Using Bayesian Networks to Model Agent Relationships

Bikramjit Banerjee, Anish Biswas, Manisha Mundhe, Sandip Debnath & Sandip Sen

Mathematical & Computer Sciences Department

University of Tulsa

600 South College Avenue

Tulsa, OK 74104, USA

{bikram, abiswas, manisha, debnath, sandip}@euler.mcs.utulsa.edu

Abbreviated title: Modeling Agent Relationships

Abstract

We envision an agent-society of the future to be as complex as a human society. Just like human societies, such multiagent systems (MAS) deserve an in-depth study of the dynamics, the relationships and the 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. We propose to represent such models as Bayesian or belief networks. We present an objective mechanism 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. We demonstrate the working of our proposed maximin entropy procedure in a multiagent scenario.

1 Introduction

Multiagent systems (MAS) may consist of self-interested agents with individual goals. Agents in an 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 situation permits. Given such a hostile environment, it becomes crucial for an agent to know whom to trust. The definition of trust, according to Gambetta [Gam90] stresses that it is fundamentally a *belief* or an *estimation*. Castelfranchi and Falcone [CF98] 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 an *a priori* or initial model.
- Engaging or observing the other 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 [SC95, ZS97]. Some of these models have been used to explore opponents' strategies [CM98]. 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. We 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. We propose to use Bayesian Networks 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 it little about the other agents. Depending on how significant and time-constrained the work at hand is, the agent will have to tradeoff progress in problem-solving with updating its model of other agents.

To illustrate this tradeoff, we will use a demonstrative example scenario. 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 we focus only on how to discover other agents' nature (in the sense of trustworthiness), and do not consider problem-solving for utility-maximization. Bayesian Networks can be used to model act

2 Bayesian Networks

A Bayesian Network [Jen96, Cha91] is a graphical method of representing relationships, i.e dependencies and independencies 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 $\mathbf{X}$ of variables and an instantiation-vector $\mathbf{x}$(that assigns a value x_i to each variable X_i in $\mathbf{X}$ 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

$$\mathbf{Pr}[\mathbf{X} = \mathbf{x}|\theta] = \prod_i \mathbf{Pr}[X_i = x_i|\mathbf{\Pi}_{X_i} = \pi_{x_i}, \theta].$$

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

$$\mathbf{Pr}[\mathbf{X}|\theta] = \prod_i \mathbf{Pr}[X_i|\mathbf{\Pi}_{X_i}, \theta].$$

This defines the joint distribution of the variables in $\mathbf{X}$, where each variable X_i is conditionally independent of its non-descendents given its parents, or conditioning variables. Bayesian Networks are useful in inference from belief-structures and observations [Cha91, AI 99]. For our purpose, we actually consider an extension of Bayesian Networks called *influence diagrams* [R.S94, RN95], that incorporate action and decision nodes besides modeling beliefs. We use Bayesian Networks 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.

