Simulating Sellers' Behavior in a Reverse Auction B2B Exchange*

Subhajyoti Bandyopadhyay[1], Alok R. Chaturvedi[2], John M. Barron[2], Jackie Rees[2], and Shailendra Mehta[2]

[1] University of Florida, 364 Stuzin Hall, Gainesville, FL 32611-1310
shubho.bandyopadhyay@cba.ufl.edu
[2] 1310 Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47907-1310
{alok, barron, jrees, mehta}@mgmt.purdue.edu

Abstract. Previous research in reverse auction B2B exchanges found that in an environment where sellers collectively can cater to the total demand, with the final (i.e. the highest-priced bidding) seller catering to a residual, the sellers resort to a mixed strategy equilibrium [2]. While price randomization in industrial bids is an accepted norm, it may be argued that managers in reality do not resort to advanced game theoretic calculations to bid for an order. What is more likely is that managers learn that strategy and over time finally converge towards the theoretic equilibrium. To test this assertion, we model the two-player game in a synthetic environment, where the agents use a simple reinforcement learning algorithm to put progressively more weights on selecting price bands where they make higher profits. We find that after a sufficient number of iterations, the agents do indeed converge towards the theoretic equilibrium.

1 Introduction

The recent industry excitement around business-to-business (B2B) e-commerce has brought into focus the large potential size of the market as well as the mechanisms that are expected to bring multiple benefits to the participants of B2B exchanges. Forrester Research predicts that there will be $1.5 trillion in goods and services transacted online among U.S. businesses domestically by 2003, and $2.5 trillion internationally [14]. Therefore, it is critical to attempt to better understand the dynamics comprising this particular type of market.

Some of the more prominent advantages B2B exchanges are expected to bring include lower costs due to automating the procurement process, reverse auctions, interoperability among users, collaborative planning and collaborative

* This research is partially funded by the National Science Foundation Grant No. DMI-0122214.

P.M.A. Sloot et al. (Eds.): ICCS 2003, LNCS 2660, pp. 365–374, 2003.

design [10]. For example, Ford announced in July 2001 that it had saved $70 million through Covisint (the online automotive exchange by the Big Three automakers) in terms of reduced paperwork and lower supplier prices, which is more than its initial investment in the exchange [10]. Bandyopadhyay, Barron and Chaturvedi [2] analyzed the competition between sellers in reverse auctions in a game-theoretic framework, and established the Nash equilibria in several scenarios. It was found that in an environment where sellers can collectively cater to the total demand, with the final (i.e. the highest-priced bidding) seller catering to a residual, the sellers resort to a mixed strategy Nash equilibrium. While price randomization in industrial bids is an accepted norm, it may be argued that managers in reality do not resort to advanced game theory calculations to bid for an order. What is more likely is that managers learn that strategy over time and finally converge towards the theoretic equilibrium. This paper tests that assertion by creating an artificial B2B exchange in the synthetic environment for analysis and simulation (SEAS) [6]. It models the sellers' behavior with artificial software agents that start bidding randomly, and use a simple reinforcement learning (RL) algorithm to "learn" the ideal strategy over time.

The competition among sellers in the environment mentioned above is different from the traditional oligopolistic Cournot competition between firms facing a downward facing demand curve, where both firms sell at the same price point. It is also distinct from a capacity-constrained Bertrand model that has been analyzed extensively in the literature, where a quantity precommitment and Bertrand competition yield Cournot outcomes that have equilibrium prices above marginal cost [12].

