A Computational Model For Online Agent Negotiation

Pu Huang **Katia Sycara**
phuang@cs.cmu.edu **katia@cs.cmu.edu**
Center for Automated Learning and Discovery
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

Agent-based on-line negotiation technology has the potential ability to radically change the way e-business is conducted. In this paper, we present a formal model for autonomous agents to negotiate on the Internet. In the basic negotiation scenario, we validate our model by showing that an agent will never make an offer that can possibly be exploited by its opponents. In our model, the negotiation process is driven by the internal beliefs of participating agents. We empirically identify the relative strength of a group of belief updating methods and show how an agent can change its behavior by adjusting some critical parameters.

1 Introduction

With the rapid growth of Internet, autonomous software agents, which can be viewed as delegates of human beings in the cyberspace, have drawn much attention in recent years because of their potential capacity to radically change the current style of practicing e-business. For example, a software agent, deployed as a delegate of its master, can shop on the Web. Once finding the commodity it is looking for, it will bargain with the owner about the price, just as a human will do. The owner, probably, is an agent as well. Compared to today's passive on-line shopping, where people themselves search the Web and finish the trade manually, we think the agent-delegated shopping will be the way of future on-line trading.

However, before the wide application of multi-agent systems in the real-world electronic commerce, several challenges, including both technical and social aspects, must be addressed [Sycara, 1998]. One of these challenges is to determine how these heterogeneous, self-interested agents should interact and negotiate with each other, given that there is no global control on the Internet. A common approach to this problem is to construct a negotiation model that guides the agents' negotiation activities.

Researchers have investigated various negotiation models from different points of view for a long time. Game theorists view a negotiation as a dynamic, incomplete information game, and try to solve the game by giving some predicted outcomes in certain conditions. Volumes of literature exist in this field. A good survey can be found in Bichler [2000]. Although successful in some problems, the game-theoretical approach is hard to extend to general problem domains, simply because the complexity and inherent uncertainty in real-world negotiation thwart accurate analysis.

In this paper, we present a computational model for on-line agent negotiation. Instead of focusing on the predicted outcomes, our model emphasizes the negotiation process itself. Specifically, we model a negotiation process as a sequential decision-making process: In every negotiation iteration, an agent checks the history of the process, updates its beliefs about its opponents and then tries to maximize its own expected payoff based on its own subjective beliefs.

When bargaining with others, a human's subjective beliefs play an important role. A buyer (she) may just believe that the seller (he) will never decrease his price offer in the next negotiation round. She may be right (because her spy told her current offer is the seller's reservation price). Or she may be wrong (because the seller intentionally leaked the false information to her spy). Whatever she is right or wrong, her mental beliefs will play a same role in the negotiation as far as the outcomes are concerned. In this paper, we conduct a series of experiments to examine the impact of different beliefs on the outcomes of a basic on-line negotiation scenario. Our approach, directly model the mental beliefs and examine their impact on the outcomes, is shared by other researchers in multiagent systems. Bazzan and Bordini [2000] studied the impact of agents' "personalities" on outcomes of the Minority Game. Sen *et al.*[2000] examined a probabilistic strategy, which can be interpreted as the agent's beliefs about other agents, used by an agent to help others in a not-perfectly-friendly environment. Gmytrasiewicz and Lisetti [2000] directly modeled an agent's "mental emotion" as a probability distribution over the possible states of the environment. Here we directly model an agent's beliefs over its opponents' action sets. The joint actions of all the agents drive the environment to shift. Every agent tries to maximize its final payoff by choosing its own optimal action based on its own subjective beliefs.

Our work is close related to the line of research on "belief-desire-intention (BDI) model", in which agents adapt themselves to the uncertain environment by using different "intention reconsideration policies". In our model, the uncertainty roots in the negotiation process itself: agents are uncertain about what actions their opponents may take and use different "belief updating methods" to interact with each other. An empirical study of different "intention reconsideration policies" can be found in [Schut and Wooldridge, 2000].

