agents' decision-making situations. For all settings, each defense unit was initially assumed to have the following uncertainties (beliefs) in its knowledge base: the other battery is fully functional and has both long and short range interceptors with probability 60%; the other battery is operational and has only long range interceptors with probability 20% (in this case, it can shoot down only distant missiles, which are higher than a specific altitude.); the other battery has been incapacitated by enemy fire with probability 10%; the other battery is unknown with probability 10%. Further, each battery has no deeper nested knowledge about the other agent's beliefs.

The warhead sizes of missiles were $470, 410, 350, 370, 420, 450$, and 430 unit for missiles A through G, respectively. In these experiments each of the two defense units was assumed to be equipped with three interceptors, if they were not incapacitated. Thus, they could launch one interceptor at a time, and did it three times during a course of one defense episode.

Table 1. The performances of RMM and human team

Cases	Team (B1-B2)	No of selected targets	Total expected damage	ANOVA
Case I	RMM-RMM	5.95 ± 0.21	717.01 ± 110.71	
(B2:both ammo,	RMM-Human	5.70 ± 0.47	797.39 ± 188.35	f=3.45
w/o comm.)	Human-Human	5.75 ± 0.44	800.45 ± 147.69	
Case II	RMM-RMM	6.00 ± 0.00	652.33 ± 58.97	
(B2:both ammo,	RMM-Human	6.00 ± 0.00	717.93 ± 94.45	f=3.43
w/ comm.)	Human-Human	6.00 ± 0.00	710.20 ± 100.92	
	RMM-RMM	5.83 ± 0.37	852.01 ± 160.79	
Case III	RMM-Human	5.75 ± 0.44	862.70 ± 120.19	f=6.96
(B2:only long,	Human-RMM	5.40 ± 0.50	895.92 ± 127.32	
w/o comm.)	Human-Human	5.30 ± 0.47	997.32 ± 145.15	
	RMM-RMM	5.88 ± 0.32	787.42 ± 110.26	
Case IV	RMM-Human	5.85 ± 0.36	842.50 ± 131.26	f=4.58
(B2:only long,	Human-RMM	5.75 ± 0.44	815.67 ± 133.60	
w/ comm.)	Human-Human	5.80 ± 0.41	908.08 ± 103.25	
	RMM-Incap.	3.00 ± 0.00	1742.22 ± 64.45	
Case V	Human-Incap.	3.00 ± 0.00	1786.86 ± 87.94	f=122.01
(B2:incap. or	RMM-Random	4.86 ± 0.58	1079.52 ± 210.64	
random)	Human-Random	4.85 ± 0.59	1115.57 ± 228.94	

Note) For all of cases, Battery1 is fully functional.
$f_{.05,2,57} = 3.15$, $f_{.01,3,76} = 4.13$

We set up 100 scenarios for RMM team and 20 scenarios for human team. We allowed for one-way communication between defense units before each salvo. If both agents wanted to send messages, the speaker was randomly picked in the RMM team, and the human team flipped a coin to determine who would be allowed to talk. The listener was silent and could only receive messages. For uniformity, in all of the anti-air defense scenarios, each battery was assumed to have a choice of the following communicative behaviors: "No communication," "I'll intercept Missile A." through "I'll intercept Missile G.", "I have both long and short range interceptors.", "I have only long range interceptors.", or "I'm incapacitated.". As human subjects, we used 20 CSE and EE graduate students.

The summary of the experimental runs we performed is shown in Table 1. Table 1 presents the average number of selected targets and the average total expected damage by RMM agents after 100 trials and by human agents after 20 trials. We focus on the performances of three different teams: RMM-RMM, RMM-human, and human-human team. To see whether the differences in results obtained by the teams are not due to chance, we performed an analysis of variance (ANOVA) on the basis of total expected damage. For this purpose we randomly chose 20 trials of the RMM team among the 100 trials available.

In Case I of Table 1, both defense units are fully functional and there is no communication. Since the computed value $f = 3.45$ in ANOVA exceeds $3.15 (= f_{.05,2,57})$, we know that the three teams are not all equally effective at the 0.05 level of significance, i.e., the differences in their performance are not due to chance with probability 0.95. Results in Case II include communication between fully functional defense units. ANOVA test shows that the differences in performance are not due to chance with probability 0.95 again. When the communication was available, the performances achieved by three teams were improved, with the RMM team performing slightly better than the other teams. The experiments in Case III and IV intended to examine how fully functional battery can cope with the situation in which the other battery has only long range interceptors. Like Case I and II, the experimental results of Case III and IV proved the effectiveness of communication. The ANOVA test indicates that the observed differences in the performance of four teams for the target selection are significant at the 0.01 level of significance, and are not due to chance with probability 99%. In Case V, the battery controlled by human or RMM cannot take advantage of communication, because the other battery is incapacitated or behaves randomly. The huge value of $f = 122.01$ tells that the performances of four different teams are clearly distinct.

6 Conclusion

We have outlined our research on decision-theoretic approach to decision-making in multiagent situations. We argued that this approach to agent design complements the traditional game-theoretic approaches, mainly used for mechanism design. We presented some of our experimental results that show the viability of our approach in chosen domain. In our current and future work we will address issues of learning other agents' models, and learning the common communication language the agents can share to effectively communicate. The reader is referred to our web page http://dali.uta.edu for details of these efforts.

