

Uncertain Knowledge Representation and Communicative Behavior in Coordinated Defense^{*}

Sanguk Noh and Piotr J. Gmytrasiewicz

Department of Computer Science and Engineering
University of Texas at Arlington
{noh, piotr}@cse.uta.edu

Abstract. This paper reports on results we obtained on communication among artificial and human agents interacting in a simulated air defense domain. In our research, we postulate that the artificial agents use a decision-theoretic method to select optimal communicative acts, given the characteristics of the particular situation. Thus, the agents we implemented compute the expected utilities of various alternative communicative acts, and execute the best one. The agents use a probabilistic frame-based knowledge formalism to represent the uncertain information they have about the domain and about the other agents present. We build on our earlier work that uses the Recursive Modeling Method (RMM) for coordination, and apply RMM to rational communication in an anti-air defense domain. In this domain, distributed units coordinate and communicate to defend a specified territory from a number of attacking missiles. We measure the benefits of rational communication by showing the improvement in the quality of interactions the communication results in. We show how the benefit of rational communication measured after the interactions is related to the expected utilities of best messages computed before the interaction takes place. Further, we compare our results to improvement due to communication achieved by human subjects under the same circumstances.

1 Introduction

This paper reports on results we obtained on communication among artificial and human agents interacting in a simulated air defense domain. For artificial agents, we advocate a decision-theoretic message selection mechanism according to which the agents make the communicative decisions so as to maximize their expected utility. Thus, they compute the expected utility of alternative communicative behaviors, and execute the one with the highest value. Following

^{*} This is an extended and revised version of a paper which was presented in Proceedings of the Third International Conference on Autonomous Agents [15]. This work has been supported by the Office of Naval Research Artificial Intelligence Program under contract N00014-95-1-0775, and by the National Science Foundation CAREER award IRI-9702132.

the principle of maximizing the expected utility [19], our agents are intended to be rational in their communicative behavior; they send only the most valuable messages, and never send messages they consider to be damaging given the circumstances at hand.

Following our work in [12, 13, 14], this paper considers a specific domain of simulated anti-air defense. We build on this earlier work, particularly on [14], that reported comparisons of decision-theoretic message selection with messages selected by humans in four simple defense scenarios. For the purpose of this paper, we implemented the communicative behaviors of our agents using KQML [11, 20]. We use a large number of randomly generated, complex defense episodes. The results include communication between our agents and coordinating human agents when both agents are fully functional, when one of the agents lacks certain kind of ammunition, and when one of the agents is incapacitated.

In the anti-air defense the agents, human or artificial, are coordinating their defense actions, in this case interception of threats, with other defense agents. The goal of defending agents is to minimize damages to their territory. To fulfill their mission, the agents need to coordinate and, sometimes, to communicate with other agents. However, since the communication bandwidth is usually limited in a battlefield environment, and disclosure of any information to hostile agents should be avoided, it is critical for a defending agent to be selective as to what messages should be sent to other agents. Endowing the agents with a decision-theoretic method to choose their own communicative behavior on-the-fly frees our agents from depending on communication protocols, frequently advocated in other work. We feel that relying on protocols computed by a designer beforehand could lock the agents into suboptimal behavior in unpredictable domain like the battlefield, in which situations that were not foreseen by the designer are likely to occur.

Our approach uses the Recursive Modeling Method proposed before in [4]. RMM endows an agent with a compact specialized representation of other agents' beliefs, abilities, and intentions. As such, it allows the agent to predict the message's decision-theoretic (DT) pragmatics, i.e., how a particular message will change the decision-making situation of the agents, and how the other agents are likely to react to it. We propose that modeling other agents while communicating is crucial; clearly, without a model of the other agents' states of beliefs it would be impossible to properly assess the impact of the communicative act on these beliefs [2]. Based on the message's DT pragmatics, our method quantifies the gain obtained due to the message as the increase in expected utility obtained as a result of the interaction.

Applying the RMM-based approach to the anti-air defense, we look at the improvement that was achieved when the communicative actions recommended were actually executed during simulated anti-air defense encounters. It turns out that rational communication improves the quality of interaction on the average by 10% for the team of two RMM agents (The average damages the units suffered decreased from 717 to 652. For details, refer to Table 3.), and for the teams of coordinating humans (The average damages decreased from 800 to 710.). Also,

the actual increase in the interaction quality is closely related to the expected utility of messages computed before hand. This is a good indication that our approach computes the “right thing”, and results in a good predictor of true benefits of communication.

As we mentioned, some of the previous approaches to communication rely on predetermined communication protocols [1, 17]. Quintero [17] specified a multi-agent protocol language which can be used to model a cooperative environment. Their rule-based agents (called LALO agents) execute rules if messages are matched to the rules in the message rule set. Barbuceanu [1] postulates that conversation plans and rules be available to the agents for performing the coordinated behavior. In their method, rewards denote the utility of finding the optimal action given a state, and are assumed to be known before run time.

Tambe [21] suggested decision-theoretic communication selectivity to establish mutual belief among agents in a team. This approach is similar to ours in that the focus is on whether or not an agent should transmit a given message (fact) to others. Their approach to selective communication uses a decision tree containing reward and penalty values that are domain knowledge. However, obtaining these values when the environment is dynamic, not fully observable or when the status of another agent is not fully predictable, may be problematic. Our framework, in contrast, does not rely on pre-determined rewards and penalties, but computes them based on possibly incomplete and uncertain models of other agents. Tambe’s work is further closely related to prior work by Cohen [2] and Grosz [8].

