Bayesian Learning in Negotiation [*†]

Dajun Zeng   Katia Sycara

The Robotics Institute
Carnegie Mellon University, Pittsburgh, PA 15213
Voice: (412) 268-8815   Fax: (412) 268-5569
zeng+@cs.cmu.edu   katia@cs.cmu.edu

Introduction

Recent growing interest in autonomous interacting software agents and their potential application in areas such as electronic commerce (Sandolm & Lesser 1995) has given increased importance to automated negotiation. Much DAI and game theoretic research (Rosenschein & Zlotkin 1994; Osborne & Rubinstein 1994) deals with coordination and negotiation issues by giving pre-computed solutions to specific problems. There has been much research reported on developing theoretical models in which learning plays an eminent role, especially in the area of adaptive dynamics of games (e.g., (Jordan 1992; Kalai & Lehrer 1993)). However, to build autonomous agents that improve their negotiation competence based on learning from their interactions with other agents is still an emerging area.

We are interested in developing autonomous agents capable of reasoning based on experience and improving their negotiation behavior incrementally. Learning in negotiation is closely coupled with the issue of how to model the overall negotiation process, i.e., what negotiation protocols are adopted. Standard game-theoretic models (Osborne & Rubinstein 1994) tend to focus on *outcomes* of negotiation in contrast to the *negotiation process* itself. DAI research (Rosenschein & Zlotkin 1994) emphasizes special protocols articulating compromises while trying to minimize the potential interactions or communications of the involved agents. Since we are motivated by a different set of research issues, such as including effective learning mechanisms in the negotiation process, we adopt a different modeling framework, i.e., a sequential decision making paradigm (Bertsekas 1995; Cyert & DeGroot 1987).

The basic characteristics of a sequential decision making model are: (1) there is a sequence of decision making points (different stages) which are dependent on each other, and (2) the decision maker has a chance to update his knowledge after implementing the decision made at a certain stage and receiving feedback so that he can make a more informed decision at the next stage. The following observations support our choice of sequential decision making as the baseline negotiation model. First, most negotiation tasks involve multiple rounds of exchanging proposals and counter-proposals. A sequential decision making framework provides readily available constructs to model the iterative nature of inter-agent interactions. Second, negotiating agents indeed receive feedback after they offer a proposal or a counter-proposal in the form of replies from the recipient agent(s). Third, a sequential decision making framework supports an open world approach. An agent doesn't need to have a complete world model at the outset of negotiation. Whenever new information comes in, irrespective of whether the agent learns the new knowledge by itself or some other information sources become available, it can make use of the newly acquired knowledge at the next decision making point. The agent can handle in the same manner more difficult situations where agents not only don't have complete information, but also the environment and other agents might be constantly changing. Last but not least, learning can take place naturally in a sequential decision making framework. This type of on-line incremental learning behavior is highly desirable in an automated negotiation program.

In this paper, we propose such a sequential decision making model, called Bazaar, which is able to learn. We address multi-agent learning issues in Bazaar by explicitly modeling beliefs about the negotiation environment and the participating agents under a probabilistic framework using a Bayesian learning representation and updating mechanism. Our ultimate research goal is to develop an adaptive negotiation model capable of exhibiting a rich set of negotiation behaviors with

modest computational effort.

A Survey of Existing Negotiation Models

Traditional single-agent decision making models typically assume that the decision maker has complete knowledge of (1) his own preference ordering or utility function, and (2) the probabilities associated with the various outcomes. When multiple agents are involved, such as in negotiation, the introduction of *strategic* interaction, however, complicates this picture. In making his decision, the rational individual must take into account the probable choices of others, whose choices are in turn contingent upon his own. This leads to the well known *outguessing regress* (Young 1975) where no accurate prediction or confident expectation about the individual choices can be produced. Therefore, the central theme of all negotiation models is avoiding this *dilemma* involved in strategic interaction.