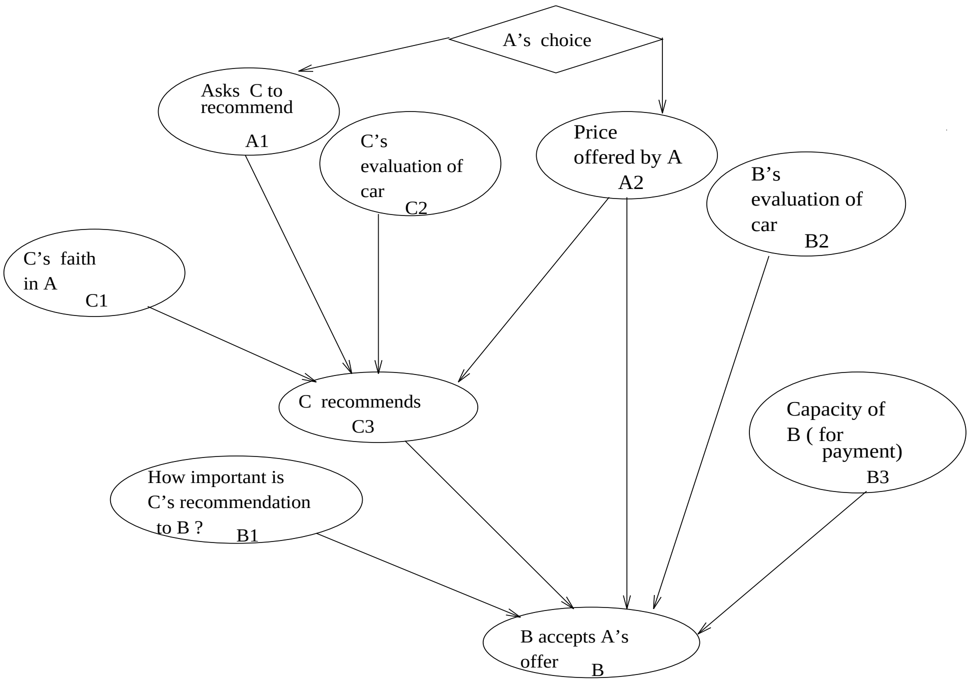


Figure 1: An example negotiation-scenario to illustrate the use of Bayesian Networks in modeling others.

- Bayesian networks allow one to learn about causal relationships. This is useful to gain 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 over fitting 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 [BSS99] 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, we have assumed a similar modeling approach, albeit for decision-making.

3 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 agent 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.

Our basic approach of eliciting information from or about an agent is as follows. We contend that often, there are actions that give out more information about an agent's strategies than other actions. From the modeling agent's viewpoint, we want 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. We visualize an *information content* in each action of an agent and define it as below.

Definition 1 *Suppose an agent A has n available actions $\{a_i\}_1^n$ represented as nodes in the modeling Bayesian Network. These need not be distinct nodes, but can be the different values of the same node. We consider the subset of parent-nodes of an action node a_i denoted*

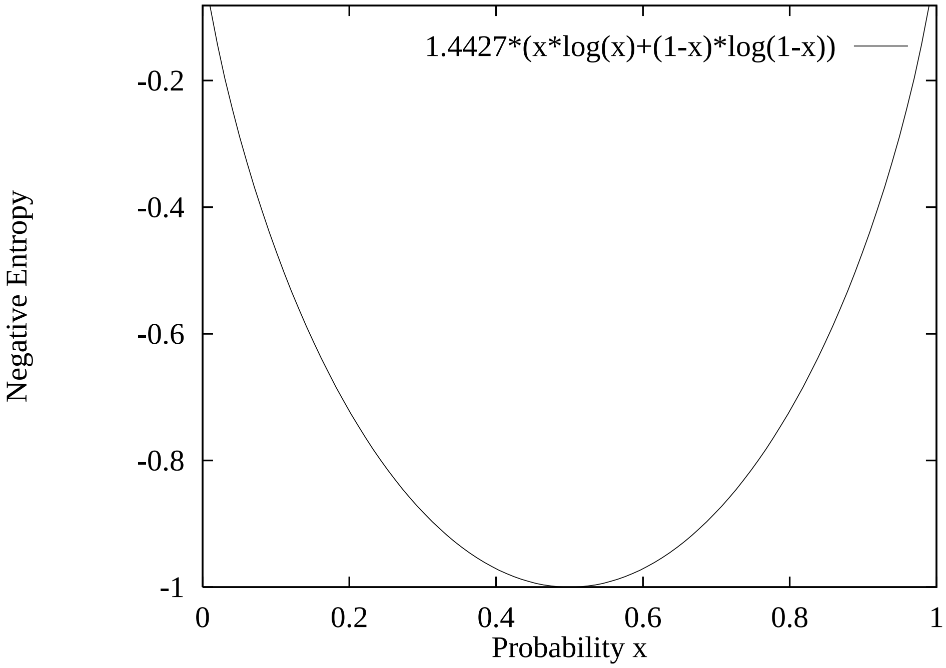


Figure 2: Plot of negative entropy-function, where the case that the probability of all options are equal (here 2 options) corresponds to minimum information content.

by $\wp(a_i)$ (i.e. $\wp(a_i) \subseteq \mathbf{\Pi_{a_i}}$) that 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*

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

We note 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 our 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 we consider is a two level game [LR57], 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 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 crucial factors that guides A's response to B's actions. We present a *maximin* [LR57] mechanism that allows B to select actions that help it to form increasingly better estimate 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

$$E_{ij} = \sum_{\wp(a_{ij})} \mathbf{Pr}[\wp(a_{ij})|b_i a_{ij}] * \log_2(\mathbf{Pr}[\wp(a_{ij})|b_i a_{ij}])$$

