MACHINE LEARNING APPROACH TO FINDING OPTIMAL STRATEGY FOR THE SEALED-BID AND HYBRID AUCTIONS

Oleksa Stepaniuk, MSc Data Science

Paper provides a methodology for finding an optimal strategy of an auction by starting from the random strategy and iteratively applying neural network. Such approach allows to finding an optimal strategy close to optimal in the case of sealed-bid auction (with uniform or exponential distribution of private values). Algorithm fails to converge in the case of a more complex hybrid auction. Despite this, results provide interesting insights into how hybrid auction is used in the public procurement market and can be of use to companies taking part in hybrid auctions.

1. Introduction

Every finite game has a mixed-strategy Nash equilibrium [1], however, there cannot be a universal algorithm for finding it [2]. As a result, researchers use specific features of games to find their equilibriums. Such an approach requires extensive training and time. In most cases, this is not an obstacle for using games and game-theoretic approaches for real-life applications – optimal strategies of most popular games (such as basic types of auctions) have long been researched and published in open access. Still, life continuously produces new games. It is possible that such a game is used in a specific situation or market before the optimal strategy for it is found by game-theorists. Also, a new game can be a “trivial” extension of the previously known game and, as a result, not attract the interest of researchers. In such a case, agents participating in a game could benefit from using a numerical method for finding the approximate optimal strategy.

The method for finding approximately optimal strategy can also be helpful for the researchers working on finding the “exact” solution as well as for modeling which of the multiple Nash equilibria is more likely in a given market. Lastly, in a real-world situation, it is possible to face an opponent that uses a strategy that is not a Nash equilibrium. Numerical methods can be used for finding the best response strategy in such a case.

One of the examples, of games used in real-life applications that do not have known Nash equilibria is a hybrid auction used for the procurement of public goods in Georgia, Ukraine, and Moldova. This auction is interesting for a number of reasons. Firstly, its specification was conditioned by practical/political reasons and not preceded by a game-theoretic analysis. Nevertheless, thousands of companies annually take part in those auctions. Moreover, all the data from each auction is available to participants and any interested researchers. As a result, the use of hybrid auction for public procurements can be regarded as a massive natural experiment. It can be expected that with time strategies of the companies participating in the auctions will converge to the evolutionary stable strategy [3].

Accordingly, this paper has a twofold goal. We want to show that numeric methods can be used for finding an optimal strategy for a sealed-bid auction and then use the same approach for hybrid auction. Secondly, we want to compare the found solution with the data from the real procurement auctions.

Paper is structured as follows: the second chapter provides a brief overview of papers related to the automatic finding of Nash equilibria and exploratory analysis of the data from procurement auctions. The third chapter describes the methodology for automatically finding Nash equilibrium for sealed-bid auction and its adaptation for hybrid auction. Chapters four, five and six provide an overview of the results, conclusions, and the discussion.

2. Related work

This chapter consists of two sections. In the first section, we provide the definition of auctions as well as discuss some of the existing algorithms for numerically finding Nash equilibriums. The second section explains the mechanics of the hybrid auction and provides an overview of existing exploratory studies of procurement systems that use it.

2.1 Auctions and numerical approaches to finding Nash equilibrium

Auctions are an example of non-cooperative games with incomplete information. Participants are assumed to be competing for owning the object of the auction. Each agent knows his private value of the auction object (how much is the object worth to him) however he is ignorant regarding the valuations of other agents. An important characteristic of auctions is the continuous action space – usually, a bid represents some amount of money which is best represented as a continuous variable. This makes a task of applying numerical and more specifically reinforcement learning approaches to solving auctions more challenging.

The basic set-up for the auction analysis is the situation with independently and identically distributed (IID) private values. In such a situation, each agent knows his private value, the number of auction participants and the probability distribution of private values of other participants. Most importantly, the valuation of one agent is not affected by the valuations of other agents. This set-up is definitely unrealistic (especially in the case of procurement auctions), however, it provides a basis for further elaboration. Such elaborations may include asymmetrical agents, risk-averse agents, agents with interdependent values, cases of collusion, etc.

In this paper, we will consider only the basic example of private value auction with two players. Specifically, we will consider sealed-bid auction and hybrid auctions. The sealed-bid auction will be used to show that the numeric method can provide a strategy close to the Nash equilibrium as described in the Vickrey, W. (1962) [4].