2 Formal models

Consider a game with m states played by n players in a limited time horizon $maxT$. In each time period $t = 0, 1, \ldots maxT - 1$, each player takes an action simultaneously. Driven by the joint actions taken of all the players, the game transfers to another state at time $t + 1$. For each player i the following information is associated with it.

Public information(shared with all other players):

- A_i : a finite set of all possible actions player i may take.
- T_i : a finite set of all possible types player i may be.
- $M = \{S, s_0, JAS, \Theta\}$: a transition automaton that defines the structure of the game. Where
  - S is a finite set of states. S consists of terminal states and nonterminal states.
  - s_0 is the initial state of the game.
  - JAS is the input alphabet of the automaton, where $JAS = \prod_{i=1}^{n} A_i$ is the Cartesian production of n players' action sets. Each member in the JAS is a n-tuple, the ith element of this tuple corresponds to the action taken by player i. The fact that the input symbols are n-tuples means that the state to which the game would transfer depends on all players. For any individual player, even after taking a certain action, it doesn't know what the next state would be. The name JAS stands for Joint Action Sets.
  - Θ : transition function of the automaton, which mapping a joint action to a state, i.e. $\Theta(s, y)$, where $s \in S, y \in JAS$, denotes the state to which the automaton would transfer from state s given the joint action y.

Private information(only known to player i itself):

- r_i : the real type of player i, $r_i \in T_i$.
- $U_i(s, r_i)$: utility function of player i, which denotes the payoff player i will get if the game ends at state $s \in S$ and its real type is r_i.
- $Prob_i(a_j)$: the probability player i believes that player j would take action a_j.

At each time period $t < maxT$, if current state is a terminal state s, the game ends and all players obtain their final payoffs $U_i(s, r_i)$. Otherwise, the game goes forward, player i computes its payoff by following formula,

$$Payoff_i^t(s) = \max_{a_i \in A_i} \{ \prod_{j=1 j \neq i}^{n} Prob_i(a_j) \times Payoff_i^{t+1}(s') \} \quad (1)$$

where $s' = \Theta(s, a)$ is the state to which the game transfers from state s, driven by the joint action $a = \prod_{k=1}^{n} a_k$ taken by all the players.

To maximize its expected payoff, player i should take action a_i^* such that

$$a_i^* = \arg \max_{a_i \in A_i} \{ \prod_{j=1 j \neq i}^{n} Prob_i(a_j) \times Payoff_i^{t+1}(s') \}. \quad (2)$$

At the last period $maxT$, the payoff is realized as utility function of player i. i.e.:

$$Payoff_i^{maxT}(s') = U_i(s', r_i). \quad (3)$$

In the classic game theory, the word "type" refers to some private information that characterizes an individual player. All the players' beliefs about what "types" others may assume form the "common knowledge" of the game, i.e., everyone knows that everyone knows that ....everyone knows those beliefs. Based on the "common knowledge", the game converges to certain equilibrium. In a Internet negotiation scenario, for instance, the shopping example mentioned before, it is general hard for the software agents to form the "common knowledge" about others' types (reservation prices). Instead, agents' actions (price offers) are always observable. In our model, players observe their opponents' actions, interpret those actions based on some subjective beliefs and then take corresponding action to maximize their own payoffs. Every player's subjective beliefs are its own "personal experiences", not shared with anyone else. Just as "experiences" of a human negotiator will determine his strategy, "beliefs" held by an agent in our model will determine its behavior in negotiation.

If a player chooses to negotiate with only one other player, the time spent by it to compute its best offer is $O(|A_1||A_2| * m * maxT)$ in each iteration, where $|A_1|, |A_2|$ are the respective sizes of the action sets of these two players. The game keeps going for at most $maxT$ periods, hence the total computation complexity is $O(|A_1||A_2| * m * maxT^2)$ at worst case. If one chooses to negotiation with multiple players simultaneously, the computation time will be exponential to the number of players. But one can solve this problem by carefully choosing the number of players based on its computational resource constraints.

