Reasoning About Other Agents' Beliefs and Desires in Competitive Scenarios

Dylan D. Doblar
Massachusetts Institute of Technology
Cambridge, MA 02139
`ddoblar@mit.edu`

Nathan P. Foss
Massachusetts Institute of Technology
Cambridge, MA 02139
`npfoss@mit.edu`

Abstract

In many competitive settings, it is advantageous to be able to predict opponents' actions, likely by inferring their internal beliefs and desires. We analyze experimental data which gives insight as to how human agents formulate bids in a blind auction setting and find that they are capable of inferring their opponents' beliefs and desires from past actions and of using this information in strategy formulation. We also propose an algorithm by which an actor can generate bids in a blind auction setting with fixed budget constraints, where both the actor and the single opponent have preferences over a specified set of features for each class of items. We devise and algorithm to estimate the opponent's internal state (their preferences over the feature set) and use it to inform future bids. We compare our algorithm to other strategies that do not account for the opponent's past bids and to experimental data from human trials.

1 Introduction

In nature, animals are frequently pitted directly against each other in competition, whether it be for food or other scarce resources. The ability to predict an opponent's actions in these scenarios would be a massive advantage. In noncompetitive everyday life, we humans make inferences about what other humans are thinking all the time. Is that person about to cut in front of me on the road? Does this person have an ulterior motive for acting the way they are? Where does my girlfriend really want to eat dinner? All these actions we're trying to predict have a latent variable in common that we have little hope of observing directly: the other agent's beliefs and desires.

There is already strong evidence that we humans have intuitive physics engines in our minds, and we use them to simulate the world for all kinds of predictive purposes (1). Therefore it is natural to think that we may also have the ability to simulate other agents, given their beliefs and desires as some research has already suggested (2). But how do we figure out what those beliefs and desires are? For a physical model of the world, it is much easier to see the possible latent properties of objects, from size and shape to weight and texture. Other agents' internal states are much more challenging to model because the only evidence we have for this rich latent space is a sparse set of observed actions.

Therefore, we propose an experiment in a contained scenario with a much simpler space of beliefs and desires. In an auction with specifically defined desirable features of object, the possibly intractable space of all beliefs and desires of an agent can be collapsed into a short vector describing how much utility an agent will gain due to each aspect of a water bottle. This simplification allows us to study specific phenomena and effects of human inference about other humans' beliefs and desires, and have a hope of modeling these interactions with computers.

2 Setting and Task

The competitive setting that we explored in both our model formulation and our human experiments is that of a blind auction. The agent is a bidder in the auction and is given a set of preferences over relevant features of the items to be auctioned off, along with a fixed budget constraint. The agent's objective is to maximize their net utility at the end of the auction — net utility is calculated as the sum of the individual utilities given by each item won, each of which is a function of the agent's preferences and the features of each item (see 2.3 for details on how utility is calculated).

2.1 Setup and Rules of the Auction

There are two participants in the auction. Each participant is assigned an arbitrary — but fair[1] — preference vector along with fixed and equal budget constraints. The values of an agent's preference vector determine how much utility they attribute to the most valuable setting of each respective parameter.

There are a fixed number of items to be auctioned off, all of which are fully observable for the duration of the auction. Items are brought up one at a time in a predetermined order and are auctioned off in the following way. For each item, both bidders privately write down their respective bids for the item. The higher bidder wins the item, and pays the average of the two bids. In the case of equal bids for an item, the bidders get to re-bid on the same item until the bids differ. Splitting the difference results in the winner paying less than their bid but more than their opponent's bid — this combined with the fixed budget constraint and finite number of items incentivizes bidders to maximize the amount their opponent has to pay for their won bottles so that the opponent has less wealth to spend on future items.