The analysis of Bandyopadhyay et al. [2] also established the nature of the equilibrium under various assumptions of the sellers' cost, capacities and the market demand. The problem is most interesting when we assume that there is no *combined* capacity constraint as such: the sellers were supplying to the entire demand before the birth of the exchange, and continue to do so after it comes into play. However, it is conceivable that the firms individually cannot supply to the entire market (in fact, it is to their benefit not to have too much capacity, when there is industry overcapacity in the first place, as then the competition reduces to a Bertrand game with all firms supplying at their marginal cost). Since the set of suppliers is limited and all are reputed in the marketplace, the buyers would not mind getting their orders fulfilled by any one or several of these suppliers. This means that while there is a competition between the firms to be the low-price bidder, it is not as extreme as a Bertrand game that results in prices equal to marginal cost. However, there remains an incentive to be the low-price bidder and have the "first invitation" to supply a requirement.

We use an agent-based simulation approach to study the sellers' behavior in a B2B market place. The use of simulation allows repeated and detailed study of the behaviors exhibited by the players in the market under various experimental treatment conditions. These artificial agents are endowed the ability to learn from previous actions by the use of a type of Reinforcement Learning (RL) algorithm described below. Artificial agents have been used to simulate human agents or players in a number of different settings such as for automated negotiations in an e-commerce environment [7, 16].

RL has been used to examine various competitive scenarios such as sealed bid k-double auction under asymmetric and incomplete information dynamics [17], market entry games [8] and rule learning in repeated games [3]. RL is an appropriate choice for the application presented in this paper due to the ability of RL agents to incorporate previous experience (either reward or no reward) into action. The model under which the artificial agents operate in this research is discussed in the next section.

The remainder of the paper is organized as follows: The model of the reverse auction is provided in Section 2. Section 3 presents the RL algorithm deployed in the simulation. Section 4 states the research assertions, hypotheses that are tested in the simulation, and provides the results. Conclusions and future research directions are discussed in Section 5.

2 The Model

SEAS uses intelligent agents (IA) to represent economic realities of electronic markets and hierarchies in a decentralized manner [5]. SEAS' intelligent agents are autonomous processes that are adaptive and behave like human agents in a narrow domain. In their respective domains, each agent has a well-defined set of responsibilities and authorities so that it can execute its tasks effectively. Examples of SEAS' IAs are: economic agents --consumer IAs, producer IAs, and regulator IAs; political agents -- government IAs, special interest IAs, etc. An agent in SEAS is equipped with reasoning, action, and communication skills required for performing their respective tasks. A SEAS IA is characterized by the knowledge it possesses.

SEAS markets are implemented in JavaSpace (Figure 1). JavaSpace is a descendant of the Linda system developed at Yale University [4]. Java Space defines the market structure. The rules for all the markets are implemented through JavaSpace, which in turn synchronizes the thread between the agents. Agents maintained in the space are updated through their working status. JavaSpace also forms the connectivity between the agents and the Database, and therefore, after completing the transactions, it updates the database.

Let us suppose that there are two buyers who bought from two sellers (one from each) before the advent of the exchange. What prevented buyers from establishing contact with both the sellers (and vice versa) are the search costs and the ongoing cost of establishing relationships within large organizations (dedicated account management teams for buyers, sales force for sellers, cost of sending individual RFQs to the entire universe of sellers, etc.) [11].[1] With the lack of competition, the sellers could afford to sell the required quantities to the buyers at their reservation price, which we assume to be the same for both buyers at *r*. With the advent of the exchange, the buyers put forward their requirements to the exchange, and the sellers can then bid for the total requirement from both buyers. The above example is just illustrative, and is not crucial to the analysis. There can be in fact only

[1] It has been estimated that in terms of reduction of paperwork alone, B2B exchanges can bring down costs per purchase order from $75-$150 to $10-$30 [10]

a single large buyer, whose requirements cannot be met by one seller – however, two sellers together have a combined capacity that is more than the buyer's requirement.

SEAS B2B model derives much of its intuition from the basic 2-player model, and therefore it is instructive to first consider the 2-player model in detail. We consider the case when both sellers have equal capacities K that is more than the respective individual requirements of the buyers, but is lesser than the total requirement of both buyers Q (i.e. $2K - Q > 0$). In such a setting, the lower priced seller is invited first to sell the required quantity, and after he has supplied his total capacity K, the other seller can then sell the residual demand $Q - K$. Both sellers have a common fixed marginal cost of production, c.