2 The Anti-Air Defense Domain

Our model of the anti-air domain¹ consists of a number of attacking missiles and coordinating defense units, as depicted in Figure 1. For the sake of simplicity we will explain our approach in the much simpler case of two missiles and two defense batteries, as depicted in Figure 2. We’ll assume that each of the batteries has only one interceptor (we conducted experiments in more complicated settings involving multiple salvos and more attacking missiles as we explain later). As we mentioned, the mission of the defense units is to attempt to intercept the attacking missiles so as to minimize damages to own territory. Let us note that our model makes coordination necessary; if the defense batteries miscoordinate and attempt to intercept the same threat, the other threat will reach its destination and cause damage proportional to the warhead size.

The agents analyze a particular defense situation in decision-theoretic terms by identifying the attributes that bear on the the quality of alternative actions available to the agent. First, each attacking missile presents a threat of a certain value, here assumed to be the size of its warhead. Thus, the defense units should prefer to intercept threats that are larger than others. Further, the defense units should consider the probability with which their interceptors would be effective

¹ This description of our domain largely coincides with one in [12, 13, 14]. We include it here again for completeness.

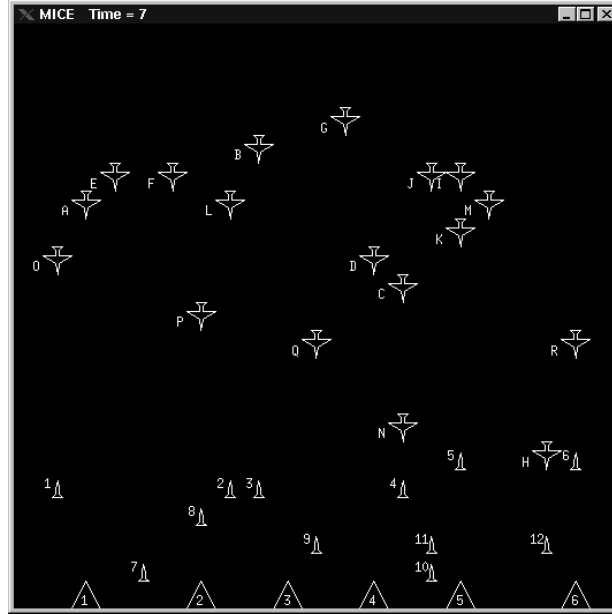


Fig. 1. A complex anti-air defense scenario.

against each of the hostile missiles. The interception probability, $P(H_{ij})$, is dependent on the angle, γ , between the missile j 's direction of motion and the line of sight between the battery i and the missile, as follows:

$$P(H_{ij}) = e^{-\mu\gamma_{ij}}, \quad (1)$$

where μ is an interceptor-specific constant (assumed here to be 0.01).

2.1 Knowledge Representation and Modeling Other Agents

The information an agent has in a defense scenario consists of general knowledge about various kinds (classes) of objects in the environment, which is contained in the agent's knowledge base (KB), and instantiations of these classes that correspond to particular objects in the agent's environment, sometimes called the world model. An example of the KB with a world model corresponding to the scenario² in Figure 2 is depicted in Figure 3. We have implemented the agent's KB and the world model, collectively called the agent's beliefs, and a hierarchical system of classes and objects that contain the properties of the objects recognized in the environment.

In Figure 3, the Anti-Air (AA) Defense Agents is the subclass of Agents, while LongRange and ShortRange Interceptors are subclasses of Interceptors.

² We view the decision making through the eyes of an individual defense unit, Battery2, and its radar-acquired data.

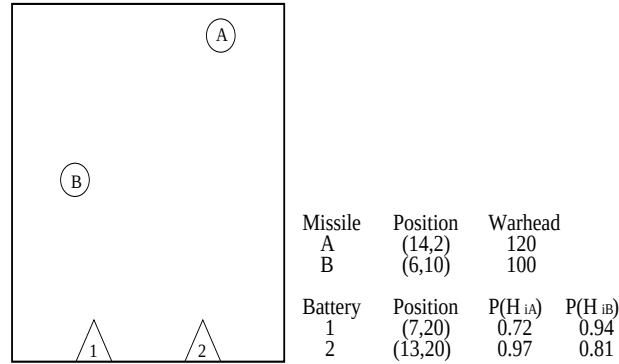


Fig. 2. An example anti-air defense scenario.

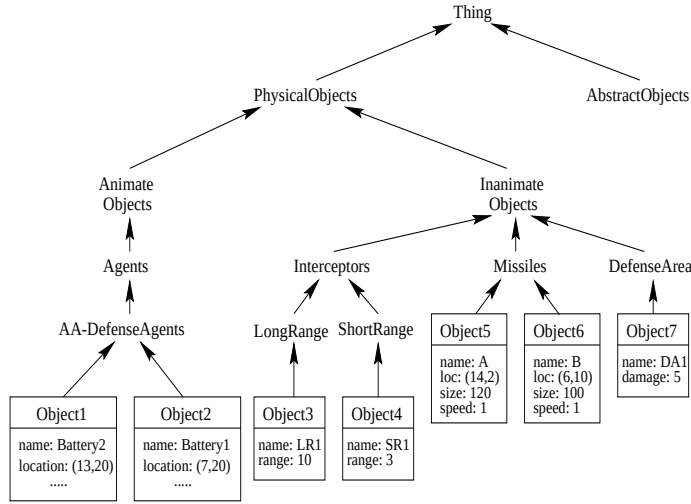


Fig. 3. An example of hierarchical knowledge base of a defense agent, Battery2.