In order to circumvent the *outguessing regress* of strategic interactions, game theoretic models make the following restrictive assumptions: [1] (1) Both the number of players and their identity are assumed to be fixed and known to everyone. (2) All the players are assumed to be fully rational, and each player knows that the others are rational (common knowledge). Each player's set of alternatives is fixed and known. (3) Each player's risk-taking attitude and expected-utility calculations are also fixed and known to each and every individual involved in decision making. These assumptions limit the applicability of game theoretic frameworks for solving realistic problems. Another important limitation of game theoretic models is that these models are fundamentally static in the sense that they primarily focus on *outcomes* in contrast to negotiation *processes*. The search for determinate rational decisions within the framework of game theory has not led to a general model governing rational choice in interdependent situations. Instead, it has produced a number of special models applicable to specific types of interdependent decision making. For instance, the most celebrated solution concept, the von Neumann-Morgenstern solution, is based on the fact that in a two-person, zero-sum game an outguessing regress can be avoided by assuming (not unrealistically) that one player knows that his opponent will "do his worst",

whatever strategy he himself selects. Analyses of the N-person cooperative game circumvent the difficulties associated with strategic interaction in a different way by introducing detailed decision rules concerning such things as the relative *power* of the players (e.g., the Shapley value and the Nash solution).

Some game theorists (e.g., (Harsanyi & Selten 1972)) have sought to achieve determinate solutions for nonzero-sum games by introducing the notion that each player may be able to assign *subjective probabilities* to the choices of the other participant. In other words, it is possible to suppose that each individual proceeds in some subjective fashion to estimate the probable choices of the other player. In essence, the individual acquires information in the process so that his choice problem reduces to a situation that is fundamentally analogous to a game against nature as in a traditional single-agent decision making situation. We view this line of research as more closely coupled with sequential decision making view of negotiation rather than orthogonal game models.

To a large extent, these theoretical models are not concerned with computational issues, i.e., how to deal with inevitable practical complexities that don't have proper analytic representations and therefore haven't found their way into the models. Some of the AI models, in this sense, can be understood as bridges between applications and abstract theoretical models. Playing games (e.g., chess, go) has been one of the major foci of AI. For certain games, game theory is able to provide a theoretically sound mathematical solution and winning strategy. The existence of the solution, however, doesn't guarantee that the player can find the solution. AI models and programs help the players locate an approximate solution strategy according to bounded rationality principles by utilizing heuristic search, heuristic evaluation, and learning techniques. Along with the emergence and development of DAI techniques, there has been increasing interest in using AI methodology and frameworks in negotiation modeling. (Sycara 1990) enriched the negotiation model by integrating AI planning, Case-Based Reasoning and other decision theoretic techniques. Multi-agent resource allocation as a special case of negotiation has been extensively explored by (Kraus & Subrahmanian 1995), in which logic framework and time constraints are taken into consideration within the traditional framework of game theory. Some recent work (e.g., (Sen & Sekaran 1995; Sandolm & Lesser 1995)) in the context of Distributed AI addresses multiagent learning issues in various settings. Our work differs from others by explicitly modeling negotiation as a sequential decision making task and using Bayesian updating as the underlying learn-

---

[1]It should be noted that some of the very recent game theoretic models are directly motivated by considerations of dropping or relaxing some of these assumptions. Although there has been interesting progress reported in the literature (e.g., (Jordan 1992)), the fundamental framework and methodology of game theory remains almost the same and it might be too early to tell whether these new results will reshape the current game theoretic framework.

ing mechanism.

Sequential Decision Making with Rational Learning

Our overall research goal is to develop a computational model of negotiation that can handle multi-agent learning and other complicated issues (e.g., multi-issue multi-criteria negotiation) that don't have straightforward and computationally efficient analytic models. We believe that a useful computational model of negotiation should exhibit the following characteristics: (1) The model should support a concise yet effective way to represent negotiation context. (2) The model should be prescriptive in nature. (3) The computational resources required for finding reasonable suggestions/solutions should be moderate, sometimes at the cost of compromising the rigor of the model and the optimality of solutions. (4) The model should provide means to model the dynamics of negotiation. (5) The model should also support the learning capability of participating agents.

Motivated by these desirable features, we have developed Bazaar, a sequential decision making negotiation model that is capable of learning. We describe how the proposed model works in a simple negotiation scenario for illustrative purposes before we present the formal description of Bazaar.

