Applied Artificial Intelligence, 14:867–879, 2000

0883-9514/00 $12.00 +.00

USING BAYESIAN NETWORKS TO MODEL AGENT RELATIONSHIPS

BIKRAMJIT BANERJEE, ANISH BISWAS, MANISHA MUNDHE, SANDIP DEBNATH and SANDIP SEN
Mathematical and Computer Science Department, University of Tulsa, Tulsa, Oklahoma, U.S.A.

An agent-society of the future is envisioned to be as complex as a human society. Just like human societies, such multiagent systems (MAS) deserve an in-depth study of the dynamics, relationships, and interactions of the constituent agents. An agent in a MAS may have only approximate a priori estimates of the trustworthiness of another agent. But it can learn from interactions with other agents, resulting in more accurate models of these agents and their dependencies together with the influences of other environmental factors. Such models are proposed to be represented as Bayesian or belief networks. An objective mechanism is presented to enable an agent elicit crucial information from the environment regarding the true nature of the other agents. This mechanism allows the modeling agent to choose actions that will produce guaranteed minimal improvement of the model accuracy. The working of the proposed maximin entropy procedure is demonstrated in a multiagent scenario.

Multiagent systems (MAS) may consist of self-interested agents with individual goals. Agents in a MAS often have limited, specialized capabilities and have to depend on other agents to achieve their goals. An agent is usually embedded in a complex, dynamic, and uncertain environment teeming with scores of others, some of whom may be past and/or potential interactors. Each agent may be driven by a plethora of objectives, though its resultant behavior can be interpreted in the context of a single rational goal of maximizing utility. In the absence of any well-established code-of-conduct for agent relationships, or enforcement of behavioral norms, agents can often find it lucrative to exploit another agent to maximize local utility, whenever the situation permits. Given such a hostile environment, it becomes crucial for an agent to know whom to trust. The definition of trust, according to

This work has been supported, in part, by an NSF CAREER Award IIS-9702672.
Address correspondence to Sandip Sen, Department of Mathematical and Computer Sciences, University of Tulsa, 600 South College Avenue, Tulsa, OK, 74104-3189. E-mail: sandip@kolkata.mcs.utulsa.edu

Gambetta (1990) stresses that it is fundamentally a belief or estimation. Castelfranchi and Falcone (1998) extend this definition to include the notion of competence along with predictability.

One way of identifying trustworthiness of other agents is by developing and deploying mechanisms to model other agents. The goal is to predict the behavior of other agents. Building detailed, up-to-date, and accurate models, however, is time-consuming and a potential detractor from actual problem-solving. The model-building process has three components:

- adopting as a priori or initial model
- engaging or observing the agent in informative interactions
- updating the initial model based on such interactions.

Each of these components involve significant time and computational cost commitments on the part of the modeling agent. The key is to estimate the true nature of other agents in as few interactions as possible.

Recently, agent modeling has received increasing attention from MAS researchers. Several probabilistic mechanisms have been developed to model agents (Sandholm & Crites, 1995; Zeng & Sycara, 1997). Some of these models have been used to explore opponents' strategies (Carmel & Markovitch, 1998). An agent using such a mechanism models others' strategies, which in turn, enables it to choose actions to maximize its payoff. Very little work, however, exists on explicitly choosing actions that aid in the model-building process. It is the plan to investigate mechanisms that will allow the modeling agent to choose actions to elicit maximal information from another agent about the latter's trustworthiness. This should provide vital information in dealing with the other agents. The use of bayesian networks is proposed to capture the relationships among the agent dispositions and their actions. The modeling agent will use its observations in tandem with its model to update its belief about other agents.

Some of the agent actions may be such that they can extract more information about other agents, though not necessarily producing the highest immediate returns. On the other hand, there may be some other actions that are of immediate benefit to the agent but tell little about the other agents. Depending on how significant and time-constrained the work at hand is, the agent will have to trade off progress in problem-solving with updating its model of other agents.