The Objects that are the leaves of the hierarchy in Figure 3 represent that the objects the defense agent Battery2 knows are present in its environment. In this case, these objects are assumed to be recognized with certainty. For example, Object2 was identified as an instance of class AA Defense Agent, while Object5 was correctly identified as an instance of class Missile.

In our example scenario, Battery2 needs to predict action of Battery1 to be able to coordinate effectively. To make this prediction, Battery2 can use the information it has about Battery1, identified as Object2 in Figure 3, and presented in more detail in Figure 4. The actual behavior of Battery1, however, depends on a decision-making situation Battery1 is facing, which in turn includes a number of factors that Battery2 may be uncertain about. For example, if Battery1

Object2
<p>simple slot: name - Battery1 slot value type: string</p> <p>simple slot: location - (7, 20) slot value type: (R,R)</p> <p>simple slot: mobility - no slot value type: {yes, no}</p> <p>function slot: utility - f(damage) slot value type: R</p> <p>simple slot: intended_action - B slot value type: string</p> <p>complex slot: KB slot value type: {KB1, KB2, KB3, No-info¹} probability distribution: {0.8, 0.1, 0.05, 0.05}</p>

Fig. 4. Details of the representation Battery2 has about Battery1.

has been incapacitated or ran out of ammunition, then it will not be able to launch any interceptors. If Battery1 is operational, Battery2 may still be not sure about ammunition Battery1 has. In this case, we assumed that Battery2 thinks that Battery1 may have both types of ammunition, i.e., long and short range interceptors, but that Battery2 is uncertain as to whether Battery1 has any short range interceptors left. If Battery1 has only long range interceptors, it would be unable to attack missile *B* which is too close, and can only attempt to shoot down missile *A*. These three models of Battery1's decision-making situations, corresponding Battery1 being fully operational and having long and short range interceptors, operational with only long range interceptors, and having no ammunition, are depicted as different states of Battery1's knowledge (KB1, KB2, and KB3) in Figure 4, with their associated probabilities 0.8, 0.1 and 0.05, respectively. The fourth model, assigning equal chance to every action of Battery1, is also included to represent the possibility that the other three models are incorrect, in which case nothing is known about Battery1's behavior; this is called a No-Information model.

Thus, the different states of knowledge of Battery1 that define the alternative decision-making situations that Battery1 may be facing, are represented by the complex slot called KB [10] in Figure 4. The value of this slot is clearly related to the case of Battery1 having ammunition of various types or it being incapacitated. For example, if Battery1 has no ammunition at all then its state of knowledge should reflect that fact. In our case, the details of the Battery1's alternative states of knowledge for each case are depicted in Figure 5. These alternatives contain the class hierarchies similar to the one in Figure 3, but they differ in the details of the object instances. In each case it is assumed that Battery1 correctly identified the incoming missiles and the defense area, but in the case in which Battery1 has both long and short interceptors there are