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

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

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

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

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

$$\mathcal{E}_i = min_j\{E_{ij}\}.$$

Lastly B selects the action b_i that maximizes this lower bound as

$$b_i : i = \arg\max_k \ \mathcal{E}_k.$$

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,

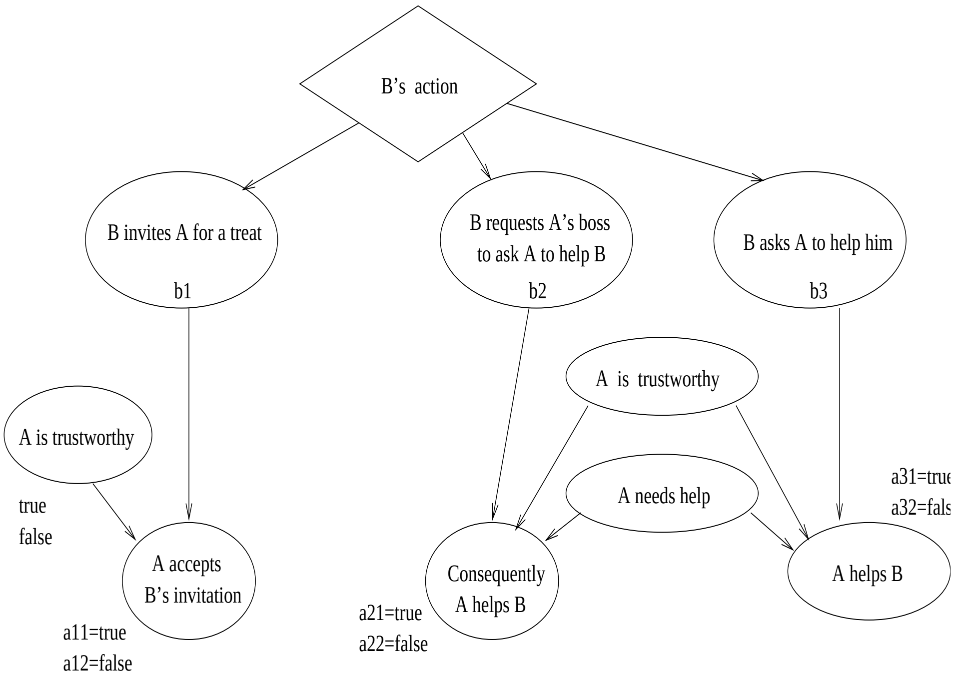


Figure 3: Bayesian Network model for the example situation.

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 we demonstrate in the following example.

4 Modeling scenario involving agent trustworthiness

Now we illustrate the use of the above-mentioned procedure with a typical agent-interaction scenario. We describe an example where agent B has to select action for elicitation of maximal information about agent A's nature. In this case we consider 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. Our approach is, however, not limited to a particular definition of trust and can be used for other definitions as well. We only need 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

- 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, we do not count this as one of the $\wp$ nodes for the corresponding actions in our calculations.
- Whether A accepts B's invitation to a treat depends on whether A is sociable. From fig 3, it can be written in accordance with our notation as $\wp(a_{1*})$. [1]
- A helps others with a reasonably high probability and without any compulsion if he is trustworthy or dependable. We assume 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 $\wp(a_{2*})$.

We assume that we have prior probabilities of these events (all events are assumed to be binary-valued) from domain knowledge. From these prior beliefs and conditional probabilities, we estimate the posterior beliefs of B regarding the nature of A i.e. whether A is trustworthy or not.

4.1 Illustration of the action-selection procedure

We have shown the subnetworks for each of B's available actions in figures 6, 4, 5, including the respective conditional-probability tables. We note that the probability-values in the

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

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

$$\mathbf{Pr}[D|a_{21}b_2] = \frac{\mathbf{Pr}[a_{21}|b_2D] * \mathbf{Pr}[D]}{\mathbf{Pr}[a_{21}|b_2]}$$

where

$$\mathbf{Pr}[a_{21}|b_2D] = \mathbf{Pr}[H] * \mathbf{Pr}[a_{21}|Hb_2D] + \mathbf{Pr}[\neg H] * \mathbf{Pr}[a_{21}|\neg Hb_2D]$$

$$\begin{aligned}\mathbf{Pr}[a_{21}|b_2] \quad = \quad & \mathbf{Pr}[HD] * \mathbf{Pr}[a_{21}|Hb_2D] \\ & +\mathbf{Pr}[H\neg D] * \mathbf{Pr}[a_{21}|Hb_2\neg D] \\ & +\mathbf{Pr}[\neg HD] * \mathbf{Pr}[a_{21}|\neg Hb_2D] \\ & +\mathbf{Pr}[\neg H\neg D] * \mathbf{Pr}[a_{21}|\neg Hb_2\neg D]\end{aligned}$$