There were a number of successful attempts to produce algorithms for automatic identification of Nash equilibriums in a certain class of games. For instance, Reeves (2005) [5] describes an iterative mechanism for finding an equilibrium for infinite games of incomplete information. Remarkably algorithm requires only one or two iterations (depending on the initial strategy) to find the Nash equilibrium of the sealed-bid auction. Unfortunately, this approach can be applied only for auctions with one round while hybrid auction has at least two rounds.

An alternative approach is described in Hu, Wellman (2003) [6]. Authors showed that Q-learning can be effectively used for solving stochastic games with the discrete action space. Also, there were a number of papers that apply deep learning techniques to a more general reinforcement learning problems with continuous action space (see [7] and [8]). It seems that the most widely used approach to reinforcement learning problems learning with the continuous space in the actor-critic method. For instance, this method is the most successful one for solving Continuous Mountain Car Challenge from OpenAI Gym [9].

In the machine learning domain, it is a common practice to evaluate the efficiency of algorithms based not only on theoretical proofs but based on the empirical tests. Taking this into account, the goal of the paper will be to apply the most straightforward, "brute force" approach to finding the equilibrium of the hybrid auction. This result then could be used as a benchmark for the comparison of the efficiency of more elaborate solutions.

2.2 Hybrid auction

Slightly different variants of the hybrid auction are used for allocating procurement contracts in Ukraine, Georgia, and Moldova. We will concentrate on the version of auction used in Ukraine. It is most suitable for the analysis due to the public availability of data. Hybrid auction used in Ukraine consists of four stages and is described in Table 1. It is important to remember that in the case of procurement auctions, companies compete for a contract by offering a price for which they are willing to deliver goods/services. The company that offers the lowest price wins the auction.

Table 1. Mechanism of the hybrid auction

ROUND	DESCRIPTION
Before auction	Auction announcement is published 15 days before the auction. Announcement includes all the requirements and the reserve price
Round 0	Before the announced deadline, all companies that want to participate in the auction submit online their documents and the initial bid. Initial bid cannot be higher than the reserve price. At this stage they do not know the number of auction participants. This is the zero, “blind” round of the auction (R0).
Round 1	With the start of the auction the initial bids of all participants are revealed. This is round 1 (R1). After the reveal all bidders successively make their next bid starting with the participant with the highest initial bid. At this stage participants instantaneously see the bids of each other.
Round 2 and Round 3	This procedure is repeated during two consecutive rounds (R2 and R3)
After auction	The bidder who offered the lowest price at the end of the third round is the “winner” of the auction. He proceeds to sign the contract if documents that he submitted are in accordance with the auction requirements. If not, such participant is disqualified and documents of the next company with the most attractive price are checked.

In 2018 in Ukraine[1], 35 thousand companies took part in 87 thousand hybrid auctions to compete for the public procurement contracts. Data on the bid of each participant in each round of the auction is publically available. There were a number of empirical studies dedicated to the analysis of Ukrainian procurement auctions. In the context of this study, the most significant is the pattern discovered by Kovalchuk (2017) [10]. Specifically that 80% of all winners of the auction were also the winners of the zero round of the auction. This pattern is reasonable – a company that wins the blind round is the last one to offer the next bid in the first round. As a result, such a company sees the bids of all other participants before choosing its own bid. In other words, in the first round, each consecutive company has more complete information about the bids of competitors. The last company has complete information about bids of others.

[1] Statistics includes only finished, competitive, above-threshold procurements. Source: ProZorro Business Intelligence module https://bi.prozorro.org

This is an important insight into the optimal strategy of the hybrid auction. However, one piece of information is missing. If it is so advantageous to win the zero round, then what should be the bid in the zero round? This is one of the questions that we want to answer in this paper. Preliminary it can be said that by offering a more attractive bid in the zero round we increase our probability of winning, however, we simultaneously decrease our expected profit from the auction. The optimal strategy should maximize expected profit.

3. Methodology

In this chapter, we will describe our approach to finding an optimal strategy for sealed-bid and hybrid auctions. As was mentioned in previous chapters we will consider cases with two participants that have symmetric IID private values. Without loss of generality, all private values will be normalized to the interval [0, 1]. The neural network will be used as a learning algorithm, though the methodology described below can be used with any other algorithm that predicts the continuous output.