Suppose two computer programs are negotiating on behalf of their users in a supply chain management scenario. Agent 1 is the producer (supplier)'s agent and Agent 2 is the buyer's agent. These two agents are involved in a negotiation process where a detailed contract concerning product mix, delivery date, price, etc., is expected to be achieved. The overall negotiation process can be modeled as exchanging proposals and counter-proposals, as typically happens in human negotiations.

Let us first view the negotiation from the supplier, i.e., Agent 1's point of view. We ignore the problem associated with locating potential buyers and assume that the existence of Agent 2 is known to Agent 1. We also assume a communication channel between Agent 1 and Agent 2 is readily available. At the outset, Agent 1 needs to come up with a solution package detailing its offer with respect to product, price, delivery date, quality, etc. How to determine the particular value of these variables depends on the following factors: (1) Agent 1's own cost and profit structure and evaluation, (2) Agent 1's understanding of the current economic situation and potential demand for its product, (3) Agent 1's model of Agent 2, and (4) Agent 1's expectation from Agent 2, such as potentially profitable future transactions.

Considering all these factors and the trade-offs among them, Agent 1 calculates the expected payoff value function associated with possible offers, and selects the offer that maximizes his payoff. Agent 2 receives the offer transmitted by Agent 1. To decide whether to accept this offer or to counter-propose, Agent 2 essentially uses a similar evaluation procedure as Agent 1.

The next step would be easy if Agent 2 decides to accept the offer. In that case, Agent 2 just needs to send an acceptance message to Agent 1, which finalizes the contract. If Agent 2 is not satisfied with the offer, it can either abort the negotiation or send back a counter-proposal. Again, the process of determining the counter-proposal is similar to that used by Agent 1 to determine the initial proposal. First, Agent 2 calculates the payoff function whose domain is all feasible offers. Then, the offer that maximizes Agent 2's payoff is selected. It should be noted that the fact that Agent 1 has sent a proposal does have an impact on the decision making process that Agent 2 goes through when deliberating his counter-proposal, since Agent 2's internal knowledge of Agent 1 and possibly the knowledge about the supply situation have been updated. Agent 1's proposal affects Agent 2's decision in a quite indirect way by causing changes in Agent 2's perception of Agent 1.

After Agent 1 receives the counter-proposal offered by Agent 2, Agent 1 first updates its model of Agent 2, then evaluates the offer in the light of newer knowledge. If it is deemed as an acceptable offer, the negotiation process is brought to an end. Otherwise, Agent 1 sends a counter-proposal based on its payoff structure and newer knowledge about its counterpart, Agent 2. Exchanges of proposals and counter-proposals will go on until one of the agents decides to accept an offer or to quit. The negotiation process can also end because of other external events such as missing an agreement deadline, etc.

Bazaar: a formal description

▶ In Bazaar, a negotiation process can be modeled by a 10-tuple $< N, M, \Delta, A, H, Q, \Omega, P, C, E >$, where,

A-1 A set N (the set of players).

A-2 A set M (the set of issues/dimensions covered in negotiation. For instance, in the supply chain management domain, this set could include product price, product quality, payment method, transportation method, etc.)

A-3 A set of vectors $\Delta \equiv \{(D_j)_{j \in M}\}$ (a set of vectors whose elements describe each and every dimension of an agreement under negotiation).

A set A composed of all the possible actions that can be taken by every member of the players set.

▷ $A \equiv \Delta \cup \{Accept, Quit\}$

A-4 For each player $i \in N$ a set of possible agreements A_i.

▷ For each $i \in N$, $A_i \subset A$.

A-5 A set H of sequences (finite or infinite) that satisfies the following properties:

▷ The elements of each sequence are defined over A.

▷ The empty sequence Φ is a member of H.

▷ If $(a^k)_{k=1,...,K} \in H$ and $L < K$ then $(a^k)_{k=1,...,L} \in H$.

▷ If $(a^k)_{k=1,...,K} \in H$ and $a^K \in \{Accept, Quit\}$ then $a^k \overline{\in} \{Accept, Quit\}$ when $k = 1,..., K - 1$.

Each member of H is a *history*; each component of a history is an action taken by a player. A history $(a^k)_{k=1,...,K}$ is terminal if there is no a_{K+1} such that $(a^k)_{k=1,...,K+1} \in H$. The set of terminal histories is denoted by Z.