References

1. Robert J. Aumann and Adam Brandenburger. Epistemic conditions for Nash equilibrium. *Econometrica*, 1995.
2. Cristina Bicchieri. *Rationality and Coordination.* Cambridge University Press, 1993.
3. Ken Binmore. *Essays on Foundations of Game Theory.* Pitman, 1982.
4. R. Brachman and H. Levesque. *Readings in Knowledge Representation.* Morgan Kaufman, 1985.
5. R. Brachman and J. Schmoltze. An overview of kl-one knowledge representation system. *Cognitive Science*, 9:171–216, 1985.
6. Adam Brandenburger. Knowledge and equilibrium in games. *Journal of Economic Perspectives*, 6:83–101, 1992.
7. Peter Cheeseman. In defense of probability. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 1002–1009, Los Angeles, California, August 1985.
8. H. Chernoff and L. E. Moses. *Elementary Decision Theory.* John Wiley, New York, 1959.

9. L. S. Coles, A. M. Robb, P. L. Sinclar, M. H. Smith, and R. R. Sobek. Decision analysis for an experimental robot with unreliable sensors. In *Proceedings of the Fourth International Joint Conference on Artificial Intelligence*, Stanford, California, August 1975.
10. Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to Algorithms*. The MIT Press, 1990.
11. Jon Doyle. Rationality and its role in reasoning. *Computational Intelligence*, 8:376–409, 1992.
12. Edmund H. Durfee and Thomas A. Montgomery. MICE: A flexible testbed for intelligent coordination experiments. In *Proceedings of the 1989 Distributed AI Workshop*, pages 25–40, September 1989.
13. Jerome A. Feldman and Robert F. Sproull. Decision theory and Artificial Intelligence II: The hungry monkey. *Cognitive Science*, 1(2):158–192, April 1977.
14. Peter Fishburn. *The Foundations of Expected Utility*. D. Reidel Publishing Company, 1981.
15. Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, 1991.
16. John Geanakoplos. Common knowledge. In *Proceedings of the Conference on Theoretical Aspects of Reasoning about Knowladge*, pages 255–315. Morgan Kaufman, August 1992.
17. Piotr J. Gmytrasiewicz and Edmund H. Durfee. A rigorous, operational formalization of recursive modeling. In *Proceedings of the First International Conference on Multiagent Systems, ICMAS'95*, pages 125–132, July 1995.
18. Piotr J. Gmytrasiewicz and Edmund H. Durfee. Rational interaction in multiagent environments: Communication. *Submitted for publication*, 1997. Available in postscript from http://www-cse.uta.edu/ piotr/piotr.html.
19. 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*, pages 166–172, July 1991.
20. Joseph Y. Halpern and Yoram Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, July 1990.
21. David Heckerman and David Maxwell Chickering. A comparison of scientific and engineering criteria for bayesian model selection. Technical Report MSR-TR-96-12, Microsoft Research, Microsoft Corporation, Redmond, WA, 1996.
22. Joseph B. Kadane and Patrick D. Larkey. Subjective probability and the theory of games. *Management Science*, 28(2):113–120, February 1982.
23. P. D. Karp, K. Myers, and T. Gruber. The generic frame protocol. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pages 768–774, August 1995.
24. Daphne Koller and Avi Pfeffer. Object oriented bayesian networks. In *Proceedings of The Conference on Uncertainty in Artificial Intelligence*, August 1995.
25. Daphne Koller and Avi Pfeffer. Probabilisitc frame based systems. In *Proceedings of the National Conference on Artificial Intelligence*, pages 580–587, Madison, WI, July 1998.
26. Y. Labrou and T. Finin. A semantics approach for KQML – a general purpose communication language for software agents. In *Proceedings of the Third International Conference on Information and Knowledge Management*, November 1994.
27. Roger B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991.
28. Allen Newell. The knowledge level. *AI Magazine*, 2(2):1–20, Summer 1981.
29. Sanguk Noh and Piotr J. Gmytrasiewicz. Agent modeling in antiair defense. In *Proceedings of the Sixth International Conference on User Modeling*, pages 389–400, June 1997.
30. Sanguk Noh and Piotr J. Gmytrasiewicz. Rational communicative behavior in anti-air defense. In *Proceedings of the Third International Conference on Multiagent Systems*, pages 214–221, July 1998.

31. P. Patel-Schneider. Practical, object-oriented knowledge representation for knowledge-based systems. *Information Systems*, 15, 1990.
32. Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufman, 1988.
33. H. Raiffa. *The Art and Science of Negotiation.* Harvard University Press, 1982.
34. Eric Rasmusen. *Games and Information.* B. Blackwel, 1989.
35. Philip J. Reny. Extensive games and common knowledge. In *Proceedings of the Conference on Theoretical Aspects of Reasoning about Knowladge*, page 395. Morgan Kaufman, 1988.
36. S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach.* Prentice Hall, 1994.
37. R. D. Shachter. Evidence absorption and propagation through evidence reversals. In *Proceedings of the Fifth Workshop on Uncertainty in Artificial Intelligence*, pages 303–310, University of Windsor, Ontario, CA, 1989.
38. Tommy C. Tan and Sergio R.D.C. Werlang. A guide to knowledge and games. In *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, 1988.
39. Michael P. Wellman. The preferential semantics for goals. In *AAAI91*, pages 698–703, July 1991.
40. P. M. Yelland. Object oriented knowledge representation. *British Telecom Technology Journal*, 11:41–51, 1993.