In addition, in the case of a hybrid auction, we will consider not a procurement auction (were bidder with the lowest bid wins) but a more usual case where participants compete by increasing their bids. We will model hybrid auction with two rounds – initial sealed-bid round followed by a "dynamic" round. Auctions with a larger number of "dynamic" rounds require much higher computational resources but represent an almost identical concept.

3.1 Sealed-bid auction

The first iteration of the sealed-bid auction:

1. For each agent, a vector of private values is determined using some probability distribution (for instance, uniform distribution). Each pair of values represent one auction.
2. Agents randomly select a bid as a percent of their private value.
3. Vectors of bids are compared. If the bid of the first agent exceeds the bid of the second agent, the first agent wins the auction. If bids are equal, the winner is determined randomly.
4. If an agent wins, his profit equals the difference between his private value and a bit. If the agent loses his profit equals zero.
5. A dataset consisting of vectors of private values, bids and profits is formed.
6. A learning algorithm is trained to predict the average expected profit given the private value and a bid as a percentage of private value.

Using the learning algorithm we can select a bid that given the private value will maximize expected profit. This bid will be the optimal bid if the opponent continues to play randomly. The algorithm for selecting optimal bid:

1. For a given private value, the expected profit for n bids evenly distributed on the interval [0,1] is predicted.
2. Two neighboring bids b1 and b2 with the highest expected profit are selected.
3. The expected profit for n bids evenly distributed on the interval [b1, b2] is predicted.
4. The steps 2-3 are repeated until an optimal bid of required precision is found.

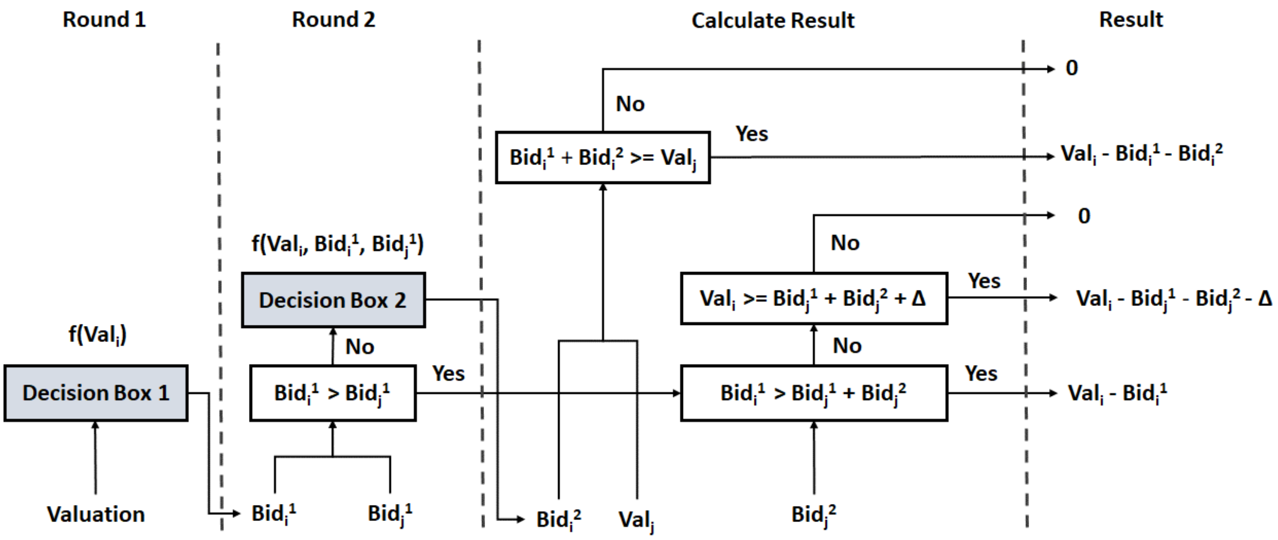
The next iteration of finding the optimal strategy for a sealed-bid auction has the same steps as the previous one, except only one of the agents plays randomly. The other is playing the optimal strategy determined during the previous iteration. The optimal strategy determined during the second iteration will be the optimal response to the strategy from the first iteration. After some number of iterations, it is expected that the algorithm will converge to a Nash equilibrium where we will have a symmetric strategy that will be the best response to itself.