3 Negotiation process

In this section, we show how to apply above model to model a basic negotiation scenario involving one buyer, one seller and one commodity. The buyer and the seller have reservation prices v_b, v_s on the commodity, respectively. In each negotiation iteration t, the buyer and the seller offer price proposals simultaneously. If the buyer's offer b_t is no less than the seller's offer s_t, the game ends. Otherwise, the game goes to next iteration until reaching the maximal time horizon $maxT$. Clearly, v_b, the reservation price of the buyer, is a crucial piece of information to characterize the buyer, thus we set the "real type" of the buyer to be v_b. Similarly, the real type of the seller is v_s. To explicitly express the states of the negotiation game, we make two assumptions:

- There is a price range $[minP, maxP]$ agreed by both players, and v_b, v_s, b_t, s_t all belong to this range.

- The players have set a minimal price increase/decrease unit before the game begins.

Given these 2 assumptions, we can change the price scale such that all the prices involved in the negotiation are integers. For example, if the price range is $[\$0, \$78.10]$ and the price increase/decrease unit is 0.1 dollar, we can scale the unit to 1 dime and then the price range becomes $[0, 781]$. So, in the follow part of this paper, we always assume that all the prices, lower and upper bound of the price range, are integers. Once we make the price range discrete, we can define other parameters of the game based on the price range. Namely:

- The state set of the game is defined as $[minP, maxP] \times [minP, maxP]$. A state s is an ordered pair (x, y), where x is the seller's price offer, y is the buyer's price offer. If $x \le y$, s is a terminal state, otherwise, it is a nonterminal state.
- The action sets and the type sets for both players are $[minP, maxP]$.
- The transition function of the game is defined as $\Theta(s, a) = a$ for nonterminal state s, i.e., from a nonterminal state, the game can transfer to any other state defined by two players joint actions. For example, if the current time is t and the current state is $(10, 5)$, since the seller's offer 10 is greater than the buyer's offer 5, it is a nonterminal state. If the seller offers 9 and the buyer offers 6, then the game transfers to state $(9, 6)$ in the next iteration $t + 1$. For a terminal state, Θ is undefined since the game ends in this state.
- Every player maintains a vector of beliefs over its opponent's action set. During each iteration, the beliefs are fixed. A player may update its beliefs between iterations. The methods of updating the beliefs will be discussed in next section.

In a real-world negotiation, a buyer always prefers to a lower price for a certain commodity. Our model incorporates this observation by requiring that the buyer's utility function to be monotonous, i.e., U_{buyer} satisfies $U_{buyer}(s_1, v_b) \ge U_{buyer}(s_2, v_b)$ for any states s_1, s_2 such that $s_1 = (x, y_1), s_2 = (x, y_2)$ and $y_1 \le y_2$. If the game ends at a nonterminal state, the players do not agree with each other and hence no trade will happen. In this case, both the buyer and the seller get nothing. So, $U_{buyer}(s, v_b)$ should equal to 0 for any nonterminal state s, i.e., there is no penalty for both players if they fail to make an agreement. If these two conditions hold, for whatever beliefs (as long as they form a probability distribution) a buyer may hold, she will never offer a price higher than her reservation price.

Proposition 1: The buyer will never offer a price higher than her reservation price v_b.

Proof: Refer to [Huang and Sycara, 2000].

By symmetry, we can get another proposition with respect to the seller:

Proposition 2: The seller will never offer a price lower than his reservation price v_s.

These two propositions validate our model in the basic negotiation scenario. They look reasonable and simple, but the interest thing here is that we achieve this reasonable behavior by pure computations instead of imposing logic rules to the model.

4 Experiments and analysis

In all experiments, we assume that all the conditions mentioned in section 3 hold. And we still stick to the basic negotiation setting: one buyer, one seller and one commodity.