Consequently, we have $\mathbf{Pr}[D|a_{21}b_2] = 0.976$. Similarly, we calculate the following probabilities

$$\mathbf{Pr}[D|a_{22}b_2] = 0.166,$$

$$\mathbf{Pr}[D|a_{31}b_3] = 0.9132,$$

$$\mathbf{Pr}[D|a_{32}b_3] = 0.1.$$

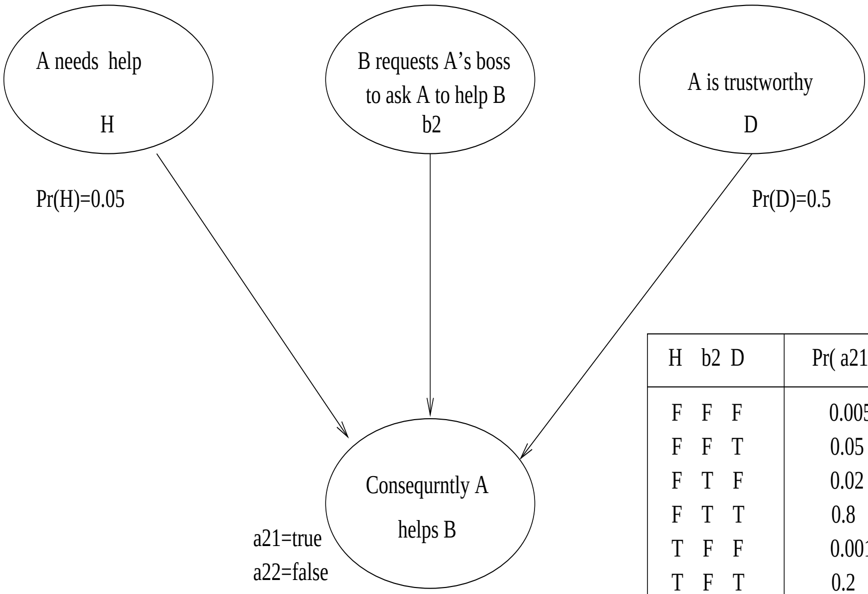
Hence, we have

$$E_{21} = -0.1132,$$

$$E_{22} = -0.6485,$$

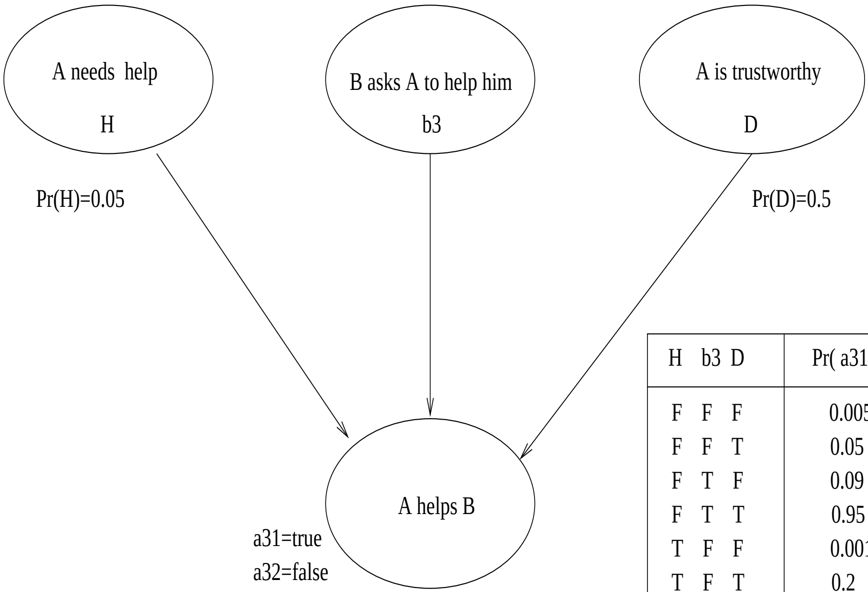
$$E_{31} = -0.4257,$$

$$E_{32} = -0.469.$$



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.

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

$$\mathcal{E}_2 = min\{-0.1132, -0.6485\} = -0.6485,$$

$$\mathcal{E}_3 = min\{-0.4257, -0.469\} = -0.469.$$

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

Here, we find 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 estimate 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 accurate estimate of A's trustworthiness has been developed.