From the modeling point of view, it is immaterial whether the sellers respond to an aggregate demand of several buyers or one single demand from a buyer. What is important to note is that the entire requirement is auctioned to the sellers, and for any unfulfilled demand, a lower priced bidder is invited before a higher priced bidder to satisfy the unfulfilled demand. It is readily apparent that with unlimited capacity, the sellers respond with a Bertrand competition in prices with the seller or sellers with the lowest marginal cost outbidding the others. This is not to the advantage of the sellers. Kreps and Scheinkman [12] (and several variants of the original model, such as [1])) show that if sellers could limit capacity, then a quantity precommitment and Bertrand competition yield Cournot outcomes that have equilibrium prices above marginal cost. At the other end of the spectrum, if the total capacity of the sellers is so limited as to be less than the total demand, it is easy to see that the sellers can sell their entire capacities at the buyer's reservation price.

The analysis shows that there exists a mixed strategy equilibrium of prices where the sellers randomize between a range of prices. The intuition behind such an equilibrium is as follows: With two similar players, there cannot be any equilibrium in pure strategies with the players settling on different prices. Settling on the same price is also ruled out, since the best response to any price is to set a price that is ε lower than that price. Thus, if any Nash equilibrium exists, it has to be mixed strategy equilibrium. It can further be shown that the support of the strategy lies between p_1 and r, where r is the reservation price for the buyer and p_1 is given by

$$.\, p_1 = \frac{(r-c)(Q-K)}{K} + c \qquad (1)$$

An intuitive way of looking at p_1 is that below this price, a seller makes less profit by "winning" (supply to capacity) than by "losing" and supply the residual at the highest possible price r (which is the best price the seller can supply the residual, since he is losing anyway).

The equilibrium strategy for either player can be expressed in terms of their price randomizing cumulative probability density function $F(p)$:

$$.\, F(p) = \frac{(p-c)K - (r-c)(Q-K)}{(p-c)(2K-Q)} \qquad (2)$$

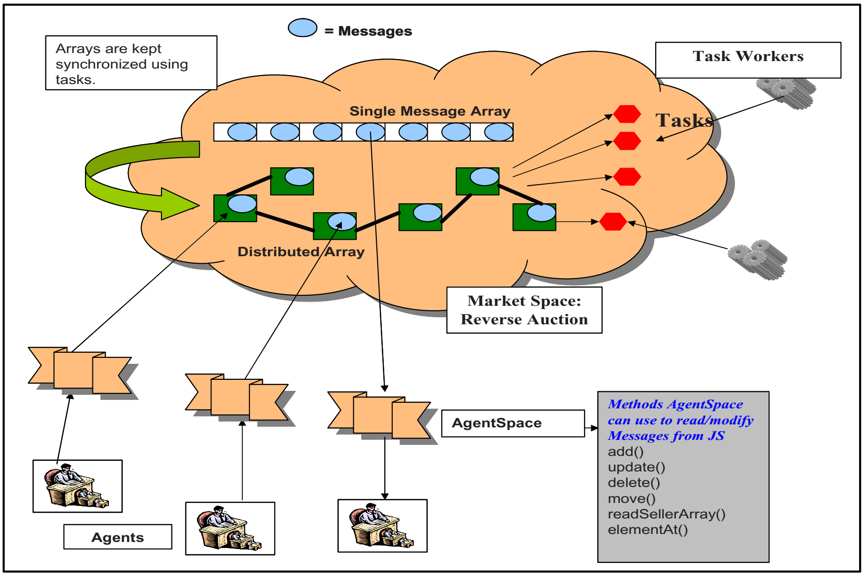

Fig. 1. SEAS Agent-based Architecture

This is a continuous probability distribution within the range (p_1, r), and effectively defines the symmetric Nash equilibrium strategy of the two players (i.e. the suppliers). The suppliers randomize their bids within this interval, such that their randomizing has a probability distribution given by *F(p)* in (2). By definition the Nash equilibrium maximizes the expected return of the suppliers.