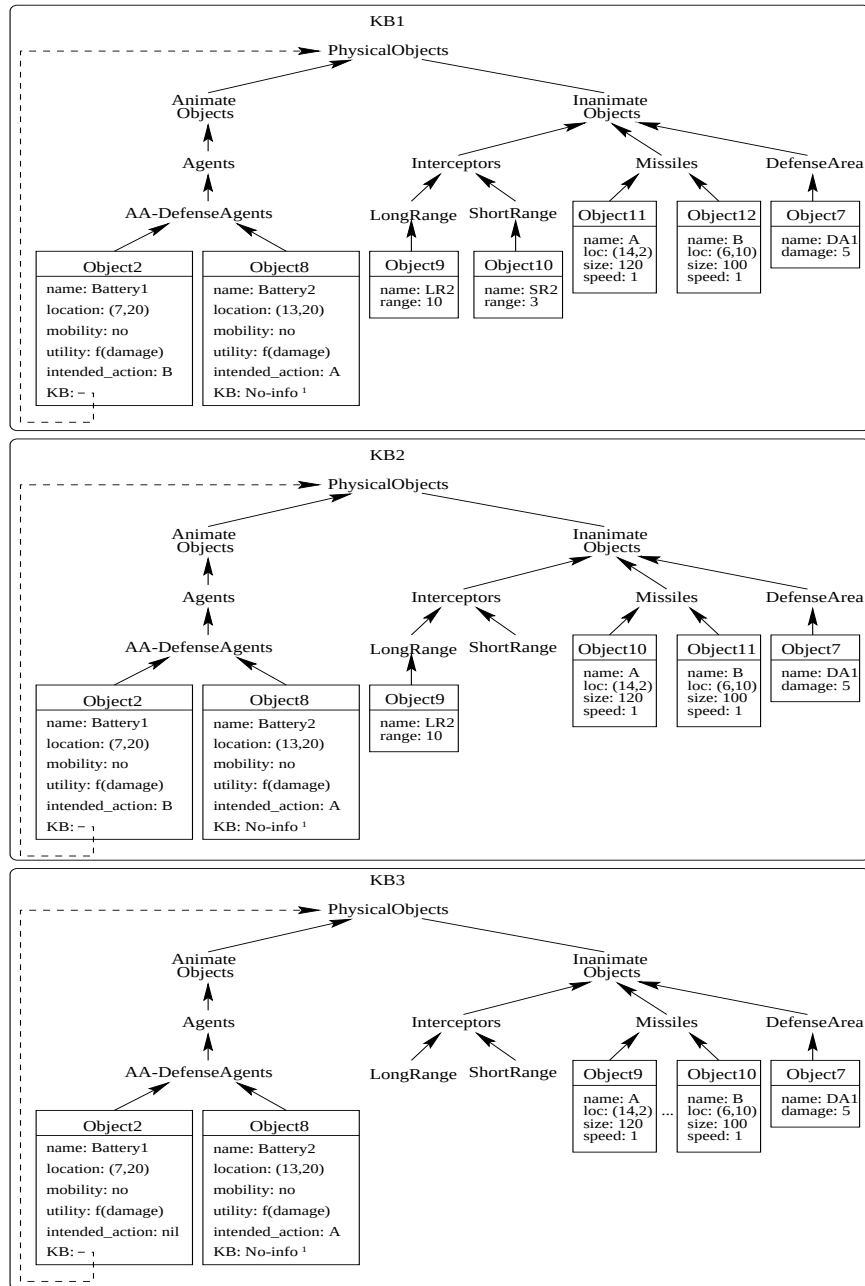


Fig. 5. Three frame based systems Battery2 has about Battery1.

two instantiations of the proper interceptor subclasses (in KB1). If Battery1 has no short range interceptors the short range interceptor class is not instantiated (in KB2), and if Battery1 is incapacitated or has no ammunition at all, none of its interceptor subclasses is instantiated (KB3). That is, they specify that if Battery1 has both types of interceptors, has only long range ones, or has none at all, the state of its knowledge are described by KB1, KB2, or KB3, respectively. The *No-info*¹ model is assigned the remaining likelihood of 5% as we explained above.

2.2 Rational Decision Making

According to our design, the agents compile the information contained in their declarative KB's into compact payoff matrices to facilitate expected utility calculations. The *payoff matrix* [9, 18] used in game theory can be seen to faithfully summarize the information contained in frames by listing the expected payoffs of possible decisions, depending on the parameters (attributes) describing the domain. In the anti-air defense domain, the expected benefit of shooting at threats can be quantified as a product of the sizes of the threats and the interception probabilities. For example, in Figure 6, the combined benefit of Battery1's shooting at the threat *B* and Battery2's shooting at the threat *A* amounts to 210.7 ($= 120 \times 0.97 + 100 \times 0.94$). This value is entered in the top payoff matrix of Figure 6. In this payoff matrix the rows represent the alternative actions of Battery2 of shooting at *A*, *B*, or not shooting, respectively, and the columns represent the alternative actions of Battery1.

In Figure 6, the top payoff matrix which is built from the information contained in Figure 3 compactly represents the decision-making situation that Battery2 is facing. To solve it, Battery2 needs to predict what Battery1 is likely to do; for example if Battery1 were to shoot at *A* it would be best for Battery2 to shoot at *B*, but if Battery1 is expected to shoot at *B* then Battery2 should definitely target *A*.

To construct the payoff matrices in the second level of recursive model structure, we retrieve the information contained Battery2's KB about the actual behavior of Battery1. For example, if Battery1 has been hit and incapacitated or has no ammunition at all, then it will not be able to launch any interceptors. This model comes from the KB3 in Figure 5 and can be represented by the probability distribution of $[0, 0, 1]$ (not shooting). If it is not incapacitated and has both types of ammunition, then its own decision-making situation can be represented as another payoff matrix. This is the left-most matrix in the second level of structure in Figure 6; it is generated from the information in KB1 in Figure 5, and it specifies that Battery1 can shoot at both targets. The information in KB2 can also be compiled into a payoff matrix (second matrix in Figure 6), which shows that, if Battery1 has no short range interceptors it is unable to attack missile *B*, and can only attempt to shoot down missile *A*. These three models, of Battery1 being fully operational and having long and short range interceptors, operational with only long range interceptors, and incapacitated, are depicted as the second level models in Figure 6, with their associated probabilities 0.8, 0.1 and 0.05,

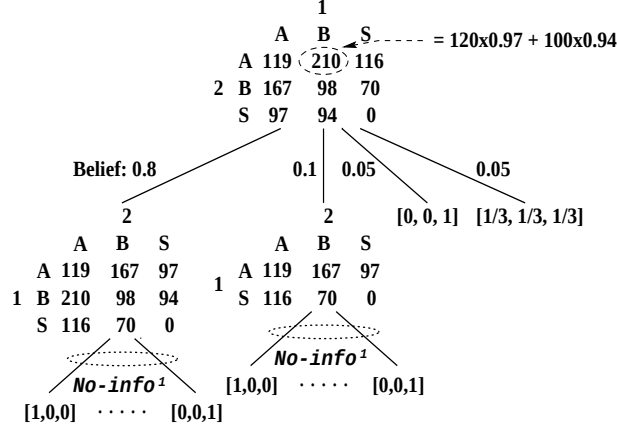


Fig. 6. A recursive model structure for Battery2's decision making.

respectively. The fourth No-Information model, assigning equal chance to every action of Battery1, is also included to represent the possibility that the other three models are incorrect, in which case nothing is known about Battery1's behavior. The probabilities associated with the alternative models can be arrived at using Bayesian learning based on the observed behavior of the other agents, as we described in [7, 13].

