Opponent Modeling in Negotiation Dialogues by Related Data Adaptation

Anonymous ACL submission

Abstract

Opponent modeling refers to the task of inferring another party's mental state within the context of non-collaborative social tasks. In a negotiation, it involves identifying the opponent's priorities, which is crucial for finding high-value deals. Discovering these priorities is helpful for automated negotiation systems deployed in pedagogy and conversational AI. In this work, we propose a transformer-based ranker for identifying these priorities from negotiation dialogues. The model takes in a partial dialogue as input and predicts the priority order of the opponent. We further devise ways to adapt related data sources for this task to provide more explicit supervision for incorporating the opponent preferences and offers, as a proxy to relying on granular utterance-level annotations. We show the utility of our proposed approach through extensive experiments based on two dialogue datasets. We particularly find that the proposed data adaptations lead to strong performance in 0-shot and few-shot scenarios. Moreover, they allow the model to perform better with access to fewer utterances from the opponent.

1 Introduction

Automated negotiation is an interesting and a challenging domain for AI research. Negotiations are key to our everyday interactions like allocating available resources, salary decisions, business deals, and legal proceedings. Being an effective negotiator is critical for automated systems deployed in complex social scenarios (Gratch et al., 2015). Such agents can make social skills training more accessible (Johnson et al., 2019a) and advance conversational AI (Leviathan and Matias, 2018).

The priorities of the opponent are typically unknown to a negotiator beforehand. Prior work argues that understanding what the opponent wants is one of the key aspects of successful negotiations (Baarslag et al., 2013). For instance, consider the scenario from the CaSiNo dataset (Chawla et al., 2021) - two participants role-play as campsite neighbors and negotiate to divide food, water, and firewood packages among each other. It is useful for a dialogue agent to know which kind of packages its opponent prefers. An accurate *model* of the opponent can enable the agent to roll out offers that work for both parties - which has implications on both the objective performance such as the points scored and subjective outcomes such as opponent satisfaction and likeness towards the agent. This can also aid in pedagogy by allowing the agent to provide concrete feedback to students who fail to incorporate the preferences of their opponents (Johnson et al., 2019b). Discovering these priorities from the interaction with the opponent is usually referred to as *Opponent Modeling* in the context of negotiations.

Most efforts in automated negotiations use highly structured communication such as in agent-agent interactions (Williams et al., 2012) and human-agent interactions based on button clicks (Mell and Gratch, 2017). Hence, the opponent models in these scenarios use frequency or Bayesian estimates by combining the structured information received from the opponent such as their explicit preferences and offers. However, this becomes non-trivial for more realistic chat-based interactions where the information is far less structured (Lewis et al., 2017; He et al., 2018).

To alleviate this problem, opponent modeling approaches for negotiations in natural language involve the collection of additional utterance-level annotations to convert the preferences and offers into a more structured format (Nazari et al., 2015), that can then be used with frequency-based methods. Unfortunately, this is expensive, requires expertise, and hurts generalizability. Further, these annotations are unavailable for agents that are deployed to end users, needing a separate NLU module which can potentially lead to error propagation in the downstream dialogue system pipeline.

To this end, we explore opponent modeling in negotiation dialogues without relying on additional utterance-level annotations. We formulate the task of opponent modeling as ranking a fixed set of issues based on a partial input dialogue, proposing a transformer-based hierarchical architecture for the same. To provide more explicit supervision for capturing preferences and offers expressed by the opponent, we devise simple and effective ways to project related data sources to this task. As opposed to multi-task learning that typically involves task-agnostic and task-specific parameters and back-to-back fine-tuning procedures that suffer from catastrophic forgetting issues, our adaptation augments the training data available to the model, allowing end-to-end joint learning and parameter sharing. We summarize our contributions below:

1. We formulate opponent modeling as a ranking task (Section 3) and propose a transformer-based model that can be trained directly on partial dialogues using a pairwise margin ranking loss (Section 4).

2. To better capture the information present in the preference and offer statements of the opponent, we adapt related data sources resulting in more labeled data for training (Section 4).

3. We define evaluation metrics by taking inspiration from prior work in negotiations, along with Dialog State Tracking and Learning-to-Rank research in NLP (Section 5). We perform experiments and analysis based on two dialogue datasets in English - CaSiNo (Chawla et al., 2021) and DealOrNoDeal (Lewis et al., 2017) with a primary focus on CaSiNo, showing the utility of the proposed methodology (Section 6).

4. We study the scope for improvement by comparing our best-performing model to a human expert, discussing common errors to guide future work (Section 6), and laying out the implications for research in human-machine negotiations (Section 9).