2.2 Item Features

In our setting, there is a single class of items, water bottles. Each water bottle has three relevant features: volume, opacity, and contents. Volume is continuous, and can range from 700mL to 1500mL. Opacity is binary — a bottle can be either opaque or transparent. The contents of the bottle is also binary; it is either pre-filled with water or empty. In general, agents prefer water bottles that are larger, opaque, and pre-filled with water (it is also possible for an agent to be indifferent for a particular characteristic).

2.3 Utility Function and Preferences

An agent j's utility function is fully characterized by the vector $[V^{(j)}, O^{(j)}, C^{(j)}]$ that represents the utility that the agent would gain from the most desirable option for each feature (maximum volume, opacity, and containing water, respectively). That is, agent j would gain $V^{(j)}$ utility winning a water bottle with maximum volume, $O^{(j)}$ utility if the bottle is opaque, and $C^{(j)}$ utility if the bottle contains water. Each $V^{(j)}$, $O^{(j)}$, and $C^{(j)}$ can take on a value from 1 to 7. The utility $V_i^{(j)}$ gained from the volume category of water bottle i is proportional to its volume (relative to the min and max possible volumes), v_i, scaled between 1 and $V^{(j)}$. Agent j's utilities gained from the opacity and contents categories for bottle i are 1 each if the bottle is transparent and empty, or $O^{(j)}$ if the bottle is opaque and $C^{(j)}$ if it contains water (respectively). Thus, net utility of an agent who has won a set of bottles $\mathcal{W}$ is

$$U(V^{(j)}, O^{(j)}, C^{(j)}) = \sum_{i \in \mathcal{W}} 1 + (V^{(j)} - 1) * v_i + 1 + (O^{(j)} - 1)\mathbb{1}_{i\ is\ opaque} + 1 + (C^{(j)} - 1)\mathbb{1}_{i\ is\ full}.$$

[1]A pair preferences is "fair" if the max net utilities (the utilities of each agent if they win all items) are approximately equal.

3 Experiment

We conducted an experiment pitting human participants against each other in this auction game, where each was instructed to maximize their own utility at the end of the auction. We performed eight trials with fair sets of preferences that span many of the possible combinations of opposing preferences that we deemed to be of interest (equal, some at extremes, and all intermediate values, among others).

Participants were shown all of the water bottles in a randomized order (we had an assortment of eight water bottles, refer to 9.1 for all of the bottles' features). The auction proceeded as described in 2.1, and once the bottles were all sold, participants were asked to estimate their opponent's preference vector.

Refer to 9.2 for an example task description given to an experiment participant with preference vector $[7, 6, 2]$.

4 Experiment Results

While our experiment was rather limited in scope and leaves plenty of room for further experimentation, we have a number of findings that suggest that people do in fact maintain an internal state representation of their opponent's beliefs and desires.

Refer to Table 1 for the results from our 8 trials.

Table 1: Auction Results

Auction	Bidder Preferences	Bottles Won	Net Utility	Guess for Opponent
1	[7,1,6]	1,2,3,8	34.75	[3,5,2]
	[7,6,2]	4,5,6,7	33.13	[7,2,6]
2	[4,2,5]	1,4,5,6	27.06	[4,1,6]
	[4,2,5]	2,3,7,8	24.38	[5,1,6]
3	[7,2,7]	1,2,5,8	33.75	[7,7,2]
	[7,6,2]	3,6,4,7	36.13	[1,6,7]
4	[6,2,2]	1,3,5,7,8	25.13	[1,7,3]
	[1,7,3]	2,4,6	27	[5,3,1]
5	[1,6,1]	4,6,7	24	[6,1,1]
	[6,2,2]	1,2,3,5,8	27.94	[3,2,7]
6	[3,1,7]	1,2,3,5,7	25.38	[1,3,7]
	[3,1,7]	4,6,8	28.25	[3,1,7]
7	[1,6,4]	1,2,5,6,8	29	[6,4,1]
	[4,4,4]	3,4,7	21.56	[2,5,5]
8	[4,4,4]	1,2,4,5	25.69	[2,6,4]
	[1,6,1]	3,6,7,8	23	[7,1,7]

There are a number of insights that can be gleaned from our results. Perhaps the most striking aspect is the high degree of accuracy in many of the participants' estimates of their opponents' preferences after an auction of only eight water bottles. Over half of the participants (nine of the sixteen total) had estimates that were within one from their opponents' true preferences in every category, and many of the others discovered the relative utility between categories (what mattered most) if not the exact magnitude.