Before checking the updating methods, we first define the players' utility functions. Through out this section, both players use the linear utility functions. The buyer's utility function is defined as:

$$U_{buyer}((x, y), v_b) = \begin{cases} v_b - \frac{x+y}{2} & x \le y \\ 0 & o.w. \end{cases}$$

The seller's utility function is defined as:

$$U_{seller}((x, y), v_s) = \begin{cases} -v_s + \frac{x+y}{2} & x \le y \\ 0 & o.w. \end{cases}$$

In above formulas, $s = (x, y)$ is a state. As said before, $x \le y$ means s is a terminal state.

Suppose at time t, the seller's offer is s_t,and the buyer's offer is b_t, we compare the following belief updating methods (Here we state the methods from the buyer's perspective. It is easy to give the corresponding formulas for the seller):

1. The buyer doesn't update at all, and always uses the uniform distribution over the price interval $[minP, maxP]$
2. The buyer sets a uniform distribution over interval $[minP, s_t]$
3. The buyer sets an exponential distribution over interval $[minP, s_t]$, and s_t has the highest probability :

   $$Prob_{buyer}(a) = \frac{1}{Z} exp\{\frac{a - s_t}{\beta(maxT - t)}\}$$

   where $a \in [minP, s_t]$, Z is a normalization factor, t is the current time, $maxT$ is the time horizon, β is a parameter.
4. The buyer sets an exponential distribution over interval $[minP, s_t]$, and $minP$ has the highest probability
5. The buyer sets a uniform distribution over interval $[b_t, s_t]$
6. The buyer sets an exponential distribution over interval $[b_t, s_t]$, and s_t has the highest probability
7. The buyer sets an exponential distribution over interval $[b_t, s_t]$, and b_t has the highest probability

In the method 3,4,6 and 7, the time horizon of the game has been taken into consideration. In the method 3 and 6, the less the time left, the more the buyer believes that the seller will not change his current offer s_t. While in method 4 and 7, the buyer does not believe that the seller is offering her a reasonable price. The less the time left, the more she believes that the seller will decrease his current offer s_t.

4.1 Both players use the same method

For the buyer, the updating method 2, 3 and 4 set probabilities over interval $[minP, s_t]$. For the seller, the interval is $[b_t, maxP]$. In Figure 1, we show the trade prices for the cases where both the buyer and the seller use the same method in this group (method 2, 3 and 4). We fix the buyer's reservation price to be 100, increase the seller's reservation price from 0 to 100 with step length 5. We set $\beta = 5$ for method 3 and $\beta = 100$ for method 4. In Figure 1, the x-coordinate is the seller's reservation price; and the y-coordinate is the final trade price. From the results, we can find that method 4 produces the "hardest" negotiator among these three methods: only when the seller's reservation price falls between 0 and 30, the players can make a deal. If the seller's reservation price is higher than 30, even though there exists a potentially wide negotiation range, they do not get an agreement. On the other hand, method 3 produces the "easiest" negotiator: for the seller's reservation price varying from 0 to 90, they can always close a deal. Method 2 stands between the "easiest" and the "hardest". The range for method 2 to make a deal is [0, 45].

What is the effect of changing β in method 3 and method 4? A little thought will confirm us that with a bigger β, method 3 will become "harder", and method 4 will become "easier". A bigger β will make the exponential distributions of methods 3 and 4 (with opposite tail directions) converge to an uniform distribution, which is the distribution used in method 2.