A-6 A function Q that associates each nonterminal history $(h \in H \setminus Z)$ to a member of N. (Q is called the *player function* which determines the orderings of agent responses.)

A-7 A set of Ω of relevant information entities. Ω is introduced to represent the players' knowledge and belief about the following aspects of negotiation:

▷ The parameters of the environment, which can change over time. For example, in supply chain management, global economic or industry-wide indices such as overall product supply and demand and interest rate, belong to Ω.

▷ Beliefs about other players. These beliefs can be approximately decomposed into three categories:

(a) Beliefs about the factual aspects of other agents, such as how their payoff functions are structured, how many resources they have, etc.

(b) Beliefs about the decision making process of other agents. For example, what would be other players' reservation prices.

(c) Beliefs about meta-level issues such as the overall *negotiation style* of other players. Are they tough or compliant? How would they perceive a certain action? What about their risk-taking attitudes? etc.

A-8 For each nonterminal history h and each player $i \in N$, a subjective probability distribution $P_{h,i}$ defined over Ω. This distribution is a concise representation of the knowledge held by each player in each stage of negotiation.

A-9 For each player $i \in N$, each nonterminal history $H \setminus Z$, and each action $a_i \in A_i$, there is an im-

plementation cost $C_{i,h,a}$. C can be interpreted as communication costs or costs associated with time caused by delaying terminal action (*Accept* or *Quit*).

A-10 For each player $i \in N$ a *preference* relation $\succeq_i$ on Z and $P_{h,i}$ for each $h \in Z$. $\succeq_i$ in turn results in an evaluation function $E_i(Z, P_{Z,i})$.

Characteristics of Bazaar

Most game-theoretic models assume that the player has infinite reasoning and computation capacity. On the one hand, this infinite rationality assumption eliminates some of the theoretical problems (e.g., the precise definition of degree of rationality is unknown) associated with modeling agents with bounded rationality; on the other hand, it is just because of assuming infinite smartness of players that outguessing regress becomes a problem, since every participating agent tries to model others in a recursive fashion (e.g., (Gmytrasiewicz & Durfee 1992)). The fact that the agents don't have infinite reasoning capacity imposes natural termination for otherwise endless outguessing regress. This is precisely the foundation of Bazaar. Bazaar ignores some aspects of the "strategic" part of a game by modeling other players explicitly (see A-7b) in terms of beliefs and uncertainty. This, along with its learning capability, differentiates Bazaar from other negotiation models. Other observations about Bazaar are:

- Bazaar aims at modeling multi-issue negotiation processes. By incorporating multiple dimensions into the action space, Bazaar is able to provide an expressive language to describe the relationships between these issues and possible trade-offs among them.

- Bazaar supports an *open* world model. Any change in the outside environment, if relevant and perceived by a player, will have an impact on the player's subsequent decision making processes. This feature is highly desirable and is seldom found in other negotiation models.

- In most of existing negotiation models, learning issues have been either simply ignored or oversimplified for theoretical convenience. Multi-agent learning issues can be addressed in Bazaar and conveniently supported by the iterative nature of sequential decision making and the explicit representation of beliefs about other agents (the distribution P over Ω).

Learning in Negotiation

The importance of learning in negotiation has been recently recognized in the game research community

as fundamental for understanding human behavior as well as for developing new solution concepts (Osborne & Rubinstein 1994; Jordan 1992; Kalai & Lehrer 1993). Theoretical results (most of which are partial and preliminary), however, are available only for the simplest game settings. Multi-agent learning has also increasingly drawn research efforts from Distributed AI community (e.g., (Mor, Goldman, & Rosenschein 1995; Sen & Sekaran 1995)). In the context of Bazaar, we are using the Bayesian framework to update the knowledge and belief that each agent has about the environment and other agents. To address the computational complexity issues with Bayesian analysis, we use the Bayesian belief network representation and updating mechanism. In addition to providing efficient updating techniques, Bayesian belief networks offer an expressive modeling language and allow easy and flexible encoding of domain-specific knowledge (Pearl 1988).