Something we were surprised to see is that there is no clear correlation between winning the auction and accurately guessing the opponent's utility. We had expected people to do noticeably better when they knew what their opponents wanted. It could be that some people did not think about this during the auction, and in fact we noticed that often when we asked for estimates of the opponent's preference, people would request to see the bottles they won and seemed make a judgment based on that. Another confounding factor could be that we did so few trials, and one of our subjects enjoyed the auction so much he did it three times and won all three despite a dismal utility-guessing record.

Notice that even over varied combinations utility functions, most final scores were quite close, within 5-10%. This seems to suggest that either the game is naturally well balanced, perhaps from the budget constraint, or that MIT students are naturally all fairly proficient at it, or both. There does not appear to be a pattern in how close different games were, but we hypothesize that with more experimentation factors not covered by the simple definition of 'fair' utilities we used will come into play. For example, subjects reported it was harder to play with utility $[4, 4, 4]$ than one with clear goals such as $[1, 6, 4]$.

As somewhat of an aside, in most cases both people got more than half their total possible utility from all the bottles. This is interesting because it shows both that auctions are an efficient way to distribute goods, and this is not a strictly competitive zero-sum game—even without cooperation, total utility gained in the end is greater than if either player had gotten all the bottles.

5 Model Design

When designing our model, we wanted to capture a couple aspects of what we observed in human behavior. We designed a few models of increasing complexity and compared them to each other in simulated auction settings. For all strategies, however the bid was generated, we added zero-mean unit-variance Gaussian noise in order to make no strategy completely deterministic, so that it is reasonable to do inference on the bids and because real world bidding is also noisy.

The first (and simplest) model that we devised was one that bid an equal amount for each item. Initially we set this amount to be the total budget divided by the number of items, but we later revised this to be double that amount because the agent can expect to win approximately half of the bottles and pay this doubled amount for each on average. Surprisingly, this model fared relatively well compared to some of the other models we tried.

For the next level of model complexity, we wanted to capture the notion that people are willing to pay a price approximately proportional to the utility that they would gain from the item. So, we devised a model that would bid exactly proportionally (and then double this proportion) for each bottle. In other words, the proportional strategy bid was

$$bid = 2 * current\_money * \texttt{util}(current\_bottle)/\texttt{util}(all\_remaining\_bottles)$$

Next, we wanted to account for the opponent's preferences and use information about what the opponent might bid to guide our agent's bids. The general idea is that for items where the opponent is going to bid more than the agent, the agent should bid more than they would originally plan to bid (but still less than the opponent would bid) to make the opponent have to pay more for the item. The opposite is true as well — the agent should bid less than they otherwise would (but more than the opponent will) to minimize the amount of money they will spend on the item. So, we implemented a model that, given an estimate for the opponent's preferences and a bid from the agent's preferences using proportional strategy, computed a bid that is between the two, contingent on a hyperparameter α dictating how aggressively the model should take the estimate into account. The bid for this 'meta-strategy' was as follows

$$bid = \alpha * \text{[estimated opponent bid]} + (1 - \alpha)\text{[bid proportional to utility]}$$

The last aspect of the model is the way in which the agent can estimate their opponent's preferences. This is done by inferring the opponent's latent preferences given their bids for all of the water bottles so far using WebPPL's `Infer` and `factor` functions to generate a distribution over possible preferences proportional to the probability of the observed bids under that preference, and finding the MAP estimate of this distribution.