3.2 Hybrid auction

We define the agent with the highest bid in the first round as the winner of the first round. The second round bid is defined as additional amount bided by the agent. As a result, the final bid equals the sum of the first round and second round bids. The agent with the highest final bid is the winner of the auction.

Mechanism of the hybrid auction includes two stages where agents have to make a decision. During the first stage, they make a bid based on their private value and expectation regarding the private value of the opponent. During the second stage, the loser of the first round offers his second bid. The other agent observes this bid and responds with his own bid. To make a modeling task simpler we separated decisions that are made based on the learning algorithm and automatic decisions. An automatic decision is defined as an obvious decision that maximizes agents expected profit (see Figure 1). It is important to notice, that in the second round only the decision of agent who lost in the first round is based on the learning algorithm. The decision of the winner of the first round is automatic.

Figure 1. Implementation of the hybrid auction mechanism*



* Δ denotes the minimum allowed increase of the bid. We use Δ=0.00001

The first iteration of the algorithm for finding the optimal strategy of the hybrid auction (hybrid auction algorithm) is similar to the first iteration of the sealed-bid algorithm:

1. For each agent, a vector of private values is determined using some probability distribution (for instance, uniform distribution). Each pair of values represent one auction.
2. Agents randomly select the first-round bid as a percent of their private value.
3. Vectors of bids are compared. If the bid of the first agent exceeds the bid of the second agent, the first agent wins the first round. If bids are equal, the winner of the first round is determined randomly.
4. The loser of the first round randomly selects the second bid as a percentage of (1-Bid^1)*Pr.Value.
5. Winner of the auction is determined using the method described in Figure 1.

After the first iteration of the algorithm, we will collect a dataset with the following structure:

Table 2. Dataset for training learning algorithm

Column	Values	Explanation
Round	0 or 1	0 denotes the first round
Previous bid value	0 or Bid^1*Pr.Value	Equals 0 in the first round
Remaining private value	Pr.Value or (1- Bid^1)*Pr.Value	Equals Pr.Value in the first round
Winner of the first round	0 or 1	Equals 0 if round=0 or agent lost in the first round
Opponent abs. bid in first round	0 or float	0 in the first round, product of opponents bid and private value in the 2nd round
Bid	$x \in [0, 1]$	Percentage of the remaining private value
Profit	$y \in [0, 1]$	Final outcome of the auction

A neural network trained on the dataset will predict the average expected profit given the explanatory variables. The output of the neural network will be used to determine the optimal first-round bid given the private value of the agent as well as optimal second-round bid in the case when an agent lost the first round.

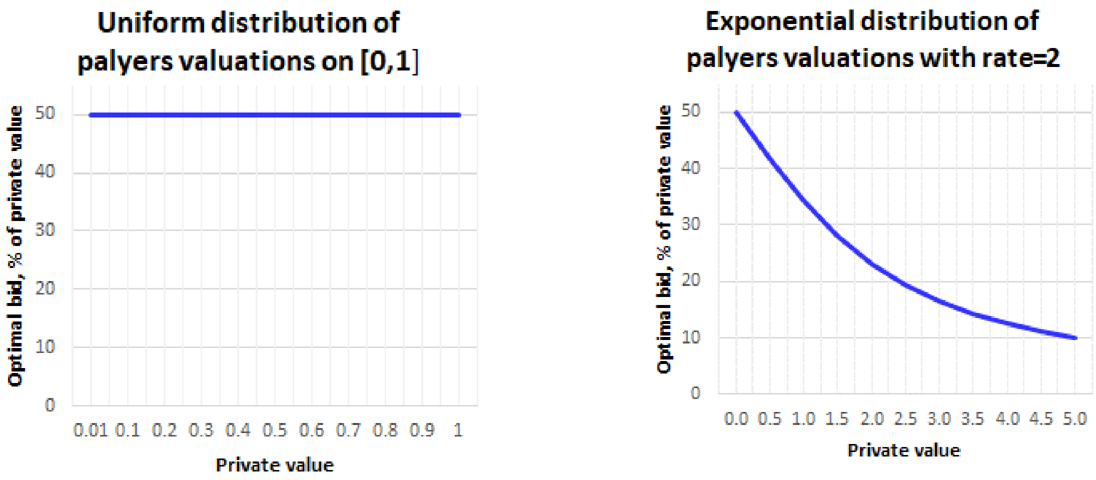
In the next iterations of the hybrid auction algorithm, one agent will continue to play randomly. For the other agent, we will choose the first round and second round bids based on the learning algorithm. Iterations are repeated until Nash equilibrium is found.