In this section, we revisit the buyer-supplier example used before to demonstrate how the Bayesian framework can be utilized in a negotiation setting. For illustrative purposes, we consider the negotiation process only from the viewpoint of the buyer and assume that the relevant information set Ω is comprised of only one item: belief about the supplier's reservation price $RP_{supplier}$[2]. It is obvious that although the buyer knows his own reservation price, the precise value of $RP_{supplier}$ is unknown to him. Nevertheless, the buyer is able to update his belief (learn) about $RP_{supplier}$ based on his interactions with the supplier and on his domain knowledge. As a result of learning, the buyer is expected to gain more accurate expectation of the supplier's payoff structure and therefore make more advantageous offers. In this example, we show how the buyer's belief about $RP_{supplier}$ can be updated during negotiation.

The buyer's partial belief about $RP_{supplier}$ can be represented by a set of hypotheses $H_i, i = 1, 2, \ldots, n$. For instance, H_1 can be "$RP_{supplier} = \$100.00$"; H_2 "$RP_{supplier} = \$130.00$". A priori knowledge held by the buyer can be summarized as probabilistic evaluation over the set of hypotheses $\{H_i\}$ (e.g., $P(H_1) = 0.2$, $P(H_2) = 0.35$, ...). The Bayesian updating occurs when the buyer receives new signals from the out-

[2]An agent's reservation price is the agent's threshold of offer acceptability. Typically a reservation price is private to each agent, and is different for each agent for each negotiation issue. For example, a supplier reservation price is the price such that the supplier agent will not accept an offer below this price; a buyer's reservation price is the price such that the buyer will not accept an offer above this price. It is obvious that if one agent knows other agents' reservation price, its proposals can be more well-targeted and result in a potentially more beneficial final outcome.

side environment or from the supplier. Along with domain-specific knowledge, these new signals enable the buyer to acquire new insights about $RP_{supplier}$ in the form of posterior subjective evaluation over H_i. In our case, the offers and counter-offers (Offer$_{supplier}$) from the supplier comprise the incoming signal; while the domain knowledge can be an observation such as: "Usually in our business people will offer a price which is above their reservation price by 17%", which can be represented by a set of conditional statements of similar form, one of which is shown as follows: $P(e_2 \mid H_2) = 0.95$, where e_2 represents "Offer$_{supplier} = \$152.1$", and H_2 "$RP_{supplier} = \$130.00$".

Given the encoded domain knowledge in the form of conditional statements and the signal (e) in the form of offers made by the supplier, the buyer can use the standard Bayesian updating rule to revise his belief about $RP_{supplier}$:

$$P(H_i \mid e) = \frac{P(H_i)P(e \mid H_i)}{\sum_{k=1}^{n} P(e \mid H_k)P(H_k)}$$

We use a numerical example to show how this updating works. For simplicity, we suppose that the buyer knows that the supplier's reservation price is either $100.00 or $130.00. In other words, the buyer has only two hypotheses: H_1: "$RP_{supplier} = \$100.00$" and H_2: "$RP_{supplier} = \$130.00$".

At the beginning of the negotiation, the buyer doesn't have any other additional information. His a priori knowledge can be summarized as: $P(H_1) = 0.5, P(H_2) = 0.5$.

In addition, we suppose that the buyer is aware of "Suppliers will typically offer a price which is above their reservation price by 17%", part of which is encoded as: $P(e_1 \mid H_1) = 0.95$ and $P(e_1 \mid H_2) = 0.75$, where e_1 denotes the event that the supplier asks $117.00 for the goods under negotiation.

Now suppose that the supplier offers $117.00 for the product the buyer wants to purchase. Given this signal and the domain knowledge, the buyer can calculate the posterior estimation of $RP_{supplier}$ as follows:

$$
\begin{aligned}
P(H_1 \mid e_1) &= \frac{P(H_1)P(e_1 \mid H_1)}{P(H_1)P(e_1 \mid H_1) + P(H_2)P(e_1 \mid H_2)} \\
&= 55.9\% \\
P(H_2 \mid e_1) &= \frac{P(H_2)P(e_1 \mid H_2)}{P(H_2)P(e_1 \mid H_1) + P(H_2)P(e_1 \mid H_2)} \\
&= 44.1\%
\end{aligned}
$$