For all of our experiments with the model, we used 20 water bottles to make the effects of different strategies clearer. We also generated random bottles every time, drawing features uniformly, to ensure the bottles we happened to have on hand didn't cause any bias. We were going to then use the exact bottles from each experiment and try to replicate them, but the models' performance on synthetic data was never satisfactory enough to advance to that stage.

6 Model Results

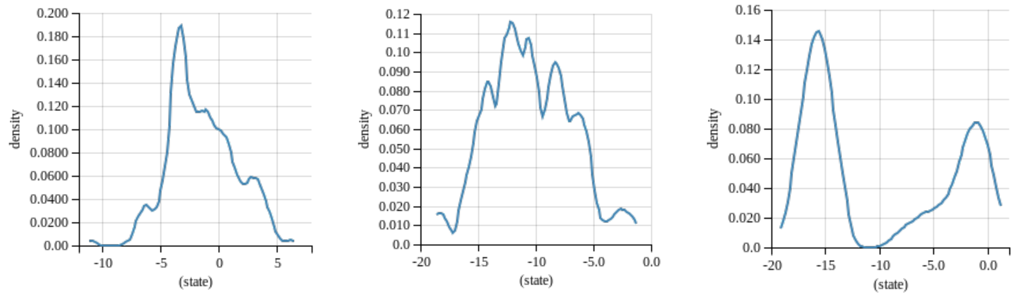

Figure 1: Proportional bidder v. constant bidder, 3 consecutive runs of 5000-sample MCMC with 20 random bottles. x-axis is the proportional bidder's final utility minus the constant bidder's.

Our most surprising result by far was the success of the simple, constant bidder. Figure 1 shows three examples of aggregated results from 5000 simulated auctions. The proportional bidder scores on average far worse, earning about 10 utility less per auction than the constant bidder. We suspect that the issue is the constant bidder getting too many water bottles almost for free when the proportional bidder bids low, and the proportional bidder wildly overpaying for bottles they do want — the constant bidder never pays more than 20 for their bottles, and the proportional bidder never pays less (until the end when the constant bidder runs out of money). These three distributions are representative of the variance we saw in the results of different simulations, which will be discussed in 6.2. Instability of results will be a recurring theme and unsolved problem for this model.

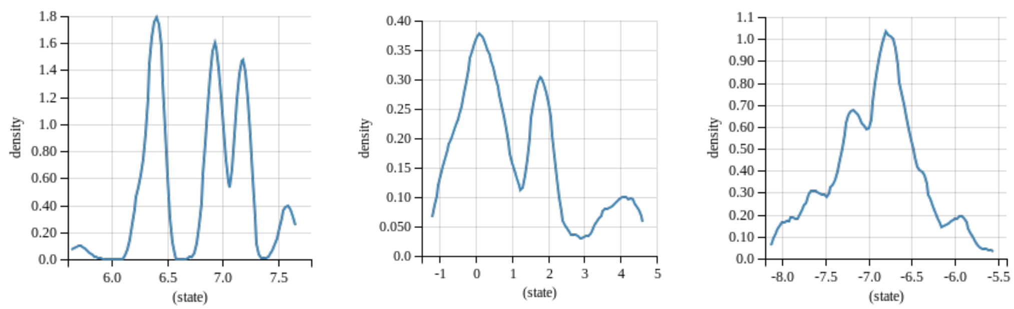

Figure 2: Proportional bidder v. proportional bidder, 3 runs of 5000-sample MCMC with 20 random bottles. x-axis is the difference in utility between player 1 and player 2.

As a sanity check, we show the results of pitting the proportional bidder against itself (with a different but balanced utility function) in Figure 2. Again the most immediate observation is the high variance in results. However, many runs show what one would expect, that the identical strategies are evenly matched.