To illustrate this tradeoff, a demonstrative example scenario will be used. Consider a situation where an agent A needs some documents that agent B has in its possession. A can either directly ask B to give the document to A, or can ask B's boss to instruct B to give the document to A. The first action of A will definitely provide more information about B's dispositions, depending on whether B obliges or not. On the other hand, the second action of A may not reveal B's actual cooperativeness, because there is an extra level of uncertainty introduced due to the mediation by B's boss. Hence, if B helps, it may be under coercion, whereas if it does not help, it may be because the boss forgot to entertain A's request and A may never know that. However, the second action may be more likely to satisfy A's immediate goal. If A has to choose between these two actions, it has to tradeoff between the likely immediate gain by choosing the second action, and long-term gain from the information extracted from B by virtue of the first action. In this case, such additional knowledge is exclusive of high immediate reward, while in some other cases it may be a side effect of the selected action.

In this paper, the focus is only on how to discover other agents' nature (in the sense of trustworthiness) and problem-solving for utility maximization is not considered. Bayesian networks can be used to model action.

BAYESIAN NETWORKS

A bayesian network (Jenson, 1996; Charniak, 1991) is a graphical method of representing relationships, i.e., dependencies and interdependencies among different variables that together define a model of a real-world situation. Technically, it is a directed acyclic graph (DAG) with nodes being the variables and each directed edge representing a dependence between two of them. In addition to its structure, a bayesian network is also specified by a set of parameters θ that quantify the network.

Consider a vector X of variables and an instantiation-vector x (that assigns a value x_i to each variable X_i in X from its domain D_i). If the immediate parents of a variable X_i is the vector Π_{X_i}, with its instantiation π_{x_i}, then

$$\Pr[X = x \mid \theta] = \prod_i \Pr[X_i = x_i \mid \Pi_{X_i} = \pi_{x_i}, \theta].$$

When the instantiation is clear from the context, the above is also written as

$$\Pr[X \mid \theta] = \prod_i \Pr[X_i \mid \Pi_{X_i}, \theta].$$

This defines the joint distribution of the variables in X, where each variable X_i is conditionally independent of its nondescendents given its parents or conditioning variables. Bayesian networks are useful in inference from belief structures and observations. For this purpose, an extension of Bayesian networks called influence diagrams is actually considered, which incorporate action and decision nodes besides modeling beliefs. Bayesian networks are used for representing belief structures, for the following major reasons:

- Bayesian networks can readily handle incomplete data sets. This is because bayesian networks offer a way to encode the correlations among the input variables.
- Bayesian networks allow one to learn about causal relationships. This is useful to gain an understanding about a problem domain. In addition, it allows to make predictions in the presence of interventions.
- Bayesian networks in conjunction with bayesian statistical techniques facilitate the combination of domain knowledge and data.
- Bayesian networks in conjunction with bayesian methods offers an efficient and principled approach to avoiding overfitting of data.
- Bayesian networks offer a method of updating the belief or the probability of occurrence of the particular event for the given causes.

An example bayesian network for a negotiation scenario (Banerjee et al., 1999) is shown in Figure 1 to illustrate how agents can use such a network to model others. In this particular example, A wants to sell its car to B, and as the negotiation for the price progresses, A updates its model regarding the factors influencing B's decision. In this paper, a similar modeling approach has been assumed, albeit for decision making.