Suppose that the buyer adopts a simple negotiation strategy: "Propose a price which is 10% below the estimated $RP_{supplier}$". Prior to receiving the supplier's

offer (\$117.00), the buyer would propose \$115.00 (the mean of the $RP_{supplier}$ subjective distribution). After receiving the offer from the supplier and updating his belief about $RP_{supplier}$, the buyer will propose \$113.23 instead. Since the new offer is calculated based on a more accurate estimation of the supplier's utility structure, it might result in a potentially more beneficial final outcome for the buyer and may also help both sides reach the agreement more efficiently.

Some observations about this example are as follows: (1) Parameters contained in domain knowledge such as the estimated percentage of the supplier's offer over his reservation price (17%) can be updated in a similar fashion. For instance, it is not unrealistic to suppose that this percentage will drop when the negotiation process continues. (2) The belief updating can be triggered by events such as discovery of externally available information in addition to the supplier's offers. For instance, if the buyer finds out during the negotiation the overall supply of the particular goods under negotiation is experiencing a tremendous increase, his estimation about the supplier's reservation price might drop without even receiving any new offers from the supplier. (3) In this example, we use the traditional Bayesian representation for illustrative purposes. Other efficient updating mechanisms utilizing more expressive representations such as the Bayesian network work essentially in the same way.

Concluding Remarks and Future work

In this paper, we present Bazaar, a sequential decision making model of negotiation in which multi-agent learning is an integral modeling construct. This model is motivated by providing a computational framework for negotiation which satisfies the following features: (1) complicated issues such as multi-issue, multi-criteria negotiation can be addressed, (2) the model provides an operational algorithm to guide offers instead of only prescribing the final outcome, and (3) learning can be easily incorporated into the model. Current work focuses on analyzing theoretically the impact of learning in various negotiation settings. We are also conducting empirical work on applying the Bazaar framework on nontrivial negotiation scenarios such as supply chain management.

References

Bertsekas, D. P. 1995. *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific.

Cyert, R. M., and DeGroot, M. H. 1987. *Bayesian analysis and uncertainty in economic theory*. Rowman & Littlefield.

Gmytrasiewicz, P. J., and Durfee, E. H. 1992. Logic of knowledge and belief for recursive modeling: Preliminary report. In *Proceedings of the National Conference on Artificial Intelligence*, 628–634.

Harsanyi, J. C., and Selten, R. 1972. A generalized nash solution for two-person bargaining games with incomplete information. *Management Science* 18:80–106.

Jordan, J. S. 1992. The exponential covergence of bayesian learning in normal form games. *Games and Economic Behavior* 4:202–217.

Kalai, E., and Lehrer, E. 1993. Rational learning leads to nash equilibrium. *Econometrica* 61(5):1019–1045.

Kraus, S., and Subrahmanian, V. S. 1995. Multiagent reasoning with probability, time, and beliefs. *International Journal of Intelligent Systems* 10(5):459–499.

Mor, Y.; Goldman, C. V.; and Rosenschein, J. S. 1995. Learn your opponent's strategy (in polynomial time). In *Proceedings of IJCAI-95 Workshop on Adaptation and Learning in Multiagent Systems*.

Osborne, M. J., and Rubinstein, A. 1994. *A course in game theory*. The MIT Press.

Pearl, J. 1988. *Proabilistic reasoning in intelligent systems*. Morgan Kaufmann Publishers, Inc., second edition.

Rosenschein, J., and Zlotkin, G. 1994. *Rules of Encounter*. Cambridge, Mass.: MIT Press.

Sandolm, T. W., and Lesser, V. R. 1995. Coalition fromation among bounded rational agents. In *Proceedings of 14th International Joint Conference on Artificial Intelligence*.

Sen, S., and Sekaran, M. 1995. Multiagent coordination with learning classifier systems. In *Proceedings of IJCAI-95 Workshop on Adaptation and Learning in Multiagent Systems*.

Sycara, K. 1990. Negotiation planning: An AI approach. *European Journal of Operational Research* 46:216–234.

Young, O. R. 1975. *Bargaining : formal theories of negotiation*. University of Illinois Press.