Inferring Private Valuations from Behavioral Data in Bilateral Sequential Bargaining

Lvye Cui, Haoran Yu*
School of Computer Science & Technology, Beijing Institute of Technology
cui_lvye@outlook.com, yhrhawk@gmail.com

Abstract
Inferring bargainers’ private valuations on items from their decisions is crucial for analyzing their strategic behaviors in bilateral sequential bargaining. Most existing approaches that infer agents’ private information from observable data either rely on strong equilibrium assumptions or require a careful design of agents’ behavior models. To overcome these weaknesses, we propose a Bayesian Learning-based Valuation Inference (BLUE) framework. Our key idea is to derive feasible intervals of bargainers’ private valuations from their behavior data, using the fact that most bargainers do not choose strictly dominated strategies. We leverage these feasible intervals to guide our inference. Specifically, we first model each bargainer’s behavior function (which maps his valuation and bargaining history to decisions) via a recurrent neural network. Second, we learn these behavior functions by utilizing a novel loss function defined based on feasible intervals. Third, we derive the posterior distributions of bargainers’ valuations according to their behavior data and learned behavior functions. Moreover, we account for the heterogeneity of bargainer behaviors, and propose a clustering algorithm (K-Loss) to improve the efficiency of learning these behaviors. Experiments on both synthetic and real bargaining data show that our inference approach outperforms baselines.

1 Introduction
With the growing popularity of e-commerce, the bilateral sequential bargaining mechanisms that allow for haggling between sellers and potential buyers have emerged on online platforms [Gujral, 2016; Huang et al., 2013]. One prerequisite for designing and optimizing these bargaining mechanisms is analyzing bargainers’ strategic behaviors. To achieve this target, it is important to infer bargainers’ private valuations on the negotiated items. For example, if a seller’s valuation on an item is known, it is easy to analyze whether the seller deceptively asks a high price during bargaining. Moreover, knowing the seller’s valuation helps quantify the efficiency and equality of the bargaining outcome.

Many prior works applied an equilibrium-based approach to infer agents’ private information (e.g., their valuations on items) [Athey and Haile, 2002; Jiang and Leyton-Brown, 2007; Athey and Nekipelov, 2010; Bajari et al., 2013]. The approach assumes that all agents are fully rational, each agent’s action is the best response to others’ actions, and their interactions always converge to a Nash equilibrium. It infers each agent’s private information by computing the inverse of his best response function given his observed behavior. However, this equilibrium-based approach cannot solve the valuation inference problem for real-world bargaining platforms. This is because the strong assumption that agents are fully rational and play equilibrium strategies may not hold in practice [Goeree and Holt, 2001; Kagel and Roth, 2020; Noti, 2021]. Bargaining platforms usually involve millions of bargainers, who have different capabilities to analyze incomplete information extensive-form games and distinct beliefs about their opponents’ information.

We focus on inferring bargainers’ private valuations based on their strategic behaviors on bilateral bargaining platforms, without relying on equilibrium assumptions. An example of these platforms is eBay’s Best Offer [Backus et al., 2020], where millions of sellers negotiate with buyers over their listed items. We use a thread to denote the price sequence generated from a seller-buyer pair’s negotiation over an item, and a thread may consist of multiple rounds. In each round,
given the opponent’s last offer, the seller and buyer can accept it, decline it, or make a counteroffer.

A seller may negotiate with multiple buyers over the price of an item (or different copies of the same item, e.g., coupon cards, commemorative stamps, artworks, and sports trading cards). The seller’s valuation on the item across these negotiations can usually be approximated as a constant, and we attempt to infer it by jointly analyzing the threads of these negotiations. Figure 1 shows an example, where a seller has three different threads after negotiating a book’s price with three buyers. Given the three thread records, we intend to infer the seller’s valuation on this book.

The main challenge of solving our problem is that sellers’ private valuations cannot be directly observed in the bargaining data. This implies that we do not have labels to guide the valuation inference process. We address the challenge and propose a Bayesian Learning-based Valuation Inference (BLUE) scheme. Our key idea is to define and derive feasible intervals of sellers’ private valuations according to their observed behaviors and leverage these feasible intervals to guide the inference of private valuations. Specifically, we make a mild rationality assumption that sellers do not choose strictly dominated strategies [Luce and Raiffa, 1989; Gopalkrishnan et al., 2018], e.g., sellers never propose or agree on prices lower than their valuations. This assumption enables us to utilize sellers’ behavior data to derive posterior feasible intervals of their private valuations, which will be used to guide the inference of these valuations.