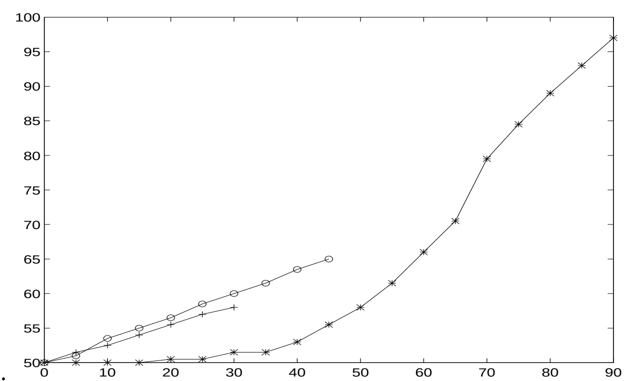

Figure 1: o:method 2, *:method 3 with $\beta = 5$, +:method 4 with $\beta = 100$

In method 2, 3 and 4, when the current price offered by the seller is s_t, the buyer sets the probabilities over range $[minP, s_t]$. Given that the buyer has offered b_t at time t, the seller will never offer a price lower than b_t after time t. Similarly, the buyer will never offer a price higher than the lowest price proposed by the seller before. Therefore, it seems more reasonable for the buyer and the seller to set the probabilities over $[b_t, s_t]$. That is what method 5, 6 and 7 do. We show the results in Figure 2. In this experiment, all the settings are the same as those in the first experiment except that the updating methods are changed to method 5, 6 and 7. We can see that this group of methods are more efficient than the methods used in the first experiment: with the buyer's reservation price fixed at 100 and the seller's reservation price varying

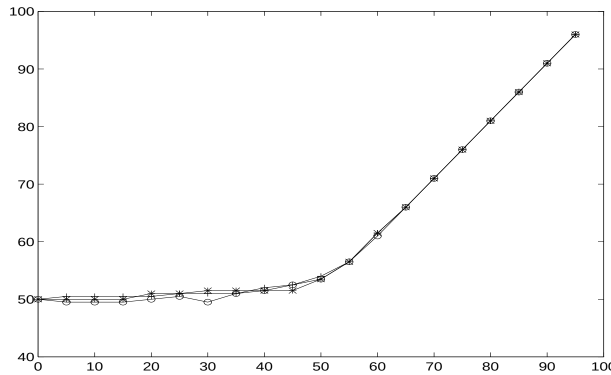

Figure 2: o:method 5, *:method 6 with $\beta = 5$, +:method 7 with $\beta = 100$

from 0 to 95, the players can always reach a deal by using any one of these three methods. An explanation to these results is that since both players have more "reasonable" beliefs on what price their opponent will offer, it is easier for them to negotiate successfully. Similarly, increasing β will make method 6 and 7 converge to method 5.

4.2 Players use different methods

In the above experiments, we showed the final trade prices if both players use the same belief updating method. In this section, we use a series of experiments to show what will happen if players use different updating methods. In all these experiments, the settings are the same as those in previous subsection.

Figure 3 shows the case where the seller uses method 4 and the buyer uses method 2. In Figure 1, we show the cases where both players use method 4 and method 2. In order to compare with those results, we show the results of those two cases here again. Compared to the case where both player use method 4, now the buyer switches to a "weaker" updating method, so the seller manages to sell his item at a higher price, and also extends the reservation price range on which he is willing to make a deal with the buyer. Compared to the case where both players use method 2, now the seller uses a harder updating method, so he still manages to sell his item at a higher price but shrinks the reservation price range on which he is willing to make a deal.

In the case where the seller uses method 2, the buyer uses method 3, as shown in Figure 4, we can find that the seller takes much advantage over the buyer. Compared to the case where both players use method 2, the fact that the buyer switches to a weaker method 3 is exploited by the seller, so he can manage to sell his item at a much higher price while extending the reservation price range simultaneously. Compared to the case where both players use method 3, although the seller decreases the reservation price range a little, the trade prices are much higher than those observed in Figure 2.