5 Conclusion

In this paper, we have presented a mechanism 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. We have provided an illustration 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. We have ignored the progress in problem-solving and focused solely on exploring the nature of the other agents. We plan to expand on this model to incorporate the

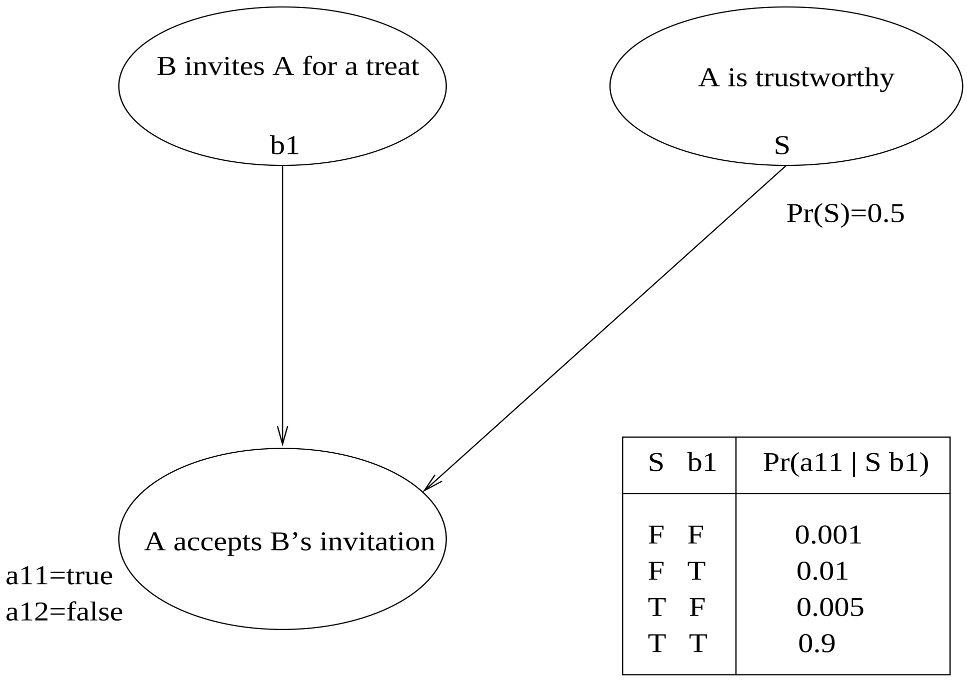


Figure 6: Portion of the Network of figure 3 for action b_1.

problem-solving criterion too, and indicate 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. We also plan to work on multiple-level decision making where a multi-level 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 knew that the modeler was trying to improve its model, and was 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 the average metrics to select actions.

Acknowledgements: This work has been supported, in part, by an NSF CAREER Award: IIS-9702672.

References

[AI 99] AI Magazine, summer 1999, volume 20, number 2, 1999. Special Issue on Bayesian Techniques.

[BSS99] B.Banerjee, S.Debnath, and S.Sen. Using bayesian networks to aid negotiations among agents. Workshop report, AAAI'99, 1999.

[CF98] Cristiano Castelfranchi and Rino Falcone. Principles of trust for MAS: Cognitive autonomy, social importance, and quantification. In *Proceedings of the Third International Conference on Multiagent Systems*, pages 72–79, Los Alamitos, CA, 1998. IEEE Computer Society.

[Cha91] Eugene Charniak. Bayesian networks without tears. *AI Magazine*, 12(4):50–63, Winter 1991.

[CM98] David Carmel and Shaul Markovitch. How to explore your opponent's stratey (almost) optimally. In *Proceedings of the Third International Conference on Multiagent Systems*, pages 64–71, Los Alamitos, CA, 1998. IEEE Computer Society.

[Gam90] D. Gambetta. *Trust.* Basil Blackwell, Oxford, 1990.

[Jen96] Finn V. Jensen. *An Introduction to Bayesian Networks.* Springer Verlag, New York, NY, 1996.

[LR57] R. Duncan Luce and Howard Raiffa. *Games and Decisions: Introduction and Critical Survey.* Dover, New York, NY, 1957.

[RN95] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach.* Prentice Hall, 1995.

[R.S94] R.Shachter. Evaluating influence diagrams. *Operations Research*, 34(36):871 – 882, 1994.

[SC95] Tuomas W. Sandholm and Robert H. Crites. Multiagent reinforcement learning and iterated prisoner's dilemma. *Biosystems Journal*, 37:147–166, 1995.

[ZS97] Dajun Zeng and Katia Sycara. Benefits of learning in negotiation. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, pages 36–41, Menlo Park, CA, 1997. AAAI Press/MIT Press.