Next, we briefly introduce our inference framework. First, we model each seller’s behavior function (which maps his bargaining history and private valuation to his decision) via a recurrent neural network. An accurate behavior function enables us to infer the seller’s valuation from his observed decision. Second, we learn these behavior functions (i.e., the parameters of the neural networks) using sellers’ behavior data. We derive feasible intervals of valuations according to the rationality assumption, and design a novel loss function based on the feasible intervals to train the neural networks. Our loss function ensures that the behaviors characterized by trained neural networks satisfy the rationality assumption and well fit the bargaining data. Third, we apply Bayes’ rule to derive the posterior distributions of sellers’ private valuations based on sellers’ behavior data and trained neural networks. In addition to above three steps, we consider the heterogeneity among sellers’ behaviors, and propose a clustering algorithm to enhance the learning of heterogeneous behavior functions.

We evaluate our method on two synthetic datasets and one large real-world dataset [Backus et al., 2020]. Compared with baselines, our method achieves more accurate and effective inference. Our contributions are two-fold:

- On the application side, we develop an effective valuation inference scheme for bilateral bargaining platforms.
- On the methodology side, we propose a novel framework for private information inference, through combining techniques from both economics and machine learning. Our framework is especially applicable to practical environments where agents cannot play equilibrium strategies due to their limited reasoning capabilities.

## 2 Related Work

### 2.1 Unknown Parameter Estimation in Games

Our work is related to the research that utilizes players’ observed actions to estimate unknown parameters of games (e.g., the payoff matrices in normal-form games). Some works tackled the estimation problem by assuming that players play according to Nash equilibria [Varian, 2007; Vorobeychik et al., 2007; Waugh et al., 2011; Blum et al., 2014; Honorio and Ortiz, 2015]. Some recent studies considered other equilibrium assumptions. For example, [Ling et al., 2018; Ling et al., 2019; Noti, 2021; Wu et al., 2022] utilized the quantal-response equilibria (QRE) or Nested Logit QRE to estimate the parameters in games (such as normal-form games and Stuckelberg games). These equilibrium-based estimation methods require strong cognitive and informational assumptions, e.g., players have full information of others’ private preferences. Such requirements are usually unsatisfied in practical bilateral bargaining, where bargainers have limited knowledge about others’ private information. In contrast, our inference scheme only requires bargainers not to choose strictly dominated strategies, and hence can handle more general human behavior in reality.

### 2.2 Private Information Inference in Auctions

Our work is also closely related to the literature that utilizes agents’ bidding data to infer their private information in auctions. [Athey and Haile, 2002; Jiang and Leyton-Brown, 2007; Athey and Nekipelov, 2010; Bajari et al., 2013] inferred bidders’ valuations on items according to their bids, by assuming that they are fully rational and play equilibrium strategies. An emerging line of research relaxes the full rationality assumption, and assumes that agents’ strategies can be captured by specific bounded rational behavior models. For example, [Noti and Syrgkanis, 2021; Nekipelov et al., 2015; Nisan and Noti, 2017] inferred the information (e.g., the value-per-click) of each bidder in sponsored search auctions, by assuming that each bidder is a no-regret learner and bids to minimize the regret. These inference schemes require a careful manual design of agents’ behavior models, and do not consider the heterogeneity of behaviors across agents. Moreover, they are mainly applicable to repeated auctions, which are distinct from online negotiations. By contrast, our scheme does not impose assumptions on the forms of agents’ behavior models. It can automatically characterize bargainers’ heterogeneous behaviors using neural networks, instead of manually designing specific unified models for them.

Other related studies include opponent modeling in automated negotiation [Hindriks and Tykhonov, 2008; Williams et al., 2011; Chen and Weiss, 2012; Baarslag et al., 2016; de Jonge, 2022] and preference elicitation [Boutilier et al., 2006; Guo and Sanner, 2010; Viappiani and Boutilier, 2010; Vendrov et al., 2020]. Different from these studies, our work focuses on price negotiation instead of multi-issue negotia-
3 Problem Formulation

In this section, we introduce the bilateral sequential bargaining, and describe the valuation inference problem.