Figure 5 shows the case where the seller uses method 7 and the buyer uses method 5. Here we only compare the results with those in the case where both players use method 5 because the results obtained by using method 5 and method 7 are very close, as shown in Figure 2. In Figure 5, we also

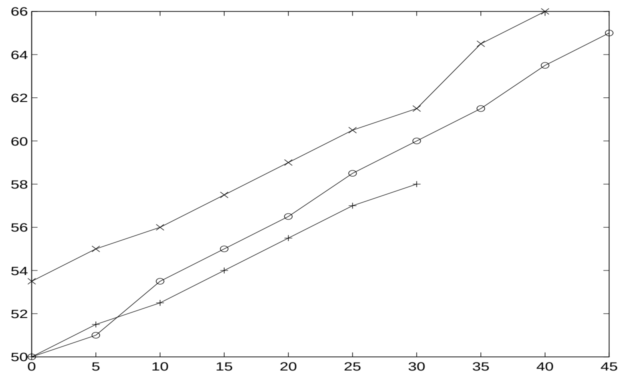

Figure 3: o:both use method 2, +:both uses method 4 with $\beta = 100$, x: the seller uses method 4 with$\beta = 100$, the buyer uses method 2

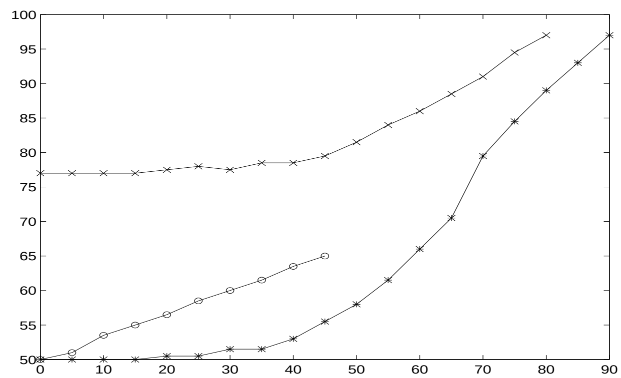

Figure 4: o:both use method 2, *:both use method 3 with $\beta = 5$, x: the seller uses method 2, the buyer uses method 3 with $\beta = 5$

show the results where the seller uses method 5 and the buyer uses method 6. In this case, since the seller uses a weaker method, the trade prices are lower than those in the case where he uses method 6.

In Figure 6, we show the trade prices where the seller uses method 4, the buyer uses method 7, and both with $\beta = 100$. We can see that the seller takes obvious advantage over the buyer.

In Figure 7. we show the trade prices where the seller uses method 2 and the buyer uses method 5.

In Figure 8. we show the results where the seller uses method 3 and the buyer uses method 6. Here the buyer uses a stronger method and she manages to buy the item at lower prices.

Till now, we have not elaborated on method 1. In Figure 9 we show the results where the seller uses method 1 and the buyer uses method 3. We can see that method 1 is close to method 3 with $\beta = 5$. But method 3 has an extra desirable feature: it performs as well as method 1 when the seller has a low reservation price, and is more flexible when the seller has a high reservation price. As shown in the figure, if the seller's reservation price exceeds 40 and both players use method 1, they can't reach a deal. But if the buyer uses method 3, they can make a deal.

To explain all these results, let's first take a close look at

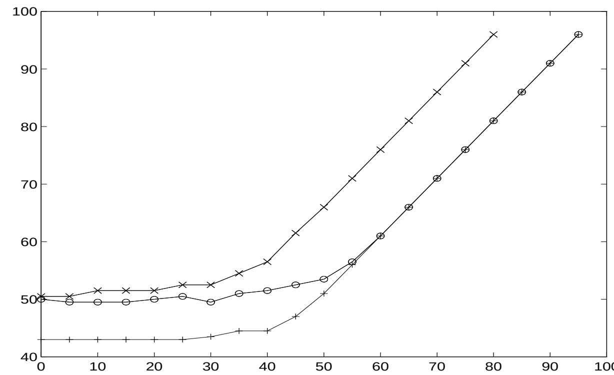

Figure 5: o:both use method 5, x: the buyer uses method 5 , the seller uses method 7 with $\beta = 100$, +: the buyer uses method 5, the seller uses method 6 with $\beta = 5$