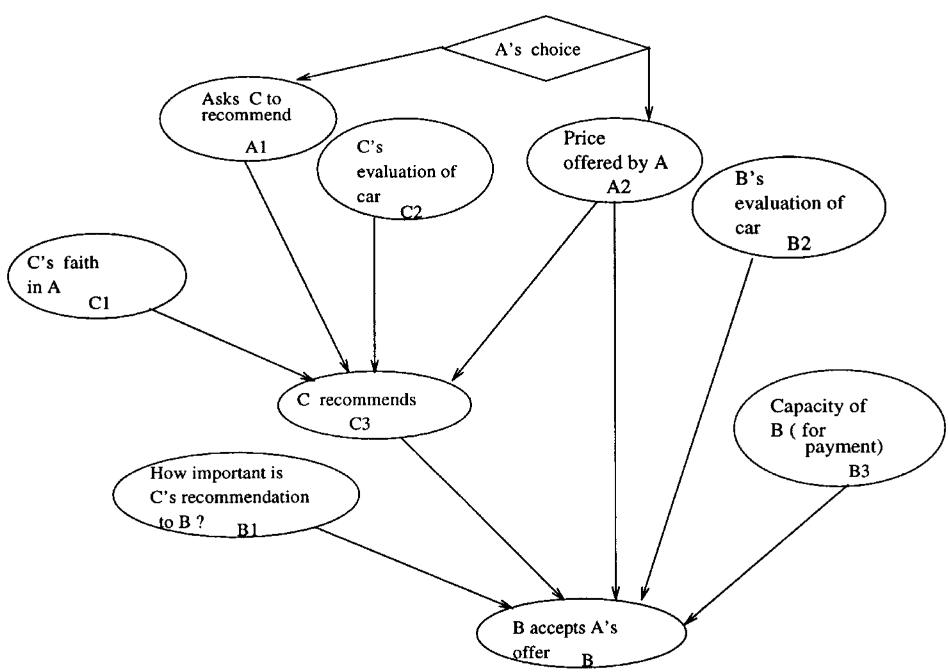


FIGURE 1. An example negotiation-scenario to illustrate the use of bayesian networks in modeling others.

CHOOSING ACTIONS TO IMPROVE MODELS OF OTHERS

The actions of agents in a multiagent environment can reveal their strategies to others. In most domains, agents are strongly coupled in the sense that the actions of an agent can influence the utility of other agents. In an open environment, a self-interested agent should be aware of the nature, dispositions, and priorities of other agents. Such knowledge can enable an agent to better plan its actions. Hence, in addition to performing its problem-solving tasks, an agents should try to elicit accurate knowledge about agents who can affect its utility and with which it frequently interacts. Actions chosen from a particular subset of available actions, or a particular order of the same action set, may reveal more information about the true nature of the agents more effectively. Our goal is to develop a mechanism for selecting the actions for the modeler so as to form better estimates of the nature of the others.

The basic approach of eliciting information from or about an agent is as follows. It is contended that often there are actions that give out more information about an agent's strategies than other actions. From the modeling agent's viewpoint, one wants to recognize the scenarios or contexts that will result in the other agents' choosing actions that reveal more information about their trustworthiness. The modeling agent should then, by its own actions, create the corresponding contexts as often as feasible, and to the extent that these do not significantly detract from its regular problem-solving activities. One visualizes an information content in each action of an agent and defines it as below.

Definition 1. Suppose an agent A has n available action $\{ai\}_1^n$ represented as nodes in the modeling bayesian network. These need not be distinct nodes, but can be the different values of the same node. One considers the subset of parent-nodes of an action node a_i denoted by $\wp(a_i)$ (i.e., $\wp(a_i) \subseteq \Pi_{ai}$) which model the dispositions of an agent like trustworthiness, cooperativeness, etc. Then information content E_i in action a_i of agent A is given by

$$E_i = \sum_{\wp(ai)} \Pr[\wp(a_i) \mid a_i] * \log_2 (\Pr[\wp(a_i) \mid a_i]).$$

One notes that this quantity lies between 0 and -1, (see Figure 2) with the minimum occurring at maximum uncertainty regarding the possible events. This corresponds to the situation which, in one's view, provides minimum information about the nature of A that B would like to know to improve his model.

The model of agent interaction that one considers is a two-level game (Luce & Raiffa, 1957), where the modeler (B) has to choose from a set of possible actions $\{b_i\}_1^m$, which lead to the other agent (A) adopting from its