3.1 Bilateral Sequential Bargaining

We use a bargaining thread to refer to the sequence of price offers made by a seller-buyer pair bargaining over an item. The thread consists of at most \( L \) bargaining rounds. Let \( s_0 \) denote the seller’s initial price offer. Given \( s_0 \), the buyer gives a counteroffer \( b_1 \) in the first round. Then, the seller and the buyer would make decisions in the following way:

- In round \( i = 1, 2, \ldots, L - 1 \), given the buyer’s price \( b_i \), the seller may (1) accept the offer, (2) make a counteroffer, or (3) decline the offer (i.e., insisting on his last offer \( s_{i-1} \)). If the seller accepts \( b_i \), the bargaining ends in round \( i \); otherwise, bargaining goes to the next round.
- In round \( i = 2, 3, \ldots, L - 1 \), the buyer can also accept, decline, or counter in response to the seller’s price \( s_{i-1} \). If the buyer accepts or declines \( s_{i-1} \), the bargaining terminates in round \( i \); otherwise, it comes to the next round.
- In the final round \( L \), the buyer accepts or declines the seller’s offer \( s_{L-1} \), and the bargaining ends.

This is known as the alternating offers protocol [Raiffa, 1982], and we show an example of \( L = 4 \) in Figure 2.

3.2 Private Valuation Inference Problem

We use \( v_m(q) \) to denote the private valuation of seller \( q \) for item \( m \). It is defined as the lowest price that seller \( q \) can accept for selling item \( m \). We consider a discrete \( v_m(q) \), since this lowest price is normally measured in dollars. This private value \( v_m(q) \) directly affects seller \( q \)’s bargaining behavior, as he seeks to sell item \( m \) with a price higher than \( v_m(q) \).

Let \( \{(x_{ih}(qm), y_{ih}(qm))\}_{i=1}^{I_h} \) denote the set of labeled data points about thread \( h \) of seller \( q \) for item \( m \), where \( I_h \) is the total number of rounds in thread \( h \). Here, \( y_{ih}(qm) \) denotes seller \( q \)’s decision in round \( i \) of thread \( h \) for item \( m \), and \( x_{ih}(qm) \) represents the sequence of all prior offers made before \( y_{ih}(qm) \) in thread \( h \). For example, considering the thread in Figure 2, when \( i = 1 \), there are two offers before the seller’s decision of \( s_1 \), which indicates that \( x_{1h}(qm) = (s_0, b_1) \) and \( y_{1h}(qm) = s_1 \). When \( i = 2 \), \( x_{2h}(qm) = (s_0, b_1, s_1, b_2) \) and \( y_{2h}(qm) = s_2 \).

We denote all threads that seller \( q \) has participated in to sell item \( m \) by \( H_m(q) = \{(x_{ih}(qm), y_{ih}(qm))\}_{i=1}^{I_h} \), where \( H_m \) is the number of threads related to item \( m \). Then, we characterize all sellers’ behaviors by thread dataset \( D_H = \{(H_m(q), M_q)\}_{q=1}^{Q} \), where \( Q \) is the number of sellers and \( M_q \) is the number of items that seller \( q \) bargains over. We describe our inference problem as follows:

**Problem 1.** Given thread dataset \( D_H \), we attempt to infer all sellers’ private valuations on items, i.e., \( \{v_m(q)|q\}_{m=1}^{M_q} \).

4 Private Valuation Inference Solution

We first explain our idea of solving Problem 1. Since a seller’s valuation is usually correlated with his bargaining behavior, we propose to model his behavior via a function \( f_{\theta}(q) \). The function (characterized by parameters \( \theta \)) maps seller \( q \)’s valuation \( v_m(q) \) and prior offers \( x_{ih}(qm) \) to his decision \( y_{ih}(qm) \). In particular, \( y_{ih}(qm) \) belongs to the set \{accept, decline, counter1, . . . , counterN\}, where set \{counter1, . . . , counterN\} includes all possible counter prices (e.g., measured in dollars). Once we learn this function, we can utilize Bayes’ rule to derive the posterior distribution \( \Pr(v_m(q)|Y_m(q); \lambda_q(q), \theta) \) for \( v_m(q) \):

\[
Pr(v_m(q)|Y_m(q); \lambda_q(q), \theta) = \frac{Pr(Y_m(q)|v_m(q), \lambda_q(q), \theta)Pr(v_m(q))}{\sum_{v_m(q)} Pr(Y_m(q)|v_m(q), \lambda_q(q), \theta)Pr(v_m(q))},
\]

where \( Pr(Y_m(q)) \) is the prior probability distribution of \( v_m(q) \), and \( \lambda_q(q) \) and \( Y_m(q) \) include the prior offers and seller behaviors of all threads related to item \( m \) and seller \( q \), respectively. Here, we assume independence among threads to derive the second equality of (1), which can reduce the computational complexity of our scheme. With \( Pr(v_m(q)|Y_m(q); \lambda_q(q), \theta) \), we can infer each \( v_m(q) \) by:

\[
v_m(q) = \arg \max_{v_m(q)} Pr(v_m(q)|Y_m(q); \lambda_q(q), \theta).
\]