Before moving to the full meta-bidder which infers the opponent's utility and adjusts its bets accordingly, we examine a partial implementation which does not have to do inference. This meta-bidder knows the opponent proportional bidder's utility function, and thus should be able to drastically outperform them. In Figure 3 we see this is indeed the case, albeit with the usual large variance in results.

6.1 Inference

Finally, we implement the full meta bidder which does inference over the opponent's utility. However, in testing the inference function alone, we find that discovery of utilities from bids is surprisingly hard, or at least not well suited to MCMC. Even with 10000 samples, and on bid sequences of 20 bottles, we could not infer a proportional bidder's utilities from their bids. The estimates were not reliably within 2 utility of ground truth for each category, which is particularly disappointing given the fact that 4 is within 2 of everything but 1 and 7. Note as well that there are only two degrees

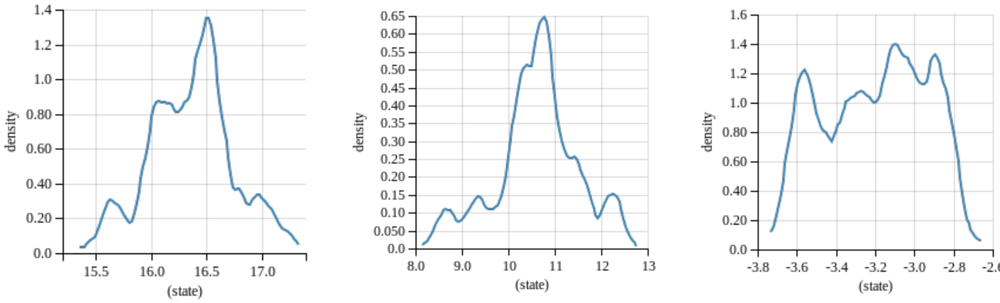

Figure 3: Meta bidder (knowing opponent's utility) v. proportional bidder, 3 runs of 5000-sample MCMC with 20 random bottles. x-axis is the meta bidder's final utility minus the proportional bidder's.

of freedom despite three possible utility values because we take advantage of the fact that the total utilities are balanced. That is, once we guess $O^{(j)}$ and $C^{(j)}$, we fill in $V^{(j)}$ as the most fair value knowing our utility. In retrospect, this space is likely small enough and MCMC unreliable enough that perhaps some sort of grid search with hill climbing would be better suited to the task.

6.2 Instability

Note that for three consecutive runs of the exact same thing the resulting distributions vary wildly and are sometimes even inverses of each other, as in the case of the second and third graphs of Figure 1. This is a bad sign for meaningful and reliable results. We suspect that this is a problem with the MCMC sampling method used. Upon inspection, the parameters only seem to change every 5-7 runs through, and when they do it's only one or two of the many sources of randomness (utility functions, bottles, etc.). Furthermore, the jaggedness of many of the distributions suggests MCMC getting stuck for many samples in a row without accepting a new point, which is a bad sign for approximation. This both slows burn-in and heavily skews the results towards an arbitrary staring point. Simulating for ten times as long, 50000 samples, took 10-15 minutes but did not show noticeable evidence of being more stable from run to run.

7 Comparison

The clearest difference between our model results and experimental results is the quality of inference about the opponent's utility function. The fact that humans were very accurate more often than not with only 8 bottles and that the model could not reliably get anywhere close with 20 bottles seems to indicate that humans are not doing anything like MCMC in their heads to solve this problem. More likely, we think, is that when people want to know what the other person wants they look at the bottles they've won so far and notice patterns in those bottles; all big or all opaque for example. This is quite distinct from using the actual bidding data that they had access to but did not seem to employ to compute some sort of likelihood of each bid given different utilities. Instead, they substituted an easier question, the heuristic of what categories that person seems to have been going for the hardest.

Our human trials also showed remarkable stability compared to our model's results, with the vast majority of final utilities falling within 5-10% of each other. This is in direct contrast to the model, which even after aggregating 5000 simulated auctions, reported wildly varying and extreme results for even identical strategies.