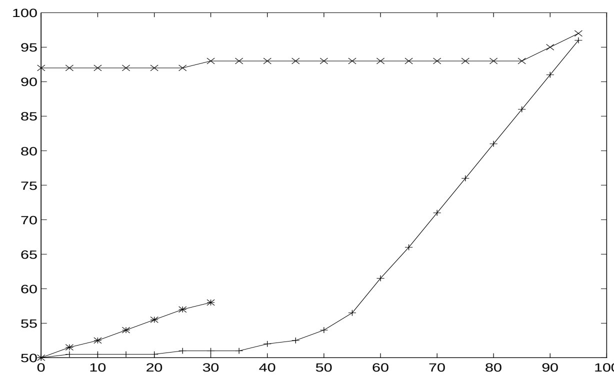

Figure 6: *:both use method 4 with $\beta = 100$, +: both uses method 7 with $\beta = 100$, x: the buyer uses method 7, the seller uses method 4, both set $\beta = 100$

the group of method 2,3 and 4. We claim that method 4 is "harder" than method 2. To show this, let's take the buyer's point of view. Method 4 and method 2 both set the probabilities over interval $[minP, s_t]$. The probability in method 4 decreases exponentially from $minP$ to s_t. This means that the buyer believes that the chance for the seller to offer a lower price is exponentially higher than the chance for the seller to offer a higher price. The probability distribution in method 2 is uniform. This means that the buyer believes that all the prices in range $[minP, s_t]$ are equally likely to be proposed by the seller. In a real-word negotiation, no one knows what kind of beliefs are "right" without extra information; but it is exactly the different beliefs that make the negotiators behave differently. Obviously, if the buyer holds the beliefs of method 4, she will be a harder negotiator than the one who holds the beliefs of method 2. Similarly, we can say that method 2 is harder than method 3. In the group of methods 5, 6 and 7, we can find that method 7 is harder than method 5, and method 5 is harder than method 6. By the same argument, we can say that method 2 is harder than method 5. Given the same β, method 3 is harder than method 6, and method 4 is hard than method 7. Of course, two methods are compared with the same β value if they involve this parameter.

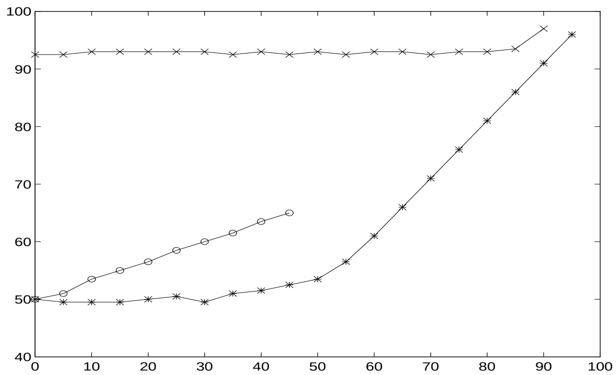

Figure 7: *:both use method 5, o: both uses method 2, x: the buyer uses method 5, the seller uses method 5

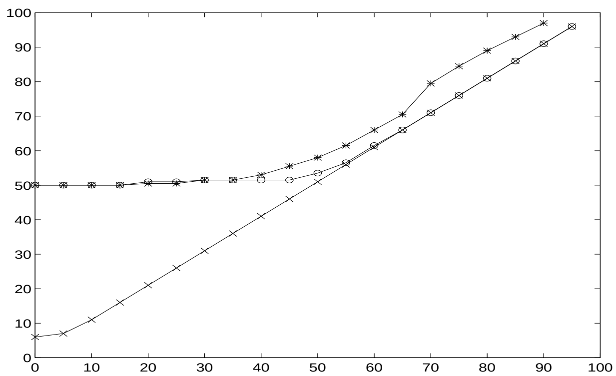

Figure 8: $\beta = 5$, o:both use method 6, *: both uses method 3, x: the buyer uses method 3, the seller uses method 6