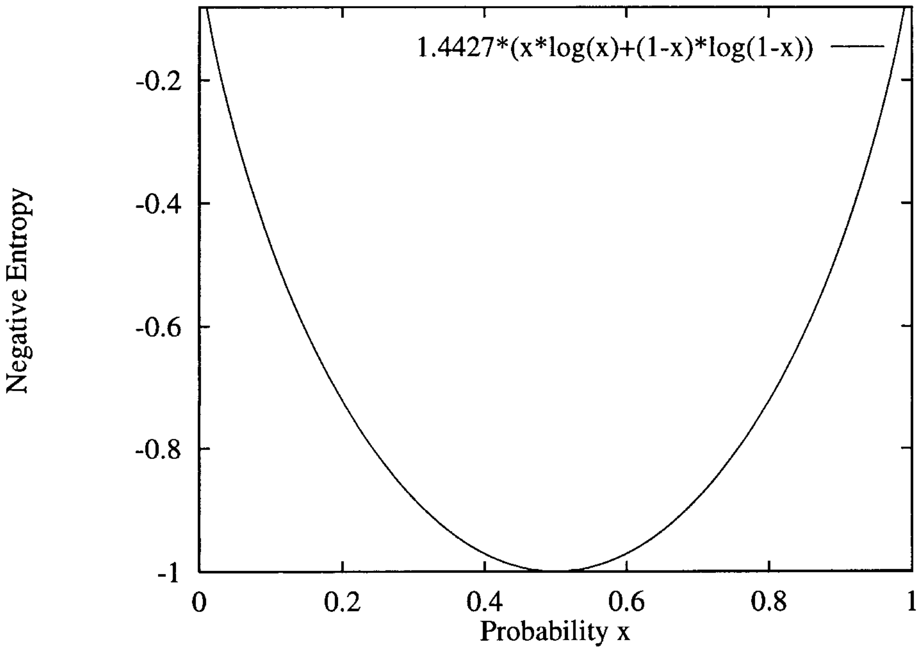


FIGURE 2. Plot of negative entropy-function, where the case that the probability of all options are equal (here two options) corresponds to minimum information content.

own set of actions. The jth action of agent A in response to the ith action of agent B, b_i, is denoted as a_{ij}, to be chosen from the set $\{a_{ij}\}_{j=1}^{j=n}$. The agent B models the factors that influence A's action-choice including its own actions, as a bayesian network. The trustworthiness of A is one of the critical factors that guides A's response to B's actions. A maximin (Luce & Raiffa, 1957) mechanism is presented that allows B to select actions that help it to form increasingly better estimates of A's trustworthiness, given its response to B's actions.

The set of actions available to A in response to each action of B are known to B and the latter has prior estimates of the probabilities of factors affecting each such action of A. Among these, $\wp(a_{ij})$ represents those parents of the action node a_{ij} that reflect the nature of A. Now given the prior probabilities of such factors, B computes the information content of A's action a_{ij} as

$$E_{ij} = \sum_{\wp(aij)} \Pr[\wp(a_{ij}) \mid b_i a_{ij}] * \log_2 (\Pr[\wp(a_{ij}) \mid b_i a_{ij}]),$$

where $\Pr[\wp(a_{ij}) \mid b_i a_{ij}])$ is computed according to Bayes rule as

$$\Pr[\wp(a_{ij}) \mid b_i a_{ij}]) = \frac{\Pr[a_{ij} \mid b_i \wp(a_{ij})] * \Pr[\wp(a_{ij} \mid b_i]}{\Pr[a_{ij} \mid b_i]}$$

and since, in general $\wp(a_{ij})$ are all independent of b_i,

$$\Pr[\wp(a_{ij}) \mid b_i a_{ij}]) = \frac{\Pr[a_{ij} \mid b_i \wp(a_{ij})] * \Pr[\wp(a_{ij})]}{\Pr[a_{ij} \mid b_i]}.$$

E_{ij} can also be looked upon as a measure of the difference between the prior and posterior probabilities of $\wp(a_{ij})$. Now, B's goal is to find the action b_i that has the maximum value for minimum information content across all of A's responses to the action b_i, i.e., B wants to maximize the minimum guarantee regarding the information obtained from A's response to b_i. To this end, B first computes the lower bound on extractable information associated with action b_i as

$$\varepsilon_i = \min_j \{E_{ij}\}.$$

Last, B selects the action b_i that maximizes this lower bound as

$$b_i : i = \arg\max_k \varepsilon_k.$$

If the prior probabilities are inaccurate, then with progressive interaction, the modeler improves its estimates of the nature (currently under consideration) of the agent being modeled, choosing actions such that convergence is achieved as rapidly as possible. Finally, when the prior and posteriors converge, the modeler moves on to explore some other traits of the other, following the same process all over again. In addition to arbitrating between conflicting actions, this procedure also suggests a choice among unrelated actions, as is demonstrated in the following example.

