How Can An Agent Learn To Negotiate?

Dajun Zeng | Katia Sycara

The Robotics Institute

Carnegie Mellon University

Pittsburgh, PA 15213

zeng+@cs.cmu.edu | katia@cs.cmu.edu

http://www.cs.cmu.edu/~softagents/

Talk Overview

- Motivations & Research Objective
- Desiderata of A Computational Model of Negotiation
- Modeling negotiation as a Sequential Decision Making Process
- Bazaar: Sequential Decision Making + Bayesian Updating
- Supply Contracting Domain
- Learning in a Simple Buyer-Supplier Scenario
- Computational Issues
- Conclusions

Motivations for Learning in Negotiation

- Importance of automated negotiation that can tolerate incomplete information and is able to adapt according to external changes in domain such as supply contracting and electronic commerce
- Much DAI and game theoretic work provides pre-computed solutions to specific problems

Research Objective

Build autonomous agents that improve their negotiation competence based on learning from their interactions with other agents

Game Theoretic Modeling of Negotiation

- Advantages:
  - Mathematical soundness and elegance
  - Thorough analysis of *strategic* interactions
  - Explicit criteria

- Many restrictive assumptions:
  - The number of players and their identity are fixed and known to everyone
  - All the players are assumed to be *fully* rational
  - Each player's set of alternatives is fixed and known
  - Each player's risk-taking attitude and expected-utility calculations are also fixed and known
- Game Theoretic Models are fundamentally *static*
- Not historically concerned with computational issues

Desiderata of A Computational Model of Negotiation

- Support a concise yet effective way to represent negotiation context
- Be prescriptive in nature
- Be computationally efficient, sometimes at the cost of compromising the rigor of the model and the optimality of solutions.
- Model the dynamics of negotiation and Learn through interactions

Characteristics of Sequential Decision Making

- A sequence of decision making points (different stages) which are dependent on each other
- The decision maker has a chance to update his/her knowledge after implementing the decision made at a certain stage and receiving feedback

Modeling Negotiation as a SDM process

- Most negotiation tasks involve multiple rounds of exchanging proposals and counter-proposals
- Negotiating agents indeed receive feedback after they offer a proposal or a counter-proposal in the form of replies
- A sequential decision making framework supports an *open world* approach.
- Learning can take place naturally in a sequential decision making framework.

Limitations of SDM

- *Strategic* interactions only partially modeled
- Fuzzy evaluation criteria

Bazaar: Sequential Decision Making with Rational Learning

- In Bazaar, a negotiation process is 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)

A-3 A set of vectors $\Delta \equiv \{(D_j)_{j \in M}\}$

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

- $\triangleright$ For each $i \in N$, $A_i \subset A$

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

- $\triangleright$ The elements of each sequence are defined over A
- $\triangleright$ The empty sequence Φ is a member of H
- $\triangleright$ If $(a^k)_{k=1,\ldots,K} \in H$ and $L < K$ then $(a^k)_{k=1,\ldots,L} \in H$
- $\triangleright$ If $(a^k)_{k=1,\ldots,K} \in H$ and $a^K \in \{Accept, Quit\}$ then $a^k \overline{\in} \{Accept, Quit\}$ when $k = 1, \ldots, K-1$

A-6 A function Q that associates each nonterminal history ($h \in H \setminus Z$) to a member of N

A-7 A set of Ω of relevant information entities

- ▷ The parameters of the environment
- ▷ Beliefs about other players:
  - (a) Beliefs about the factual aspects of other agents
  - (b) Beliefs about the decision making process of other agents
  - (c) Beliefs about some meta-level issues such as the overall *negotiation style* of other players

A-8 For each nonterminal history h and each player $i \in N$, a subjective probability distribution $P_{h,i}$ defined over Ω

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 implementation cost $C_{i,h,a}$

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 $E_i(Z, P_{Z,i})$

- **Solution Concept:** Adaptive feedback control from Dynamic Programming

Domain: Supply Contracting

- Supply Contracting is an emerging area in Operations Management
  - Motivation: Manufacturing companies need to ensure smooth and inexpensive supply of raw material and components that are needed to produce and assemble the final product.

- Supply contracting is an ideal evaluation domain for Bazaarsince:
  - Significant in its own right
  - Quantitatively-oriented
  - Some strategic parts of supply contracting have been ignored in analytic modeling and in fact are being ignored in practice
  - Opportunity for learning: uncertainties involved in various stages of supply contracting, e.g., uncertainty in demand and supply

Learning in a Simple Buyer-Supplier Scenario

- Assumptions:
  - The relevant information set Ω has only one item: belief about the supplier's reservation price $RP_{supplier}$ (from the buyer's perspective)
  - The buyer's partial belief about $RP_{supplier}$ is represented by two hypotheses:
    * H_1 = "$RP_{supplier} = \$100.00$"
    * H_2 = "$RP_{supplier} = \$130.00$"
  - A priori knowledge: $P(H_1) = 0.5, P(H_2) = 0.5$

– Domain Knowledge: “Usually in our business people will offer a price which is above their reservation price by 17%”, part of which is encoded as:

  * $P(e_1 \mid H_1) = 0.95$
  * $P(e_1 \mid H_2) = 0.75$

  where e_1 denotes the event that the supplier asks $\$117.00$ for the goods under negotiation

– The buyer adopts a simple negotiation strategy: “Propose a price which is 10% below the estimated $RP_{supplier}$”

- Suppose that the supplier offers $\$117.00$
- Given this signal and the domain knowledge, the buyer can calculate the posterior estimation of $RP_{supplier}$ as follows:

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

- 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

Initial Theoretical Results

A player who uses the Bayesian mechanism to update his beliefs about the unknown parameters of the game and other player's strategies in a subjectively rational fashion performs at least as well as without the Bayesian learning

Computational Issues

- Efficiency
  - Bayesian Network
- Convergence, . . . ,
  - Experimental study of solution quality, time to reach an agreement, etc.

Conclusions

- “In-between” game-theoretic models and single agent decision making models
- Bazaar aims at modeling multi-issue negotiation processes
- Bazaar supports an *open* world model
- Address multi-agent learning utilizing the iterative nature of sequential decision making and the explicit representation of beliefs about other agents