4. Results

4.1 Sealed-bid auction

We will be testing the efficiency of our approach using the sealed-bid auction, were valuations of players are distributed using the uniform or exponential distribution. Values of optimal bids (as derived in [11]) in these two cases are provided in Figure 2. As can be seen, if private values are distributed uniformly, the optimal bid as a share of valuation does not depend on valuation – it always equals to 50%. The optimal strategy for the exponential distribution of private values is quite different. If the rate of exponential distribution equals 2, then mean valuation equals 0.5 and the probability of having a private value greater than 2.3 is only 1%. As a result, even player with a large private value will not make a large bid. For instance, it is optimal for a player with valuation equal to 4 to bid only 0.5, i.e. 12.5% of his valuation.

Figure 2. Values of analytically derived optimal bids, sealed-bid auction with two playersx



4.1.1 Only bids are used for training

Considering the fact that optimal bid in the case of uniform distribution does not depend on the valuation, we will first test the simplest model with only one input – the value of bid (number from 0 to 1). Using this information neural network will predict the average expected outcome without knowing the actual private value of the player.

Figure 3 shows the result of the first iteration when both agents randomly select bid. Red dots indicate the actual outcome of the auction, blue line denotes expected outcome predicted by the neural network. As can be seen from Figure 4, the algorithm converges in 4 iterations. The mean value of optimal bid oscillates around 47% (mean value of optimal bids from last 16 iterations). Even though the algorithm never reaches optimal analytically derived bid value of 50% it is quite close. Consequently, an agent would be better off by following the strategy provided by the algorithm than by playing randomly.

Figure 3. Result of the first iteration of algorithm. Sealed-bid auction with uniform distribution of bids, only bid is used for training

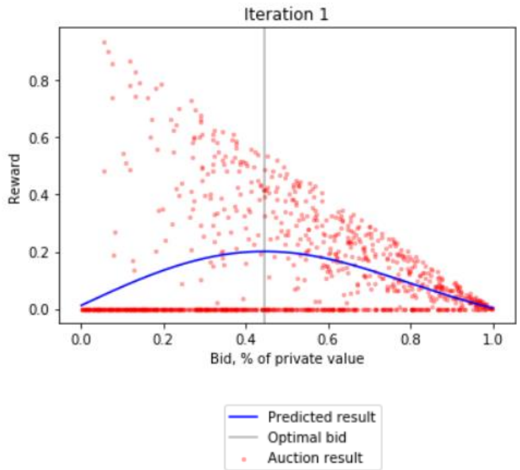
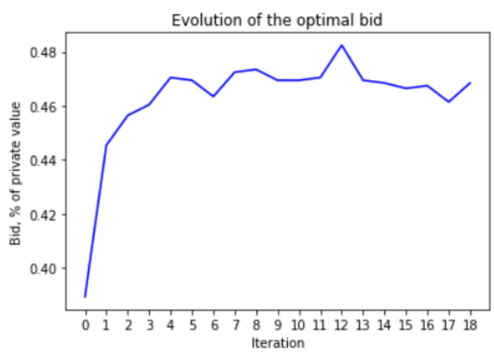


Figure 4. Convergence of the optimal bid value.



During the implementation of this model we encountered one difficulty – due to the random character of sample generation, algorithm periodically converged to an optimal bid of 0% or 100%.

We managed to solve the problem by substituting the new extreme value of optimal bid by the optimal bid of the previous iteration. In this way we allowed the algorithm to continue the progress.