MODELING SCENARIO INVOLVING AGENT TRUSTWORTHINESS

Now the use of the above-mentioned procedure is illustrated with a typical agent-interaction scenario. An example is described where agent B has to select action for elicitation of maximal information about agent A's nature. In this case, one considers an agent trustworthy if it responds positively to one's request for help. A negative response (refusal to help) will decrease trust in that agent, in the absence of any defendable reason, e.g., that the agent was busy in something more crucial to its utility. This approach is, however, not limited to a particular definition of trust and can be used for other definitions as well. One only needs a characterization of the sequence of actions according to the definition adopted. In this example (see Figure 3), agent B has to choose from the following set of actions:

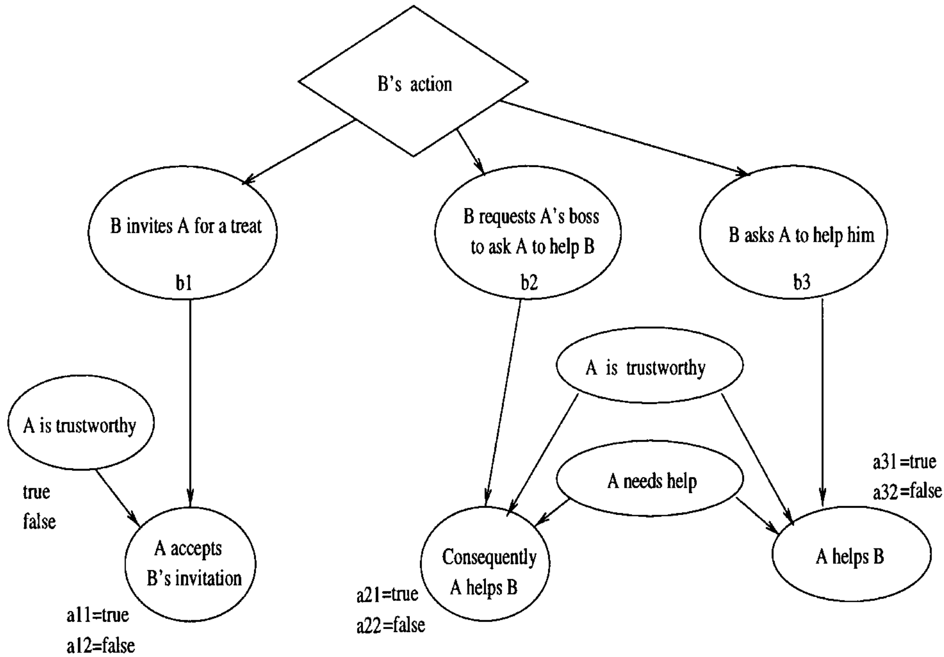


FIGURE 3. Bayesian network model for the example situation.

- B asks A to help him (b_3)
- B invites A for a treat (b_1)
- B requests A's boss to ask A to help B (b_2).

In this case, the other possible events capable of affecting A's actions are:

- A may need help and, hence, may offer help to others hoping for help in return (node H in Figure 3). To keep matters simple, one does not count this as one of the $\wp$ nodes for the corresponding actions in the calculations.
- Whether A accepts B's invitation to a treat depends on whether A is sociable. From Figure 3, it can be written in accordance with the notation as $\wp(a_{1*})$.[1]
- A helps others with a reasonably high probability and without any compulsion if he is trustworthy or dependable. It is assumed that this information about A's nature is of vital importance to B and so (from Figure 3) it can be written as $\wp(a_{3*})$ and also as $\wp(a_{2*})$.

One assumes that one has prior probabilities of these events (all events are assumed to be binary-valued) from domain knowledge. From these prior beliefs and conditional probabilities, one estimates the posterior beliefs of B regarding the nature of A, i.e., whether A is trustworthy or not.