To summarize, the structure illustrated in Figure 6, called the recursive model structure, represents Battery2's model of its decision-making situation, arrived from the information contained in the agent's KB. The recursive model structure contains the agent's own decision-making situation (top payoff matrix), and the alternative models Battery2 has of Battery1's decision-making situation. The nesting of models could continue, to include the models Battery2 thinks Battery1 uses to model Battery2, but in this case we assumed that Battery2 has no information about the models nested deeper, and the recursive model structure in Figure 6 terminates with No-Information models. Under the assumption that the nesting of models is finite, the recursive model structures, like the one in Figure 6, can be solved using dynamic programming. The solution proceeds bottom-up and results in expected utility values being assigned to Battery2's alternative behaviors, the best of which can then be executed. For a more detailed solution procedure and further discussion, see [4, 12].

3 Communication in Anti-Air Defense

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) the act brings about. We model DT pragmatics using the RMM representation to investigate the utility of the

communicative acts [6, 5].³ The transformation in the agent’s decision-making situation, as represented by RMM’s recursive model structure, 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. By evaluating the alternative communicative acts in this way, the agent can select and send the highest utility message—the message that causes the greatest gain in the expected utility of the agent’s action.⁴

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:

$$U(M) = U_{p^M}(Y) - U_p(X). \quad (2)$$

where $U_p(X)$ is the utility of the best action, X , expected before sending the message, and $U_{p^M}(Y)$ is the utility of the best action, Y , expected if the message were to be sent. Further details and discussion is contained in [6, 5]. We now apply DT pragmatics to our anti-air defense domain.

3.1 Intentional and Modeling Messages

Our implementation of communicating autonomous agents is closely related to BDI theories that describe the agent’s Beliefs, Desires, and Intentions [2, 20]. In a coordinated multi-agent environment, agents’ communicative acts remove the uncertainty of their *beliefs* about the other agents, and inform their current *intentions* to others. Therefore, in the current implementation, we classify the messages into two groups: intentional messages and modeling messages [5]. Intentional messages describe the intended actions of the agent, and modeling messages describe the environment or the status of agent relevant to the achievement of their goal. Both types of messages are intended to remove some of the uncertainty present in the absence of communication.

The language we use to implement communicative acts is KQML, and we use two performatives provided in KQML: *Attempt* and *Assert* [20]. When our agents want to send intentional messages, they use *Attempt*. In case of modeling messages, they use *Assert*. For simplicity we assume here that the agents can mutually understand communication primitives, as in [1], but this assumption can be easily relaxed, as described in [5]. In air-defense domain, when Battery1 wants to tell Battery2 that, for example, it has both long and short range interceptors, the modeling message will be as follows:

³ Since our formalism is probabilistic, it naturally handles cases when the meaning of a message is itself uncertain.

⁴ The notion of the utility of a message we use here differs from the notion of the value of information considered in decision theory [16, 19]. 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.

```
(assert :sender Battery1
      :receiver Battery2
      :content has(Battery1, long_range_interceptors),
               has(Battery1, short_range_interceptors))
```

The intentional message that declares, for example, that Battery2 will attempt to intercept the threat B , can be defined similarly:

```
(attempt :sender Battery2
      :receiver Battery1
      :content intercept(Battery2,  $B$ ))
```

The choice of attempting to send a message in a coordinated multi-agent environment will be made by autonomous agents themselves. In other words, if the agent computes that it can perform a beneficial communicative act, they will attempt to send the message to others. However, we allow only one agent to communicate at a time. Thus, if two agents simultaneously want to communicate the actual speaker is picked at random; this simulates a collision and back-off on the same communication channel. We also assume that agents are sincere, i.e., not lying. This restriction preserves the truthfulness assumption, and keeps the DT pragmatics of the messages well-defined. When an agent receives a modeling message it updates its *beliefs* according to the message received and executes a newly decided *intention*. If an agent receives an intentional message, the agents make a decision by optimizing their action given the sender's intended behavior contained in the message.

3.2 Expected Utilities of Messages: Examples

With the concrete example of an anti-air defense in Figure 2, we now present the decision-theoretic evaluation of the alternative messages performed by our agents. For simplicity we focus on the communicative behavior of the Battery2. Assume that Battery2 chooses among the five communicative behaviors below:

No Comm: No communication
M1: I'll intercept missile A .
M2: I'll intercept missile B .
M3: I have both long and short range interceptors.
M4: I have only long range interceptors.

Figure 7 depicts the change of the decision-making situation expected due to Battery2's sending a message, *M1*, "intercept(Battery2, A)". As we mentioned, this change is the message's decision-theoretic (DT) pragmatics. It illustrates that, as a result, Battery2 expects Battery1 to know that Battery2 will intercept missile A , which is represented as a $[1, 0, 0]$ probability distribution Battery1 would use to model Battery2's actions.