4.1.2 Bids and private values are used for training

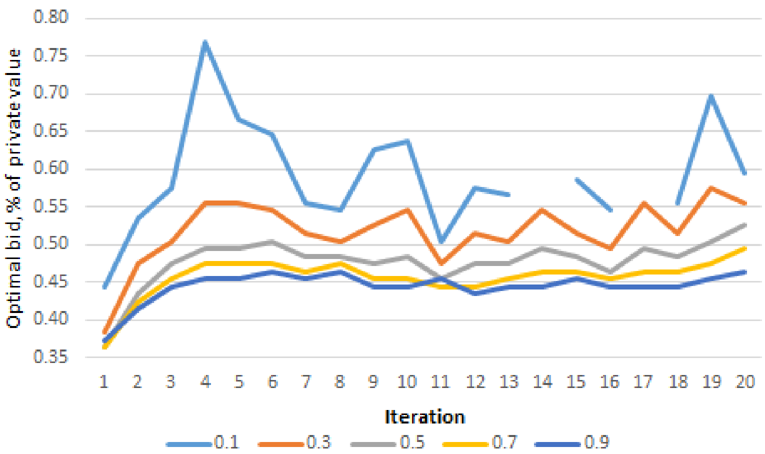
In the previous model, after the initial iteration, the opponent always chose one optimal bid regardless of his private value. This section introduces a more complicated model, where the opponent chooses an optimal bid based on his private value. This adds substantial computational cost to the model. Specifically, we are using a vector of one million private values and for each of them, we predict the expected profit from one hundred possible bids. As a result, at each iteration, we predict the expected profit from 100 million combinations of private value and a bid.

As previously, we encountered a problem of algorithm converging to either 0 or 1. We tried to mitigate the problem by substituting the cases where neural network predicts extreme optimal bid by average optimal bid of other cases. This, however, does not change the trajectory of convergence and after several iterations, all of the optimal bids converge to 0 or 1. The tendency was also unaffected by changes in the activation function (logistic, relu, tanh).

The problem was solved by using a deeper neural network (3 inner layers with 10, 5 and 3 units) and a system of 'indicator private values.' Specifically, after each iteration we calculated optimal bid for five private values: 0.1, 0.3, 0.5, 0.7, 0.9. If any of the optimal bids equaled 0 or 1 we repeated iteration using the previous version of the neural network.

Figure 5 shows the convergence of optimal bids for these five valuations. The average optimal bids for last 15 iterations are (starting from private value of 0.1: 59%, 53%, 49%, 46% and 45%. As can be seen from the graph, the convergence path for larger private values is much more stable, there is practically no variation in the optimal bid for private values of 0.5, 0.7 and 0.9. Optimal bid for the private value of 0.1 is especially volatile. Twice algorithm predicted that it is optimal to bid 0% if valuation equals 0.1, thus two times algorithm repeated the previous iteration. Interestingly all optimal bids are close to the true optimal bid of 50%. However, for private values above 0.5 recommended optimal bid is less than true optimal and for private values above 0.5, it is greater than true optimal.

Figure 5. Convergence of the optimal bid for different private values. Uniform distribution



In the case of exponential distribution of private values, optimal bid does depend on the private value. This makes a learning task more challenging, the algorithm has to learn a non-linear relationship between the bid, private value, and expected profit. Figure 6 depicts evolution of optimal bids for four private values: 0.3, 0.5, 0.7, 0.9. Optimal bid for the private value of 0.1 was even more volatile and did not converge, thus, it was omitted from the graph. Remaining optimal bids have a common trajectory for the first ten iterations and then diverge as was expected.

Figure 6. Convergence of the optimal bid for different private values. Exponential distribution

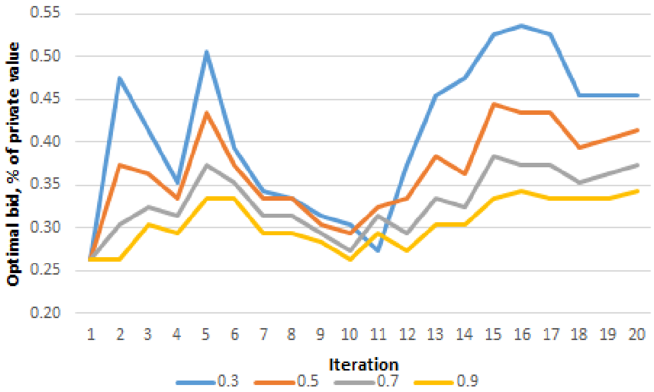


Table 3 provides a comparison of analytically calculated optimal private bids and average optimal bids from the last ten iterations of the algorithm. On average predicted values of optimal bids differ from the true values by 2.5 percentage points.