Illustration of the Action-Selection Procedure

The subnetworks have been shown for each of B's available actions in Figures 4, 5, 6, including the respective conditional-probability tables. One notes that the probability values in the table of Figure 4 have lower values than corresponding elements in Figure 5 wherever the action node of B has true value. This is because of the additional uncertainties that were mentioned earlier. However, the probability values remain identical wherever the action node has false value, because the other influencing factors are common and affect A's decision alike.

Based on these probability values, B computes the posterior probability of A being trustworthy, given B selects action b_2 and A selects action a_{21} as

$$\Pr[D \mid a_{21}\, b_2] = \frac{\Pr[a_{21} \mid b_2\, D] * \Pr[D]}{\Pr[a_{21} \mid b_2]},$$

where

$$\Pr[a_{21} \mid b_2\, D] = \Pr[H] * \Pr[a_{21} \mid H b_2\, D] + \Pr[\neg H] * \Pr[a_{21} \mid \neg H b_2\, D]$$

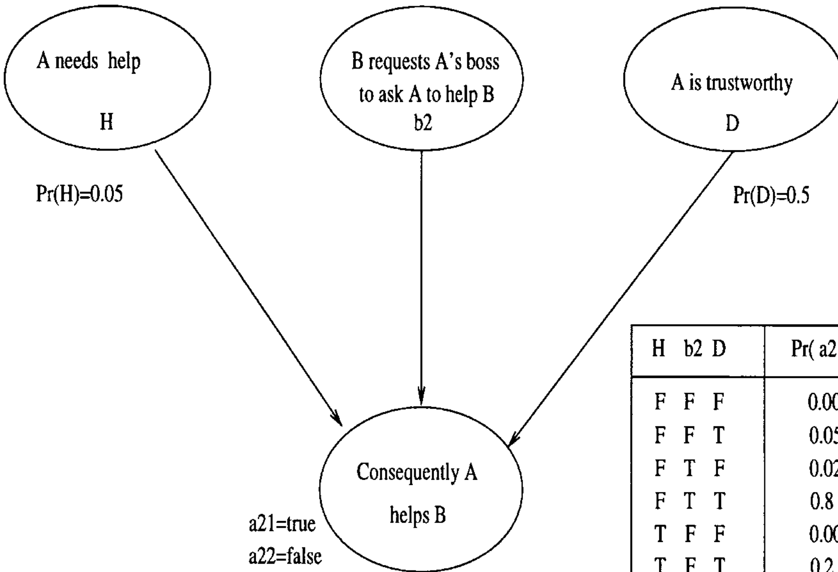
$$\begin{aligned}\Pr[a_{21} \mid b_2] = {} & \Pr[HD] * \Pr[a_{21} \mid H b_2\, D] \\ & + \Pr[H \neg D] * \Pr[a_{21} \mid H b_2 \neg D] \\ & + \Pr[\neg H D] * \Pr[a_{21} \mid \neg H b_2\, D] \\ & + \Pr[\neg H \neg D] * \Pr[a_{21} \mid \neg H b_2 \neg D].\end{aligned}$$

Consequently, one has $\Pr[D \mid a_{21}\, b_2] = 0.976$. Similarly, one calculates the following probabilities:

$$\Pr[D \mid a_{22}\, b_2] = 0.166,$$
$$\Pr[D \mid a_{31}\, b_3] = 0.9132,$$
$$\Pr[D \mid a_{32}\, b_3] = 0.1.$$