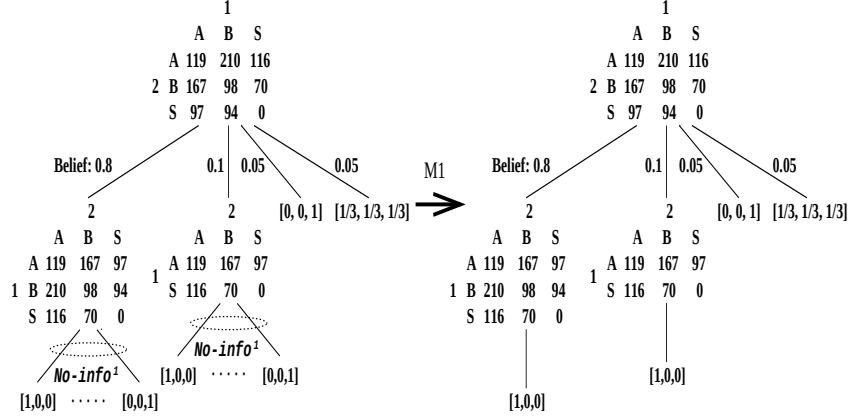


Fig. 7. DT pragmatics of $M1$.

To compute the value of communication according to Equation (2), one has to solve both modeling structures in Figure 7 and compare results. Before communication (the left part of Figure 7), Battery2 computes that if Battery1 is fully operational then the probability distribution over Battery1's actions A , B , and S is $[0.45, 0.55, 0.0]$, and if Battery1 has only long range interceptors the probability distribution over Battery1's actions A , B , and S is $[1, 0, 0]$. These results are obtained using logic sampling, as described in [12]. Combining these with the other distributions on the second level as

$$0.8 \times [0.45, 0.55, 0.0] + 0.1 \times [1, 0, 0] + 0.05 \times [0, 0, 1] + 0.05 \times [1/3, 1/3, 1/3] \\ = [0.48, 0.46, 0.06],$$

one can easily compute that Battery2's best option is to shoot at missile A , with the expected utility, $U_p(A)$, of 160.7:

$$U_p(A) = 0.48 \times 119 + 0.46 \times 210 + 0.06 \times 116 = 160.7.$$

After sending the message, $M1$ (the right part of Figure 7), the action of Battery2 becomes known, and its behavior is represented by the probability distribution $[1, 0, 0]$. Then, if Battery1 is fully operational it will choose to shoot at missile B , i.e., the probability distribution over Battery1's actions becomes $[0, 1, 0]$. If Battery2 has only long range interceptors it will intercept missile A . These probability distributions are combined with the model of Battery2 being incapacitated and with the fourth No-Information model:

$$0.8 \times [0, 1, 0] + 0.1 \times [1, 0, 0] + 0.05 \times [0, 0, 1] + 0.05 \times [1/3, 1/3, 1/3] \\ = [0.12, 0.82, 0.06].$$

The resulting distribution is Battery2's overall expectation of Battery1's actions, given all of the remaining uncertainty. Propagating these results into Level 1, the combined probability distribution describing Battery1's actions is used to compute the expected utility of Battery2's action of shooting A . We have:

Similar computations show that the message $M2$ is damaging in this case, that is, it has a negative expected utility. In other words, it is a bad idea for Battery2 to commit itself to shooting down missile B in this scenario. Thus, if the communication is expensive or the communication channel is limited, Battery2 should give the priority to the message $M1$ since it has the highest expected utility in this case.

4 Experiments

In our implementation, the anti-air defense simulator with communication was written in Common LISP and built on top of the MICE simulator [3]. Our experiments were intended to compare the performance achieved by RMM team with that of human team, and to test how RMM team can coordinate with human team without pre-defined protocols that human would have to learn for effective interactions. To demonstrate a baseline performance in the experiment, we also use a random strategy as one of target selection methods. The random team randomly intercepts targets without any reasoning.

In the experiments 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. As we mentioned, apart from payoff matrices, the agent's beliefs about environment and other agent's abilities are maintained in a general purpose knowledge base. 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. The positions of defense units were fixed and those of missiles were randomly generated. Each of two defense units was assumed to be equipped with three interceptors, if they were not incapacitated. Thus, they could launch one interceptor at a given state, and did it three times during a course of one defense episode.

We set up 100 scenarios for RMM team and 20 scenarios for human team. As we mentioned, we allowed for one-way communication at a time between defense units. Thus, 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 will be allowed to talk. The listener was silent and could only receive messages. As human subjects, we used 20 CSE and EE graduate students. Each of them was presented with the scenarios, and was given a description of what was known and what was uncertain in each case. The students were then asked to indicate which of the 11 messages was the most appropriate in each case.

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.”

To evaluate the quality of the agents’ performance, the results were expressed in terms of (1) the number of selected targets, i.e., targets the defense units attempted to intercept, and (2) 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 $\{(1 - P(H)) \times \text{warhead_size}\}$ to the total damage. If a missile was not targeted, it contributed all of its warhead size value to the damage.

4.1 Communication by RMM and Human

Among the various episodes we ran we will consider two illustrative examples, depicted in Figure 9, to examine the coordination achieved by RMM and human team in more detail. In these examples each defense unit was fully functional and has both long and short range interceptors. The warhead sizes of missiles were included in Figure 9.