Derivation of (1) requires an accurate behavior model \( f_{\theta}(q) \) to compute \( Pr(y_{ih}(qm)|v_m(q); x_{ih}(qm), \theta) \). To achieve this, we design a novel loss function to optimize \( \theta \) given thread dataset \( D_H \). Our loss function is defined based on a posterior feasible interval of \( v_m(q) \). Next, we introduce this feasible interval in Section 4.1, and study the optimization of \( \theta \) in Section 4.2. To simplify the explanation, we first assume all sellers have the same behavior pattern, i.e., \( \theta = \theta(q) \) for all \( q \in \{1, \ldots, Q\} \). In Section 4.3, we extend our solution to the case where sellers have distinct behavior patterns.
4.1 Feasible Interval of Valuation

In the economics and game theory literature, strictly dominated strategies refer to the strategies that always result in lower payoffs than alternative strategies regardless of other players’ strategies [Luc and Raiffa, 1989; Gopalan and et al., 2018; Qin et al., 2019; Maschler et al., 2020]. We make a rationality assumption that sellers never choose strictly dominated strategies. Specifically, if the bargaining ends with an agreement, the seller’s payoff equals the agreed price minus his valuation; otherwise, his payoff is zero. Under our rationality assumption, accepting the buyer’s offer implies that the seller’s valuation is no greater than the agreed price. This is because once the seller’s valuation is greater than the price, accepting becomes a strictly dominated strategy and the seller will not choose this decision. Formally, our rationality assumption is as follows.

**Assumption 1.** In the bargaining process: (i) if a seller accepts a buyer’s offer, the seller’s valuation is no greater than the price; (ii) if a seller declines the buyer’s offer in the last round, the seller’s valuation is no less than the price; (iii) a seller never proposes a price less than his valuation.

We define the feasible interval of a seller’s valuation as the set of all possible valuation values satisfying Assumption 1. Given thread dataset \(D_H\), we could derive the bounds \(c_m^{(q)}\) and \(d_m^{(q)}\) of the feasible interval of valuation \(v_m^{(q)}\) as follows:

\[
\begin{align*}
\{c_m^{(q)}\} &= \max\left\{c_{m1}^{(q)}, c_{m2}^{(q)}, \ldots, c_m^{(q)}, \ldots, c_{MHm}^{(q)}\right\}, \\
\{d_m^{(q)}\} &= \min\left\{d_{m1}^{(q)}, d_{m2}^{(q)}, \ldots, d_m^{(q)}, \ldots, d_{MHm}^{(q)}\right\}.
\end{align*}
\]

Here, we calculate \([c_m^{(q)}, d_m^{(q)}]\) as the intersections of the feasible intervals of all \(H_m\) threads. For thread \(h\), we calculate its bounds \(c_m^{(q)}\) and \(d_m^{(q)}\) according to seller \(q\)’s decision \(y_{ih}^{(qm)}\) in last round \(I_h\): (i) When \(y_{ih}^{(qm)} = 0\), we have \(c_m^{(q)} = d_m^{(q)} = y_{ih}^{(qm)}\); (ii) Otherwise, we have \(c_m^{(q)} = 0\) and \(d_m^{(q)} = \min\left\{y_{ih}^{(qm)}, y_{ih}^{(qm)}, \ldots, y_{ih}^{(qm)}\right\}\).

### 4.2 Learning of Homogeneous Behavior

Given thread dataset \(D_H\) and the set of our derived bounds \(\{c_m^{(q)}, d_m^{(q)}\}_{m=1}^{M_H}\), we design a novel loss function to learn \(\theta\). Our loss function consists of the following two parts:

- The first part is related to the rationality assumption.