Table 3. Comparison of predicted and true values of optimal bids. Exponential distribution of private values

Private value	True optimal bid, %	Average predicted optimal bid, %	Difference, perc. points
0.3	45.03	45.25	0.22
0.5	41.80	39.29	-2.51
0.7	38.70	34.85	-3.85
0.9	35.75	31.92	-3.83

4.2 Hybrid auction

In this section, we apply the algorithm tested on a sealed-bid auction to a hybrid auction described in the Methodology. In this auction, an agent makes a decision twice: during the first round based on his private value and during the second round based on his private value and the bid of the opponent in the first round. Importantly, in the second round, we consider only those instances, when the agent lost the first round but his private value is larger than the opponent's bid in the first round. In other cases the decision of agent in the second round is automatic.

One of the difficulties that we encountered while applying the algorithm to the hybrid auction is determining the impact of changes in the neural network on the algorithm efficiency. Changes of the activation function and the depth of the neural network did not have a measurable impact on the algorithm convergence. The testing was further constrained by the computational requirements of the algorithm – one iteration required, on average, 5 minutes and each testing requires multiple iterations.

Figures 7 and 8 show the path of optimal bids for the first and second round through 20 iterations. As can be seen, the optimal bid for the first round is very stable and close to the optimal level of the sealed-bid auction (50%). Interestingly, the optimal bids for private values of 0.5, 0.7 and 0.9 are by

3-5 percentage points lower than their sealed-bid counterparts. This is an expected result – hybrid auction provides additional information during the second round, thus it makes sense to bid less during the first round. Optimal bids for the second round do not converge. The average value of second round optimal bids in the cases when the algorithm provides acceptable result is 68%.

Figure 7. Convergence of the optimal bid for different private values. First round of the hybrid auction

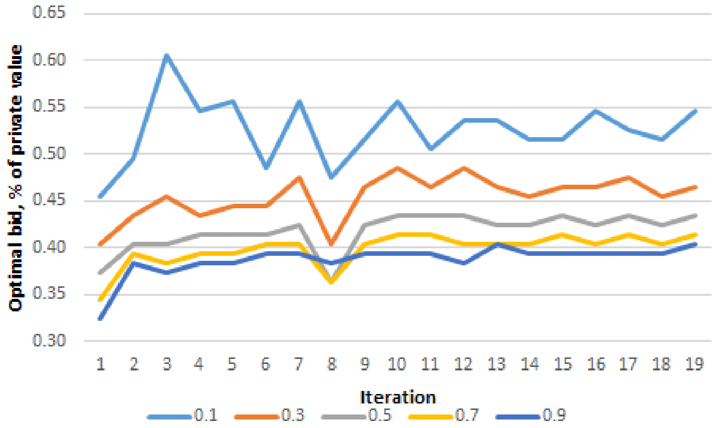
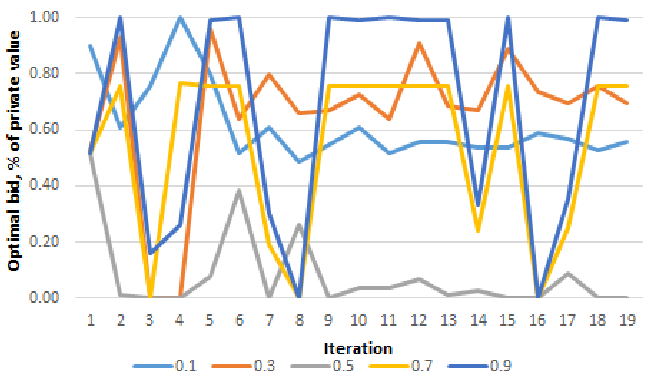


Figure 8. Convergence of the optimal bid for different private values. Second round of the hybrid auction



5. Conclusions

In this paper, we investigated possible ways of finding the optimal strategy of auctions – a subtype of games with incomplete information. Specifically, we tested whether it is possible to determine Nash equilibrium of the sealed-bid auction by starting with a random strategy, finding the optimal response to random strategy using the neural network and then iteratively repeating the process until optimal strategy converges. For the simplest model with uniform distribution of private values and only bids used as an input for neural network, algorithm converged to optimal bid equal to 47% (3 percentage points below the optimal level).