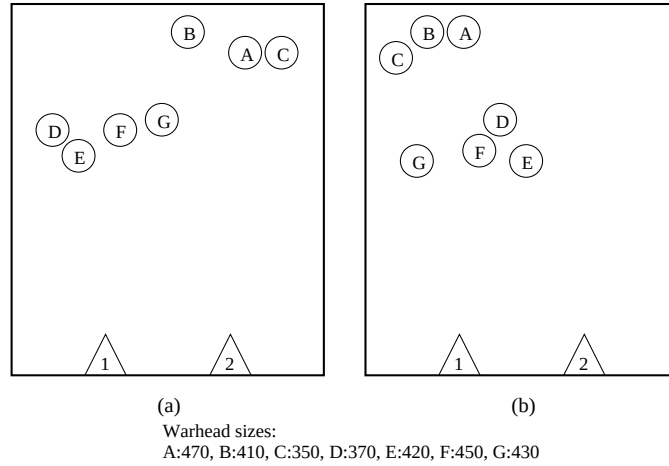


Fig. 9. Two scenarios for the comparison of performances achieved by RMM and human team.

The result of interaction in scenario (a) in Figure 9 is presented in Table 1. Without communication, the RMM batteries 1 and 2 shot at threat *F* and *A*, respectively, during the first salvo; at *E* and *B*, respectively, during the second salvo; and at *D* and *C*, respectively, during the third and final salvo, as depicted in Table 1. The total damage sustained by the RMM team in this encounter was

721.8. The choices made by a human team without communication is similarly displayed in the upper right corner of Table 1; the damage suffered by the human team was 731.2. For comparison, we include the optimal firing sequences and its resulting damage, determined by an external super-agent with all of the information about the status of multi-agent environment. The lower portion of

Table 1. Performance analysis for scenario (a).

Scenario (a)	Optimal Firing Sequences	
Targets	$F, A; E, B; G, C$	
Total damage	626.1	
Scenario (a)	No Communication	
Agent	RMM	Human
Targets	$F, A; E, B; D, C$	$D, B; E, A; F, C$
Total damage	721.8	731.2
Scenario (a)	Communication	
Agent	RMM	Human
Message 1	has(B2,both_amm0)	intercept(B2,A)
Message 2	intercept(B2,B)	intercept(B2,C)
Message 3	intercept(B1,E)	intercept(B1,G)
Targets	$F, A; G, B; E, C$	$F, A; E, C; G, B$
Total damage	630.0	647.2

Table 1 illustrates what happened when the agents could exchange messages. Before the first salvo Battery2 was allowed to talk. The RMM agent in charge of Battery2 sent the message was “I have both long and short range interceptors,” and shot at target A . Upon receiving the message, Battery1 controlled by RMM intercepted F . In case of the human team, Battery2’s initial message was “I will intercept target A ”, and the human team shot at targets F and A during the first salvo. The messages exchanged and the firings in the following two salvos are also shown in the lower portion of Table 1. As expected, the performance with communication was better than one without communication for both teams; the RMM suffered damages of 630.0, while the human team scored 647.2.

As we mentioned, the difference of total damage in RMM and human teams with and without communication shows the benefit of communication. In this scenario, the expected utilities of communicative acts executed by the RMM team were $U(\text{has}(B2, \text{both_ammo})) = 18.04$, $U(\text{intercept}(B2, B)) = 41.05$, and $U(\text{intercept}(B1, E)) = 32.49$, which sums up to 92.03. This amount is closely related to the benefit of the communication, i.e., $91.8 (= 721.8 - 630.0)$. That shows that, in this scenario, the expected utilities of the messages transmitted were an adequate estimate of the benefits of communication actually realized. As the results show, the human team’s score was very similar to that of RMM team, but humans chose different communicative acts. For example, the message

chosen by the human player before the first salvo “I will intercept target A ” has the expected utility of 12.17. This message is, therefore, useful, although slightly suboptimal from the decision-theoretic point of view. In scenario (b)

Table 2. Performance analysis for scenario (b).

Scenario (b)	Optimal Firing Sequences	
Targets	$A, E; G, F; B, D$	
Total damage	663.9	
Scenario (b)	No Communication	
Agent	RMM	Human
Targets	$F, E; B, D; G, G$	$A, E; F, F; G, D$
Total damage	1020.3	1001.8
Scenario (b)	Communication	
Agent	RMM	Human
Message 1	intercept($B2, E$)	has($B2, both_ammo$)
Message 2	intercept($B1, G$)	intercept($B1, B$)
Message 3	intercept($B1, B$)	intercept($B1, G$)
Targets	$A, E; G, F; B, D$	$A, E; B, F; G, D$
Total damage	663.9	708.4

(see Table 2) the targets are clustered in front of Battery1, unlike in scenario (a) in which targets are scattered. In this case, communication is more critical because it is more likely that targets could be intercepted redundantly without communication, resulting in greater overall damage incurred. Intuitively, as most targets head toward Battery1, the target that Battery2 selects as the biggest threat is likely to be also the most threatening to Battery1. As shown in Table 2, when communication is available, redundant target selection was prevented and the total expected damages were drastically reduced. In this scenario the sum of the expected utilities of the three messages sent by the RMM team is 246.57, while the benefits of communication actually obtained is 356.4. The result of RMM team was optimal, as described in Table 2.