Given \(\theta\), we can infer the seller’s valuation \(v_m^{(q)}\). With accurate \(\theta\), the inferred \(v_m^{(q)}\) will fall into its feasible interval \([c_m^{(q)}, d_m^{(q)}]\). Thus, our loss function’s first part is the following negative log-likelihood function:

\[
\sum_{q=1}^{Q} \sum_{m=1}^{M_H} \sum_{h=1}^{I_h} \log \text{Pr}(c_m^{(q)} \leq v_m^{(q)} \leq d_m^{(q)} | y_{ih}^{(qm)}, x_{ih}^{(qm)}, \theta).
\]

Here, \(\text{Pr}(c_m^{(q)} \leq v_m^{(q)} \leq d_m^{(q)} | y_{ih}^{(qm)}, x_{ih}^{(qm)}, \theta)\) is the probability that the inferred valuation \(v_m^{(q)}\) belongs to the feasible interval \([c_m^{(q)}, d_m^{(q)}]\), given the prior offers \(x_{ih}^{(qm)}\), the seller decision \(y_{ih}^{(qm)}\), and the parameters \(\theta\).

### 4.3 Extension for Heterogeneous Behavior

In the presence of heterogeneous behavior, a natural approach is to learn a behavior model for each seller utilizing only his thread data. However, each seller may only have a small number of bargaining threads, which may be insufficient for behavior learning. Considering that many sellers have similar behavior patterns, we propose an algorithm to cluster sellers, as shown in Algorithm 2.

Specifically, we partition sellers into \(K\) clusters (denoted by \(\{C_k\}_{k=1}^{K}\)), and the sellers in \(C_k\) are assumed to have the same behavior pattern. Our algorithm iterates between two
\[ V^{(1)}_{\theta_t} = - \sum_{q=1}^{Q} \sum_{m=1}^{M_q} \sum_{h=1}^{H_m} \sum_{i=1}^{I_h} \partial \log \Pr \left( c_m^{(q)} \leq v_m^{(q)} \leq d_m^{(q)} \middle| y_{ih}^{(qm)}, x_{ih}^{(qm)}, \theta_t \right) \frac{\partial \theta_t}{\theta_t} \]

\[ \sum_{q=1}^{Q} \sum_{m=1}^{M_q} \sum_{h=1}^{H_m} \left( \sum_{v_m^{(q)}} \Pr \left( y_{ih}^{(qm)} \middle| v_m^{(q)}, x_{ih}^{(qm)}, \theta_t \right) \Pr \left( v_m^{(q)} \middle| \theta_t \right) - \frac{\partial \Pr \left( y_{ih}^{(qm)} \middle| v_m^{(q)}, x_{ih}^{(qm)}, \theta_t \right)}{\partial \theta_t} \Pr \left( v_m^{(q)} \middle| \theta_t \right) \right) \]


\begin{algorithm}
\caption{K-Loss Clustering Algorithm}
\begin{algorithmic}
\State \textbf{Input:} Thread dataset \( D_K \), cluster number \( K \), maximum epoch times \( E \).
\State \textbf{Output:} \( \{C_1, \ldots, C_K\} \).
\Statex \hspace{1em} 1: \text{Randomly assign sellers into } K \text{ clusters } \{C_k\}_{k=1}^K.
\Statex \hspace{1em} 2: \text{Let } e = 0.
\Statex \hspace{1em} 3: \text{while } e \leq E \text{ and the clustering does not converge do}
\Statex \hspace{1em} \hspace{1em} 4: \text{for } k = 1, \ldots, K \text{ do}
\Statex \hspace{1em} \hspace{1em} \hspace{1em} 5: \text{Learn behavior function } g_{\theta}^{(C_k)}, \text{ only based on the}
\Statex \hspace{1em} \hspace{1em} \hspace{1em} \hspace{1em} \hspace{1em} \text{thread data related to the sellers in cluster } C_k.
\Statex \hspace{1em} \hspace{1em} \hspace{1em} 6: \text{end for}
\Statex \hspace{1em} \hspace{1em} 7: \text{Re-assign all sellers according to the prediction losses}
\Statex \hspace{1em} \hspace{1em} \hspace{1em} \hspace{1em} \hspace{1em} \text{of } K \text{ learned models on the seller thread data, i.e.,}
\Statex \hspace{1em} \hspace{1em} \hspace{1em} \hspace{1em} \hspace{1em} \text{assign seller } q \text{ to cluster } C_k \text{ if } g_{\theta}^{(C_k)} \text{ has the smallest}
\Statex \hspace{1em} \hspace{1em} \hspace{1em} \hspace{1em} \hspace{1em} \text{prediction loss on his bargaining behavior data.}
\Statex \hspace{1em} \hspace{1em} 8: \text{end while}
\Statex \hspace{1em} 9: \text{return } \{C_1, \ldots, C_K\}
\end{algorithmic}
\end{algorithm}