For a more complex model, when both private value and bid were used as an input to the neural network, the algorithm was also within 5 percentage points from the optimal value. Interestingly algorithm recommended a more risk-averse behavior than optimal if the private value is above 0.5 and a more risk-loving strategy when the private value was above 0.5. In the case of the exponential distribution of private values, the algorithm also proposed optimal bids within 4 percentage points from the optimal level. It should be noted, that algorithm was able to learn a non-linear relationship between the private value and the level of optimal bid.

Lastly, we applied the approach tested on the sealed-bid auctions to the hybrid auction. Unfortunately, we concluded that the straightforward approach that worked for simpler auctions does not allow to find the optimal strategy of the hybrid auction. Specifically, despite stable results for the first round, optimal bids in the second round are highly volatile and in the majority of iterations converge to either 0 or 1. Still, judging from the successful iterations, the optimal strategy suggested by the algorithm in the case of uniform distribution of bids consists in bidding approximately 50% of private value in the first round. If the agent loses the first round, the bid of the second round should be approximately 70% of the difference between remaining private value and the bid of the opponent in the first round.

6. Discussion

This paper is an example of difficulties that arise when working with reinforcement learning problems that have continuous action space. In such cases agent has an infinite number of choices and adding a new dimension or a new agent exponentially increases the complexity of the problem. As a result, straightforward statistical methods are inefficient. In the case of this paper, even samples with millions of observations were not enough to capture the complexity of the hybrid auction. As a result, a more sophisticated approach should be used, for instance, actor-critique method or Bayesian approaches as suggested by the literature.

Despite problems with convergence, our experiments provided an interesting counterpoint to the empirical data from hybrid auctions that are conducted in the Ukrainian public procurement system. Specifically, winners of the first round did not have a substantially higher probability of winning an auction. This can certainly be the result of the volatile results of our algorithm, however, this can also indicate that in a real market either agents with higher private values are more likely to bid more in the first round, or agents are generally risk-averse and unwilling to offer high bids after losing the first round. These are the promising hypothesis that can be tested.

Lastly, it should be noted that the algorithm described in the paper can have practical value to agents taking part in the hybrid auctions despite its inability to converge to the Nash equilibrium. In a real setting, a more important task is to find an optimal response to a given opponent or market situation that do not necessarily represent Nash equilibrium. This is especially true for the public procurement market where data about past auctions is available, thus a probability distribution of opponents private values can be inferred. The algorithm would be especially useful for markets with homogeneous goods, for instance, the natural gas market.

References

[1] Nash, J. (1951). Noncooperative games. The Annals of Mathematics, 54:289–295.

[2] Papadimitriou, C.H. (2007). The Complexity of Finding Nash Equilibria, in Nisan N. (ed.) Algorithmic game theory. 1st ed. Cambridge: Cambridge University Press, pp. 49-50.

[3] Smith, J. and Price, G. (1973). The Logic of Animal Conflict. Nature, 246(5427), pp.15-18.

[4] Vickrey, W. (1962). Auctions and Bidding Games. Recent Advances in Game Theory, no. 29 in Princeton Conference Series, pp. 15–27. Princeton, NJ: Princeton University Press

[5] Reeves, D. M. (2005). Generating Trading Agent Strategies: Analytic and Empirical Methods for Infinite and Large Games. Ph.D. Dissertation. University of Michigan, Ann Arbor, MI, USA. AAI3192760.

[6] Hu, J., Wellman, M.P. (2003). Nash Q-learning for general-sum stochastic games. Journal of Machine Learning Research 4 (December 2003), 1039-1069.

[7] Lazaric, A., Restelli, M., Bonarini, A. (2007). Reinforcement learning in continuous action spaces through sequential monte carlo methods, Advances in Neural Information Processing Systems 2007.

[8] Lee, K., Kim, S., Choi, J., & Lee, S. (2018). Deep Reinforcement Learning in Continuous Action Spaces: a Case Study in the Game of Simulated Curling. Proceedings of the 35th International Conference on Machine Learning, PMLR 80:2937-2946, 2018.

[9] openai/gym. (2019). Retrieved from https://github.com/openai/gym/wiki/MountainCarContinuous-v0

[10] Kovalchuk, A. (2017). Exploratory analysis of the ProZorro hybrid auction. Working paper, Kyiv School of Economics.

[11] Krishna, V. (2009). Auction theory. Boston: Elsevier Academic Press, pp.16-17.