4.2 Performance Assessments

The summary of all of the experimental runs is shown in Table 3. Table 3 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 perform 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 3, both defense units

Table 3. The performances of RMM and human team

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

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

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 the probability of 0.95. Case II included communication between fully functional defense units. ANOVA test showed 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.

5 Conclusions and Future Work

We implemented a probabilistic frame-based knowledge base to represent the uncertain information agents have about the environment and about the other agent present, and presented the implementation and evaluation of the decision-theoretic message selection used by automated agents coordinating in an anti-air defense domain. We measured the increase in performance achieved by rational communicative behavior in RMM team, and compared it to the performance of the human-controlled defense batteries. The results are intuitive: as expected, communication improves the coordinated performance achieved by the teams. It is interesting to see the differences between the communicative behaviors exhibited by RMM and human agents. While human communicative behaviors were often similar to those selected by the RMM agents, there are telling differences that, in our experimental runs, allowed the RMM team to achieve a slightly better performance.

Apart from these differences, we found (as also reported in [14]) that in cases when the RMM calculations show one message that is clearly the best among all of the others, our human subjects were more likely to choose this message themselves.

The results above show that the decision-theoretic message selection exhibits psychological plausibility since it is likely to agree with human choice of communicative behavior. Further, the decision-theoretic calculations allow for a meaningful prediction of the benefits of communication actually obtained as the result of the interaction. In our current work we are extending our approach to rational coordination and communicative behavior to more complex situations, such as one depicted in Figure 1. In such scenarios we will address the need to make our algorithms flexible under time pressure.

References

1. M. Barbuceanu. Coordinating agents by role based social constraints and conversation plans. In *Proceedings of the 14th National Conference on Artificial Intelligence*, pages 16–21, Providence, Rhode Island, July 1997. AAAI Press/MIT Press.
2. 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.
3. E. H. Durfee and T. A. Montgomery. MICE: A flexible testbed for intelligent coordination experiments. In *Proceedings of the 1989 Distributed AI Workshop*, pages 25–40, Sept. 1989.
4. P. J. Gmytrasiewicz and E. H. Durfee. A rigorous, operational formalization of recursive modeling. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 125–132, Menlo Park, 1995. AAAI Press/The MIT Press.
5. P. J. Gmytrasiewicz and E. H. Durfee. Rational interaction in multiagent environments: Communication. *Submitted for publication*, 1997. Available in postscript from <http://www-cse.uta.edu/~piotr/piotr.html>.

6. P. J. Gmytrasiewicz, E. H. Durfee, and D. K. Wehe. The utility of communication in coordinating intelligent agents. In *Proceedings of the 9th National Conference on Artificial Intelligence*, pages 166–172, July 1991.
7. P. J. Gmytrasiewicz, S. Noh, and T. Kellogg. Bayesian update of recursive agent models. *An International Journal of User Modeling and User-Adapted Interaction*, 8(1/2):49–69, 1998.
8. B. J. Grosz and C. Sidner. Plans for discourse. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*. MIT Press, 1990.
9. R. L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. John Wiley and Sons, New York, 1976.
10. D. Koller and A. Pfeffer. Probabilistic frame-based systems. In *Proceedings of the 15th National Conference on Artificial Intelligence*, pages 580–587, Madison, Wisconsin, July 1998.
11. 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*, Nov. 1994.
12. S. Noh and P. J. Gmytrasiewicz. Multiagent coordination in anti-air defense: A case study. In M. Boman and W. V. de Velde, editors, *Multi-Agent Rationality – MAAMAW'97 Workshop, Lecture Notes in Artificial Intelligence Vol. 1237*, pages 4–16. Springer, New York, May 1997.
13. S. Noh and P. J. Gmytrasiewicz. Coordination and belief update in a distributed anti-air environment. In *Proceedings of the 31st Hawai'i International Conference on System Sciences for the Uncertainty Modeling and Approximate Reasoning minitrack*, volume V, pages 142–151, Hawai'i, Jan 1998. IEEE Computer Society.
14. S. Noh and P. J. Gmytrasiewicz. Rational communicative behavior in anti-air defense. In *Proceedings of the Third International Conference on Multi Agent Systems*, pages 214–221, Paris, France, July 1998.
15. S. Noh and P. J. Gmytrasiewicz. Implementation and evaluation of rational communicative behavior in coordinated defense. In *Proceedings of the Third International Conference on Autonomous Agents*, pages 123–130, Seattle, Washington, May 1999.
16. J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman, San Mateo, California, 1988.
17. A. Quintero. Multi-agent system protocol language specification. Technical report, Universidad de los Andes, Colombia, 1996.
18. J. 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, Arizona, Feb. 1988.
19. S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, Englewood Cliffs, New Jersey, 1995.
20. I. A. Smith and P. R. Cohen. Toward a semantics for an agent communications language based on speech-acts. In *Proceedings of the 13th National Conference on Artificial Intelligence*, pages 24–31, Portland, Oregon, Aug. 1996.
21. M. Tambe. Agent architectures for flexible, practical teamwork. In *Proceedings of the 14th National Conference on Artificial Intelligence*, pages 22–28, Providence, Rhode Island, 1997. AAAI Press/MIT Press.