steps: (i) learning the behavior model of sellers in each cluster, and (ii) updating the assignment of sellers to \( K \) clusters based on the fitness between the sellers’ behavior data and the learned behavior models. Particularly, we use \( g_{\theta}^{(C_k)} \) to denote the behavior pattern of sellers in cluster \( C_k \). This \( g_{\theta}^{(C_k)} \) is similar to \( f_{\theta} \) and can be learned via Algorithm 1. Since we are mainly concerned with the clustering performance in Algorithm 2, one approach to accelerating the learning of \( g_{\theta}^{(C_k)} \) in line 5 is that we assume uniform valuations across sellers and simplify \( g_{\theta}^{(C_k)} \) as a mapping from seller’s bargaining thread histories to seller’s decision.

After clustering sellers into \( K \) clusters, we further learn their behavior functions (captured by \( f_{\theta}: (v_m^{(q)}, x_{ih}^{(qm)}) \rightarrow y_{ih}^{(qm)}, \forall q \in C_k \) via Algorithm 1. Then, we infer their private valuations utilizing learned functions based on (1) and (2).

5 Experiments

5.1 Dataset Description

We generate synthetic data by simulating the process of bilateral bargaining (depicted in Figure 2). To model the diversity of sellers’ and buyers’ behaviors, we consider three kinds of patterns, i.e., simple, random and payoff-maximization.

- For the simple pattern, an agent accepts the opponent’s offer as long as he can obtain a non-negative payoff.
- For the random pattern, an agent first estimates its opponent’s private valuation, and then randomly takes actions from those that give it a non-negative expected payoff.
- For the payoff-maximization pattern, an agent chooses the action that maximizes his payoff, given his belief about the opponent. Specifically, the agent first updates his belief about the opponent’s valuation based on the opponent’s prior actions. Then, the agent reasons about the opponent’s potential response in the new round, and chooses his action to maximize his expected payoff.

We synthesize two thread datasets (i.e., synthetic data I and synthetic data II), based on different distributions of agents’ valuations. In synthetic data I and synthetic data II, all agents’ valuations follow uniform and categorical distributions, respectively. We let the support of the valuation distribution be \( \{10, 14, \ldots, 98\} \). In each of synthetic data I and synthetic data II, we simulate 900 sellers (with bargaining behaviors uniformly chosen among the three patterns) and there are about 120,000 bargaining threads.

We also conduct experiments on a large dataset collected from eBay’s Best Offer platform [Backus et al., 2020]. It contains concrete bargaining information, e.g., seller ID, item ID, thread ID, and each thread’s complete history. If a seller only has a limited number of decisions recorded, it is nearly impossible to infer his valuations on different items. Thus, we filter out these sellers, and our real dataset includes 32,099 sellers with 358,641 threads.

5.2 Experiment Settings

We compare six learning-based valuation inference methods (in Appendix A, we show the results of two sampling-based inference methods):

- **BLUE** (**B**ayesian **L**earning-based **V**aluation **I**nference): We assume all sellers have homogeneous behaviors, and implement Algorithm 1 for valuation inference.
- **BLUE-C** (**B**LUE with **C**lustering): We first cluster sellers via Algorithm 2, and then make inference for the sellers in each cluster via Algorithm 1.
- **DL** (**D**ual **L**earning): Dual learning [Qin, 2020; He et al., 2016] is a learning framework that utilizes the primal-dual structure of AI tasks to advance the learning
Table 1: Clustering on Two Synthetic Datasets: D-I and D-II represent synthetic data I and synthetic data II, respectively.