The analysis will be similar for the n-player model, where the highest bidder supplies the residual, and the rest supply to capacity ($(n-1)K < Q < nK$). The support of the strategy for the sellers is given by (p_1^n, r), where

$$p_1^n = \frac{(r-c)(Q-(n-1)K)}{K} + c\,, \tag{3}$$

and the expression for the distribution function is given by

$$F_n(p) = \left[\frac{(p-c)K-(Q-(n-1)K)(r-c)}{(p-c)(nK-Q)}\right]^{1/n-1} \tag{4}$$

While price randomization in industrial bids is an accepted norm, it might be difficult to accept that managers go through advanced game theory calculations (and in any case, the real-life situations are far more varied than the simplified model scenarios that make any game theory analysis extremely complex) to determine their bids. It is conceivable that sellers learn from their past experiences to bid in a fashion that maximizes their surplus. It is this assertion that we test in the remainder of this paper.

3 The Algorithm

To test our assertion, we model the competing sellers as artificial software agents. Like human subjects, we propose that these agents "understand" the following (without resorting to knowing game theory):

1. There are two opposing forces in the pricing strategy – a higher price (towards *r*) means greater per-unit profit, but also brings about a higher probability of "losing" to the competition (in terms of being the first invited bidder to supply the demand).
2. It does not make any sense to price below p_1, as is clear from the above analysis.
3. Since there is a need to balance between higher probability of winning and higher per-unit profit, there is no a priori reason to rule out any price between p_1 and *r*, and further, there is reason to believe that price randomizing might be the ideal (or equilibrium) solution.

The "sellers" thus start off initially with a totally random pricing strategy (i.e. the price distribution is uniform in its support), with the hope of learning over time about the ideal nature of the randomization. This is the same assumption as Erev and Roth [8] employ in their experiments and refer to as the "initial propensities" of the players for their pure strategies.

If we denote the average profit of player *j* (*j*=1,2) in division *k* (*k*=1,…10) in simulation round *t* as $\Pi_{jk}(t)$, then the probability $p_{jk}(t+1)$ of choosing that division in round *t+1* is given by

$$p_{jk}(t+1) = \frac{\Pi_{jk}(t)}{\sum_{k=1}^{10} \Pi_{jk}(t)} \tag{5}$$

The game is then repeated a sufficient number of times so that the players can hopefully learn sufficiently to converge to the ideal distribution. The experiment can be repeated with other values of *Q*, *K*, *c* and *r*.

4 Hypothesis Testing, Results, and Discussion

For testing the assertion, we selected various values of *Q*, *K*, *c* and *r*. In the 2-player model equilibrium, we can see from the expressions of p_1 and *F(p)* that the drivers of interest are the values ($2K-Q$) and ($r-c$). If the combined capacity (2*K*) is barely more than the demand (*Q*), the sellers have little incentive to lower prices, while if there is a large amount of overcapacity, the sellers would greatly reduce prices (since losing would mean catering to a very small residual demand). The difference ($r-c$) on the other hand would determine the range of the support of prices. We keep Q (the total quantity demanded) fixed at 100 units and r (the reservation price) fixed at \$80. The values of K chosen reflect the amount of overcapacity: at K=65, we have moderate overcapacity, prompting moderate competition; at K=80, the possibility of supplying only a small fraction of the demand (i.e. if the supplier bids the higher price, he ends up supplying the residual of only 20 units) should prompt more severe competition. For each of these values of K, we choose the value of c, the marginal cost, as \$60.