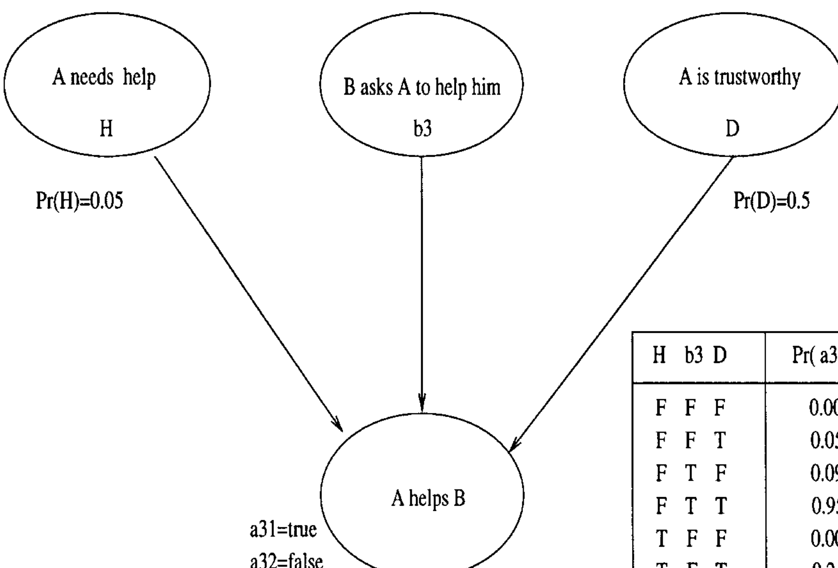
Hence, one has

$$\begin{aligned}E_{21} &= -0.1132,\\ E_{22} &= -0.6485,\\ E_{31} &= -0.4257,\\ E_{32} &= -0.469.\end{aligned}$$



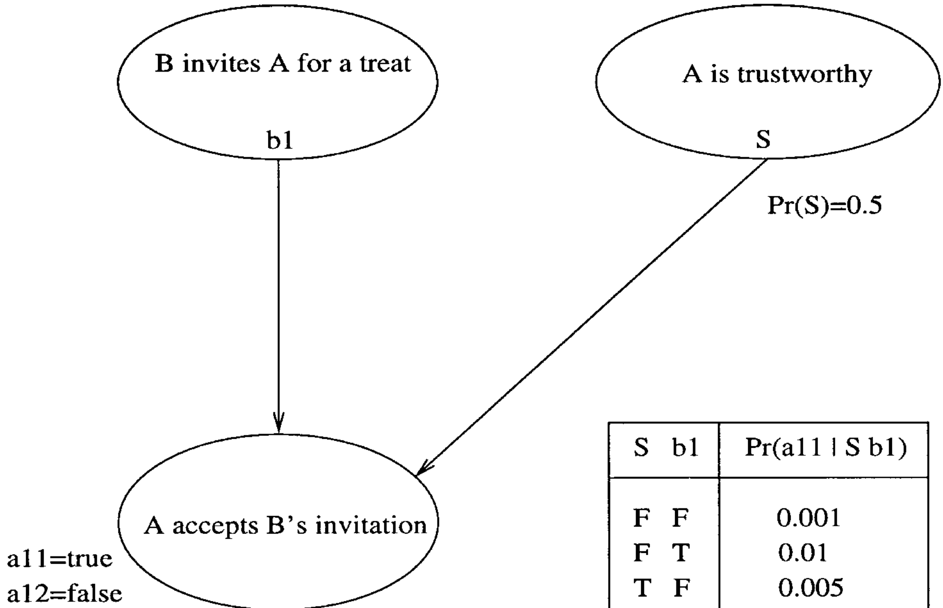
H	b2	D	Pr(a21| H b2 D)
F	F	F	0.005
F	F	T	0.05
F	T	F	0.02
F	T	T	0.8
T	F	F	0.001
T	F	T	0.2
T	T	F	0.01
T	T	T	0.9

FIGURE 4. Portion of the network of Figure 3 for action b_2 .



H	b3	D	Pr(a31| H b3 D)
F	F	F	0.005
F	F	T	0.05
F	T	F	0.09
F	T	T	0.95
T	F	F	0.001
T	F	T	0.2
T	T	F	0.1
T	T	T	0.99

FIGURE 5. Portion of the network of Figure 3 for action b_3 .



S	b1	Pr(a11 | S b1)
F	F	0.001
F	T	0.01
T	F	0.005
T	T	0.9

FIGURE 6. Portion of the network of Figure 3 for action b_1.

The information content, as expected intuitively, is higher for actions of A which allow B to update its prior Pr[D] = 0.5 by the maximum amount (either increase or a decrease), and B should choose the action that maximizes the minimal increase in this prior. One sees that

$$\varepsilon_2 = \min\{-0.1132, \ -0.6485\} = -0.6485,$$

$$\varepsilon_3 = \min\{-0.4257, \ -0.469\} = -0.469.$$

Clearly, action b_3 is preferred to b_2 for maximal updating of the prior probability Pr[D], as contended earlier. In addition, one has also considered the action b_1. It can be shown that $\Pr[S|a_{11}b_1] = 0.989$ and $\Pr[S|a_{12}b_1] = 0.1$. One has again assumed the prior probability of A being trustworthy to be 0.5.