5 Conclusions

We have developed a computational model for on-line agent negotiation. In the basic negotiation scenario, we validate our model by showing that an agent will never make an offer that will possibly compromise its own ability to gain a benefit. We achieve this objective by pure computations, not by imposing logical constraints on the agent model. One of the difficulties presented to a negotiation agent on the Internet is that it has little information about its opponents. To achieve mutual interactions while defending its own benefits at the the same time, an agent can adopt the "beliefs" mechanism to adjust its behavior. We show that with different "internal beliefs", agents can behave differently, just as human beings: some are hard negotiators, while others are very willing to make a deal with his/her opponent. By simulations, we empirically show the relative strength of a group of "internal belief updating" methods. Our experiments show that an agent should make a trade-off in negotiation: if it is too hard , it may lose the chance to earn more profits by making a deal with its opponent; while if it is too weak, the agent probably just earns marginal profits even it gets a deal with its opponent. We show that we can adjust the parameter β to make this trade-off in our model. The advantage of our model is that it is flexible and easy to implement. To show different "personalities" in a negotiation, one only need to plug in suitable "subjective beliefs" to one's agents.

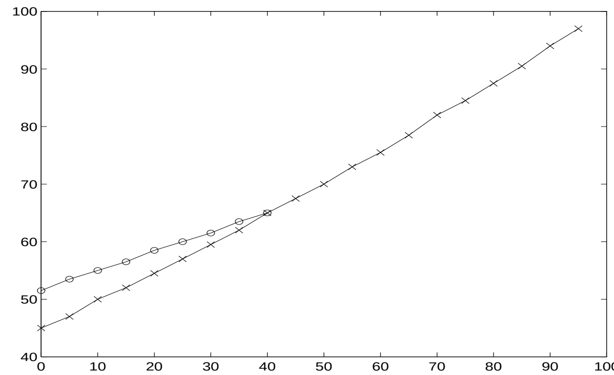

Figure 9: o:both use method 1, x: the buyer uses method 3, the seller uses method 1

References

[Bazzan and Bordini, 2000] Ana Bazzan and Rafael Bordini. Wayward Agents in a Commuting Scenario (personalities in the Minority Game). In *Proceedings of the 4th International Conference on MulitAgent Systems*, pages 55-62, Boston, MA, July 2000.

[Bichler, 2000] Martin Bichler. A Roadmap to Auction-based Negotiation Protocols for Electronic Commerce. In *Proceedings of the 33th Hawaii International Conference on System Science*, Maui, Hawaii. 2000.

[Fudenberg and Tirole, 1995] Drew Fudenberg, and Jean Tirole. *Game Theory*. The MIT Press, Cambridge, MA, 1995.

[Gmytrasiewicz and Lisetti, 2000] Piotr J. Gmytrasiewicz and Christine L. Lisetti. Using Decision Theory to Formalize Emotions in Multi-Agent Systems. In *Proceedings of the 4th International Conference on MulitAgent Systems*, pages 391-392, Boston, MA, July 2000.

[Huang and Sycara, 2000] Pu Huang and Katia Sycara. A Computational Model for On-line Agent Negotiation (extended version). *Working Paper, the Robotics Institute, Carnegie Mellon University*, Pittsburgh, PA, 2000.

[Schut and Wooldridge, 2000] Martijn Schut and Michael Wooldridge. Intention Reconsideration in Complex Enviroments. In *Proceedings of the 4th International Conference on Autonomous Agents*, pages 209-216, Barcelona, Catalonia, Spain, June 2000.

[Sen *et al.*, 2000] Sandip Sen, Anish Biswas and Sandip Debnath. Believing others: Pros and Cons. In *Proceedings of the 4th International Conference on MulitAgent Systems*, pages 279-285, Boston, MA, July 2000.

[Sycara, 1998] Katia Sycara. Multiagent Systems. *Artifical Intelligence*, 19(2):79–92, 1998.