For each of the pairs of values of K and c, we calculate p_1 and corresponding subdivision limits, which are shown in the *Bin* column. After running the simulation as described above, we find out the number of times the prices are picked in each subdivision, and this is given in the Frequency (O_i) column. The theoretical cumulative distribution gives us the theoretically expected number of observations in each subdivision, and this is presented in the E_i column. We compute the χ^2 statistic as $\sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i}$, and this is compared with the corresponding chi-square value with $p = 0.05$ (16.92). Formally stated, we would reject the null hypothesis that the data follows the distribution specified in (2), if the calculated χ^2 exceeded the corresponding χ^2 value with a significance level α of 0.05:

H$_0$: $F_n(p) = F_n^*(p)$ where $F_n^*(p)$ is the experimentally generated distribution.

H$_1$: $F_n(p) \neq F_n^*(p)$

The simulation results show that as expected from the theoretical results, lower prices are preferred over high prices (Tables 1 & 2). We run a chi-square goodness of fit test with each of the simulation settings. As the results show (the 'Chi-sq.' column in the tables compute the χ^2 statistic, whose sum is shown in the final row, and this sum is compared to the corresponding chi-square value with p = 0.05 which is shown in the final column), the fit with the theoretical distribution is always very good. In all the nine cases, we do not reject the null hypothesis that the experimental frequency distribution follows the theoretical probability distribution. In other words, the simulation runs results in the agents learning over time to come very close to the ideal distribution with every set of values of the parameters K and c.

It is therefore observed that the artificial software agents start off selecting their prices uniformly throughout the interval of (p_1, r), but gradually learn over time to select lower prices with monotonically higher probabilities. In fact, the final frequency distributions show that the learning is 'perfect' within margins of statistical error. The results have interesting ramifications in real-world scenarios. Managers might not have the luxury of learning over a large number of observations themselves as in these simulations, but they can utilize the "organizational memory" (i.e. the experiences of him as well as his predecessors) to effectively build the learning capability over time. Managers also have their own intuition, which these artificial agents lack that might result in accelerated learning towards equilibrium (and therefore optimal) behavior.

5 Conclusions

While we are currently addressing many of these issues in our ongoing research, the results of our simulations with the 2-player model show considerable promise. We hope that these results spur the interest of using automated agents that will enable organizations to effectively compete in the increasing number of electronic transactions. While one of the main attractions of B2B exchanges remains in their ability to automate the processes by which organizations can participate in electronic transactions with each other, the problem of overseeing each and every one of them is still very much an issue. This problem will likely exacerbate in future as more and more organizations start to utilize these electronic services. While the algorithms that need to be used in real-world scenarios will be much more complex than those presented in this research, we think that organizations might over time develop such algorithms of increasing sophistication. At first, very basic transactions having routine processes would be entrusted to such learning mechanisms. As algorithms get more complex, and simultaneously organizations also gain confidence in such mechanisms, more complex transactions would probably be entrusted. Organizations might also develop processes by which unusual procedures set off triggers for either human intervention or even a complete abort.

Table 1. Simulation run results with $K = 65$ units, $c = \$60$

Bin	Frequency (O_i)	Th. Cum. Fr.	Th. Freq. dist. (E_i)	Chi-sq.	Chi-sq. value (p=0.05)
71.69230769	176	171.0526316	171	0.143093117	
72.61538462	152	317.0731707	146	0.244855636	
73.53846154	133	443.1818182	126	0.376585912	
74.46153846	109	553.1914894	110	0.009266784	
75.38461538	86	650	97	1.206752397	
76.30769231	78	735.8490566	86	0.717628032	
77.23076923	75	812.5	77	0.035558781	
78.15384615	67	881.3559322	69	0.050024511	
79.07692308	68	943.5483871	62	0.052534014	
80	56	1000	56	0.351041475	
	1000		1000	3.187340659	16.92