<table>
<thead>
<tr>
<th>Methods/Datasets</th>
<th>G1 Precision</th>
<th>G1 Recall</th>
<th>G2 Precision</th>
<th>G2 Recall</th>
<th>G3 Precision</th>
<th>G3 Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K$-Means/D-I</td>
<td>0.647 ± 0.088</td>
<td>0.629 ± 0.156</td>
<td>0.480 ± 0.109</td>
<td>0.479 ± 0.067</td>
<td>0.543 ± 0.072</td>
<td>0.809 ± 0.209</td>
</tr>
<tr>
<td>AE-Image/D-I</td>
<td>0.698 ± 0.084</td>
<td>0.755 ± 0.118</td>
<td>0.736 ± 0.088</td>
<td>0.666 ± 0.136</td>
<td>0.993 ± 0.008</td>
<td>0.979 ± 0.027</td>
</tr>
<tr>
<td>$K$-Loss/D-I</td>
<td>0.992 ± 0.006</td>
<td>0.998 ± 0.002</td>
<td>0.998 ± 0.002</td>
<td>0.989 ± 0.011</td>
<td>1.000 ± 0.000</td>
<td>1.000 ± 0.000</td>
</tr>
<tr>
<td>$K$-Means/D-II</td>
<td>0.994 ± 0.085</td>
<td>0.300 ± 0.191</td>
<td>0.467 ± 0.077</td>
<td>0.657 ± 0.080</td>
<td>0.730 ± 0.086</td>
<td>0.952 ± 0.028</td>
</tr>
<tr>
<td>AE-Image/D-II</td>
<td>0.701 ± 0.042</td>
<td>0.776 ± 0.079</td>
<td>0.763 ± 0.044</td>
<td>0.661 ± 0.084</td>
<td>0.976 ± 0.015</td>
<td>0.993 ± 0.007</td>
</tr>
<tr>
<td>$K$-Loss/D-II</td>
<td>0.995 ± 0.003</td>
<td>0.993 ± 0.002</td>
<td>0.998 ± 0.002</td>
<td>0.992 ± 0.003</td>
<td>1.000 ± 0.000</td>
<td>1.000 ± 0.000</td>
</tr>
</tbody>
</table>

5.3 Experiment Results

Heterogeneous Behavior Clustering

We evaluate the clustering performance of different methods on two synthetic datasets, where we know the ground truth of different sellers’ behavior patterns. Recall that in both synthetic data I and synthetic data II, we simulate 900 sellers whose behaviors are randomly chosen among three pattern groups. We denote the three pattern groups by G1, G2, and G3. Table 1 summarizes the clustering precisions and recalls of three clustering approaches after five runs of experiments on G1, G2, and G3. Our K-Loss achieves the best performance in all cases, and can accurately assign the sellers with the same bargaining pattern into the same cluster.

Valuation Inference on Synthetic Data

On the synthetic datasets, we characterize the valuation inference performance by the mean squared error (MSE), where the error of an inferred valuation is its squared distance to the actual valuation. Figure 3 shows the convergence performance of different inference schemes. It plots the MSEs achieved on synthetic validation data after running these schemes for different numbers of iterations. We test six random seeds, and the shadows in Figure 3 indicate the standard deviations of MSEs. Our BLUE-C and BLUE converge to much lower MSEs than other schemes on both datasets.

Figure 4 presents different methods’ average MSEs after six runs of experiments on the testing data of two synthetic datasets. BLUE-C achieves the lowest MSE (less than 15.0) on both datasets, and other methods have relatively large MSEs (with the smallest MSE being 72.5). BLUE also outperforms all comparison methods on both datasets. We also

Implementation Details

For both synthetic data and real data, we randomly select 80% of all threads for training, 10% for validation, and 10% for testing. We model each behavior function $f_\theta$ via a gated recurrent unit (GRU). The Adam optimizer with a learning rate of 0.001 is applied for our network training. The epoch number $T$ is set to 500 with a batch size of 64, and the weight factor $\alpha$ is set to 0.6. The source code and data are available at: https://github.com/cuiyve/Bargaining-project.

5.3 Experiment Results

Heterogeneous Behavior Clustering

We evaluate the clustering performance of different methods on two synthetic datasets, where we know the ground truth of different sellers’ behavior patterns. Recall that in both synthetic data I and synthetic data II, we simulate 900 sellers whose behaviors are randomly chosen among three pattern groups. We denote the three pattern groups by G1, G2, and G3. Table 1 summarizes the clustering precisions and recalls of three clustering approaches after five runs of experiments on G1, G2, and G3. Our K-Loss achieves the best performance in all cases, and can accurately assign the sellers with the same bargaining pattern into the same cluster.

Valuation Inference on Synthetic Data

On the synthetic datasets, we characterize the valuation inference performance by the mean squared error (MSE), where the error of an inferred valuation is its squared distance to the actual valuation. Figure 3 shows the convergence performance of different inference schemes. It plots the MSEs achieved on synthetic validation data after running these schemes for different numbers of iterations. We test six random seeds, and the shadows in Figure 3 indicate the standard deviations of MSEs. Our BLUE-C and BLUE converge to much lower MSEs than other schemes on both datasets.