Those two large discrepancies in particular resoundingly reject our model, or at least this implementation of it. The proportional bidder may be a much too oversimplified approximation of what people do as well, because we noticed that people almost never bet less than 7 or more than 30, even when a bottle was nearly worthless or had all the characteristics they cared most about. They seemed to have a clipping mechanism which pulled them back to a mean of 15-20 which we did not include in our synthetic strategies. It could be that this is actually better than betting purely proportionally, and our model would have done better with such a mechanic.

8 Conclusion

When we began this project, we expected the proportional bidder and proportional meta bidder to be highly accurate approximations to human behavior and that most people would fall into one of those two modes of bidding. We also expected it to be a very logical and successful strategy, and were shocked to see it pushed around so decisively by the constant bidder.

The most interesting results ended up being how poorly the model reflected real life, and how much more time it took than humans only to come to a far inferior result. People are very intuitively talented at this auction game, and often impressively aware of their opponent's preferences after bidding on only 8 water bottles. The best next directions of inquiry to extend this result would be to revisit the proportional strategy in favor of a more moderate (but still proportional) bidder, and to improve inference on opponent utilities by using something other than MCMC over the utility values. A big unresolved question is why there was so much variation in results, likely due to poor MCMC mixing, and how to solve that distribution approximation problem.

Studying and modeling human behavior through simple competitive games such as this one may not have shown much success this time, but we suspect it could be a fruitful avenue of exploration in the future to learn about how people model other peoples' internal beliefs and desires, and what they do with those inferences.

Acknowledgments

Many heartfelt thanks to the TAs for spending so much time talking to us and working through our weekly new project ideas. Their guidance and advice was invaluable for both picking a final direction and being able to pursue it.

References

[1] Schwettmann, S., Tenenbaum, J., Kanwisher, N. (2018) Evidence for an Intuitive Physics Engine in the Human Brain. *Conference on Cognitive Computational Neuroscience*

[2] Baker, C.L., Saxe, R., Tenenbaum, J.B. (2009) Action understanding as inverse planning *Cognition*

9 Appendix

9.1 Experiment Bottle Features

Table 2: Experiment Bottle Features

Bottle #	Volume	Opaque	Pre-Filled
1	1500mL	N	N
2	750mL	N	Y
3	1150mL	N	N
4	1200mL	Y	Y
5	750mL	N	N
6	700mL	Y	Y
7	700mL	Y	N
8	700mL	N	Y

9.2 Experiment Handout

You have recently come upon an unexpected fortune ($80) given to you by a Nigerian Prince, and you find yourself a bidder in a special form of a blind auction, where the two participants each write down their bid hidden from the other. Then, whoever bid the highest gets the item, but only has to pay the mean of the two bids—so they still pay more than their opponent was willing to pay, but get to split the difference and not pay the maximum amount they are willing to pay.

You and your opponent bidder each have $80 to use, and you each have private preferences over three distinct features: volume, opacity, and contents. Your utility is as follows:

On a scale of 1 to 7, volume has an importance of 7 to you. This means you get 7 utility for the largest bottle (1500mL), and 1 utility from the smallest bottle (700mL), with the bottles in between scaled appropriately.
Opaque bottles give you an additional utility of 6, and only 1 if they are transparent.
If the bottle comes with water already in it, that gives you additional utility 2, and 1 if it does not (again the scale is 1 to 7).

Your score at the end of the auction is the sum of all the utilities you gained from each water bottle you won the bid for. Whoever has the highest score at the end, wins.

(Note that you only care about your net utility at the end of the auction; you do not care about any money that you have left over).

You do not know your opponent's preferences, but they are guaranteed to be balanced, i.e. give the same sum over all of the bottles as yours.

Each item is on display for you to inspect for the duration of the auction, and items are brought up one by one in a predetermined order for you and your opponent to bid on.