Here, one finds that the action b_1 is the most favored among the actions available to B. With increasing exploration by B into A's trustworthiness, its estimates are going to be better. As B develops more accurate estimates of A's trustworthiness, this improved knowledge allows B to be more effective in its problem-solving activities. B can also decide to explore other aspects of A's nature once an accurate estimate of A's trustworthiness has been developed.

CONCLUSION

In this paper, a mechanism has been presented to enable bayesian networks-based modelers to select actions that lead to more accurate models about the nature of another agent. The mechanism involves the use of a maximin procedure for action selection that guarantees a minimum level of improvement in estimation of an agent's trustworthiness irrespective of whatever action the latter selects. An illustration has been provided of the working of this procedure with a running example.

The knowledge of another agent's nature may be extremely significant in guiding the modeling agent's problem-solving activities, given the open and competitive environment it is situated in. The progress in problem-solving has been ignored and focus has been solely on exploring the nature of the other agents. An expansion on this model is planned to incorporate the problem-solving criterion too, and an indication on how the tradeoff between these two metrics is to be achieved for action selection. This will provide a unified framework by which exploratory actions are incorporated as an integral part of routine problem-solving for achieving the goal of maximizing long-term utility. Work on multiple-level decision-making is also planned where a multilevel tree structure is generated for each action available to an agent.

The maximin action-selection method is conservative in nature. To guarantee a certain improvement in model estimate it can ignore large improvements. This approach is completely justified if the other agent knows that the modeler is trying to improve its model, and is then deliberately trying to take actions to minimize such increases. When such an assumption is untenable, the modeler can choose the action that produces the maximum average improvement. An interesting avenue would be to experimentally evaluate the relative effectiveness of the maximin and average metrics to select actions.

NOTE

1 a_{i*} stands for the node that can take values a_{ij} $\forall j$.

REFERENCES

AI Magazine. Summer 1999. Special Issue on Bayesian Techniques, 20(2).

Banerjee, B., S. Debnath, and S. Sen. 1999. Using bayesian networks to aid negotiations among agents. In the working notes of *AAAI'99*, *Workshop on Negotiation: Settling Conflicts and Identifying Opportunities* (also available as AAAI Technical Report WS-99-12), pp. 44–49, 1999.

Castelfranchi, C., and R. Falcone. 1998. Principles of trust for MAS: Cognitive autonomy, social importance, and quantification. In *Proceedings of the Third International Conference on Multiagent Systems*, 72–79, Los Alamitos, CA, IEEE Computer Society.

Charniak, E., Winter 1991. Bayesian networks without tears. *AI Magazine* 12(4):50–63.

Carmel, D., and S. Markovitch. 1998. How to explore your opponent's strategy (almost) optimally. In *Proceedings of the Third International Conference on Multiagent Systems*, 64–71, Los Alamitos, CA, IEEE Computer Society.

Gambetta, D. 1990. *Trust*. Oxford: Basil Blackwell.

Jensen, F. V. 1996. *An introduction to bayesian networks*. New York: Springer-Verlag.

Luce, R. D., and H. Raiffa. 1957. *Games and decisions: Introduction and critical survey*. New York: Dover.

Russell, S., and P. Norvig. 1995. *Artificial intelligence: A modern approach*. Englewood Cliffs, NJ: Prentice Hall.

Shachter, R. 1994. Evaluating influence diagrams. *Operations Research* 34(36):871–882.

Sandholm, T. W., and R. H. Crites. 1995. Multiagent reinforcement learning and iterated prisoner's dilemma. *Biosystems Journal* 37:147–166.

Zeng, D., and K. Sycara. 1997. Benefits of learning in negotiation. In *Proceedings of the 14th National Conference on Artificial Intelligence*, 36–41, Menlo Park, CA, AAAI Press/MIT Press.