Figure 4 presents different methods’ average MSEs after six runs of experiments on the testing data of two synthetic datasets. BLUE-C achieves the lowest MSE (less than 15.0) on both datasets, and other methods have relatively large MSEs (with the smallest MSE being 72.5). BLUE also outperforms all comparison methods on both datasets. We also
notice that clustering-based inference methods perform better than those without clustering, e.g., DL-C is better than DL.

In addition to MSEs, we compare different inference methods in terms of the fractions of inferred valuations satisfying the rationality assumption. Specifically, we define Rationality-Compatible Inference Rate (RCIR) as the percentage of inferred valuations that belong to the feasible intervals derived based on rationality assumption. Intuitively, a method achieving a higher RCIR can infer valuations more accurately. Figure 5 shows the average RCIRs of different methods after six runs of experiments on the testing data of synthetic datasets. The RCIRs of our schemes (BLUE-C and BLUE) are above 90\%, and those of other methods are below 67\%. This result is consistent with Figure 4, and implies that RCIR is a good indicator of the inference performance.

**Valuation Inference on Real Data**

Our real data from eBay does not record sellers’ actual valuations. Hence, we cannot directly compare the MSEs of different methods on the real data. Instead, we compare different methods using two criteria: the RCIR and the ratio of satisfying the secret bounds (as defined later).

In order to justify our choice of the recurrent neural network (GRU), we investigate the performance of Transformer under our BLUE-C scheme. Figure 6a presents the RCIRs of the two neural networks at different iterations on the real validation data. They converge to similar RCIRs. Since Transformer has more learning parameters, its training is much slower than GRU. Hence, we employ GRU to implement all inference schemes when performing other experiments.

Figure 6b shows the RCIRs of different inference schemes at different iterations on the real validation data. We observe that our BLUE-C and BLUE converge to much higher RCIRs than other schemes.

We show the performance of different methods on real testing data. Figure 7a presents different methods’ RCIRs. Our BLUE-C achieves the highest RCIR (i.e., 89.0\%). In Figure 7b, we compare different methods using a new criterion. To enable auto-declining and auto-accepting of offers, each seller can inform eBay of an upper bound and a lower bound of its private valuation. For example, if a buyer’s offer exceeds the upper bound, eBay will accept the offer on behalf of the seller. We call these secret bounds. They are also recorded in eBay’s dataset, and are closely related to the sellers’ actual private valuations. Figure 7b shows the ratio of inferred valuations satisfying these secret bounds. Our BLUE-C still achieves the highest ratio (i.e., 77.0\%), while the ratios under all baselines are no greater than 66.5\%.

### 6 Conclusion

In this paper, we proposed a novel framework for inferring bargainers’ private valuations on items in bilateral sequential bargaining. Our inference framework works in general scenarios where bargainers may not play equilibrium strategies due to their limited reasoning capabilities and knowledge. It can automatically characterize bargainers’ behaviors through neural networks rather than manually designed models. Our future research will explore the theoretical foundations of our methods and extend them to account for time-discounted valuations. Another extension is generalizing our framework to multilateral bargaining, where an agent negotiates the price of a homogeneous product with different opponents simultaneously to elicit a more favorable offer [Thomas and Wilson, 2002; Thomas and Wilson, 2014].

### A Comparison with Sampling-Based Methods

We consider two sampling-based methods: Random: selecting random values from the feasible intervals; Mean: using the means of the bounds of feasible intervals. We evaluate them based on the mean squared errors on two synthetic datasets (D-I, D-II) and the ratios of satisfying the secret bounds on real data (D-R). Table 2 shows that our scheme outperforms these two methods.

---

**Table 2: Comparison with Sampling-Based Methods.**

<table>
<thead>
<tr>
<th></th>
<th>Random</th>
<th>Mean</th>
<th>BLUE-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-I</td>
<td>66.64 ± 2.81</td>
<td>41.22 ± 2.17</td>
<td>12.97 ± 1.98</td>
</tr>
<tr>
<td>D-II</td>
<td>69.89 ± 4.43</td>
<td>42.89 ± 1.73</td>
<td>14.71 ± 3.24</td>
</tr>
<tr>
<td>D-R</td>
<td>63.54% ± 0.33%</td>
<td>65.49% ± 0.60%</td>
<td>77.03% ± 0.97%</td>
</tr>
</tbody>
</table>

---

**Figure 5: RCIRs of Six Methods on Synthetic Testing Data.**

**Figure 6: Convergence of RCIRs on Real Validation Data.**

**Figure 7: Comparison Between Methods on Real Testing Data.**
Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 62202050) and the Beijing Institute of Technology Research Fund Program for Young Scholars.

References