Table 2. Simulation run results with K = 80 units, c = $60

Bin	Frequency (O_i)	Th. Cum. Fr.	Th. Freq. dist. (E_i)	Chi-sq.	Chi-sq. value (p=0.05)
66.5	284	307.6923077	308	1.824307692	
68	197	500	192	0.114492308	
69.5	140	631.5789474	132	0.538947368	
71	99	727.2727273	96	0.114229904	
72.5	79	800	73	0.541022727	
74	61	857.1428571	57	0.260357143	
75.5	48	903.2258065	46	0.079749309	
77	36	941.1764706	38	0.100264137	
78.5	35	972.972973	32	0.322752385	
80	21	1000	27	1.344027027	
	1000		1000	5.24015	16.92

References

1. Allen, B., Deneckere, R., Faith, T. and Kovenock, D., "Capacity precommitment as a barrier to entry: a Bertrand-Edgeworth approach," *The North American Winter Meetings of the Econometric Society*, January 1992
2. Bandyopadhyay, S., Barron, J.M. and Chaturvedi, A.R., "A game-theoretic analysis of competition among sellers in an online exchange," Working Paper, 2002
3. Bell, A. M., "Reinforcement Learning Rules in a Repeated Game," *Computational Economics,* Vol. 18, pp. 89-111, 2001.
4. Carriero, N. and Gelernter, D., How to write parallel programs: A guide to the perplexed, ACM Computing Surveys, 21, 3 (September 1989), 323-357.
5. Chaturvedi, A and Mehta, S., The SEAS Environment 2, Technical Report, Purdue University, West Lafayette, IN, 2002.
6. Chaturvedi, A. R. and Mehta, S., "Simulations in Economics and Management: Using the SEAS Simulation Environment," Communications of the ACM, March 1999.
7. Epstein, J. M. and R. Axtell, Growing Artificial Societies: Social Science from the Bottom Up, Brookings Institution Press, Washington DC, 1996.
8. Erev, I. and Rapoport, A. "Coordination, 'Magic,' and Reinforcement Learning in a Market Entry Game," Games and Economic Behavior, Vol. 23, pp. 146-175, 1998.
9. Erev, I. and Roth, A.E., "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *The American Economic Review*, Vol. 88 No. 4, pp. 848-881, 1998
10. Helper, S. and MacDuffie, J.P. "B2B and Modes of Exchange: Evolutionary and Transformative Effects," *The Global Internet Economy*, edited by Bruce Kogut (forthcoming)

11. Kerrigan, R., Roegner, E.V., Swinford, D.D. and Zawada, C.C. "B2Basics," *Mckinsey Quarterly* No. 1, 2001, pp. 45-53
12. Kreps, D. and Scheinkman, J. "Quantity precommitment and Bertrand competition yield Cournot outcomes", *The Bell Journal of Economics*, 1983, pp. 326-337
13. Kuhn, H.W. and Nasar, S. "Editor's Introduction to Chapters 5, 6 and 7", *The Essential John Nash*, pp. 48
14. Larson, P. "B2B E-Commerce: The Dawning of a Trillion-Dollar Industry", *The Motley Fool's Internet Report*, March 2000
15. Nash, J.: "Non-Cooperative Games" *Annals of Mathematics*, Vol. 54, September 1951, pp. 286-295
16. Oliver, J. "A Machine Learning Approach to Automated Negotiation and Prospects for Electronic Commerce," *Journal of Management Information Systems,* Vol. 13, no. 3, pp. 83-112, 1996.
17. Rapoport, A., Daniel, T. E. and D. A. Searle, "Reinforcement-Based Adaptive Learning in Asymmetric Two-Person Bargaining with Incomplete Information," *Experimental Economics,* Vol. 1, pp. 221-253, 1998.
18. Roth, A.E. and Erev, I. "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, Special Issue: Nobel Symposium, 8(1), pp. 164-212, 1995
19. Sutton, R. S. and A. G. Barto, *Reinforcement Learning: An Introduction*, The MIT Press, Cambridge, MA 1998
20. Thorndike, E.L. "Animal Intelligence: An Experimental Study of the Associative Processes in Animals," *Psychological Monographs*, 2(8), 1898.