2 Task Overview

We start by describing a common and useful abstraction for studying negotiations in scientific literature, known as the multi-issue bargaining task or MIBT (Fershtman, 1990). The negotiations that we focus on are based on this abstraction. Consider two negotiators P_1 and P_2 who negotiate over m issues: $I_1, I_2, \dots, I_m$. Each issue I_i is associated with a total number of items T_i. The goal is to reach an agreement by dividing all the items for every issue such that each item is assigned to one of the negotiators. Assume P_p receives a reward r_{pi} for every item of an issue I_i. The primary objective is then to maximize the total reward R_p:

$$R_p = \sum_{i=1}^{m} r_{pi} t_i, \quad (1)$$

where $t_i \in \{0, 1, 2, \dots, T_i\}$ is the number of items of issue I_i that P_p is able to negotiate for, at the end. In several realistic applications, other subjective goals can be important as well such as the overall satisfaction of the opponent and their liking of the partners once the negotiation is over.

The rewards r_{pi} are based on a priority order that is defined before the negotiation. Typically, these are unknown to the opponent. The task of opponent modeling involves discovering these priorities from the interaction. Our focus is on two datasets: CaSiNo (Chawla et al., 2021) and DealOrNoDeal (Lewis et al., 2017), both based on this MIBT design. CaSiNo is grounded in a camping scenario, containing negotiations over three issues: *food*, *water*, and *firewood*, while DealOrNoDeal involves three arbitrarily-defined issues: *books*, *hats*, and *balls*. Our main goal is to perform opponent modeling for CaSiNo. To this end, we adapt DealNoDeal along with the available metadata in CaSiNo for data augmentation.

3 Problem Formulation

We define the problem from the *perspective* of a specific negotiator (referred to as *self*, hereafter), and aim to model the priorities of the *opponent*. Assume that a conversation C contains an alternating sequence of N utterances between the negotiator self S and the opponent O. The partial conversation is C_k, which is obtained after S observes k utterances from the opponent[1]. The goal is to predict the priority order of the opponent over a predefined set of issues, for each possible value of k. Specifically, we build the model M, with $Y_O = M(C_k)$, where Y_O is the desired priority order of the opponent.

Our motivation for training with partial input dialogues comes from downstream applications in

[1] C_k will contain either $2k$ or $2k-1$ utterances, depending on who starts the conversation.

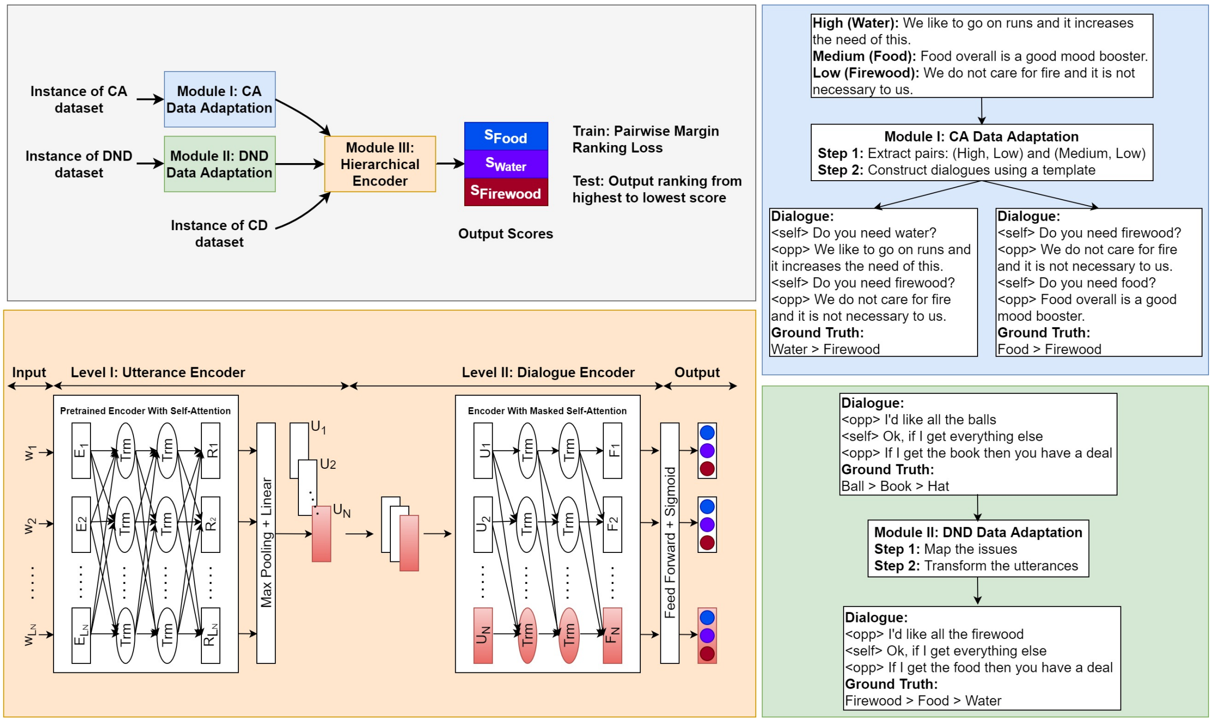


Figure 1: An overview of our proposed methodology for opponent modeling in negotiations dialogues. Module I: Adaptation of the CaSiNo Arguments (CA) dataset, Module II: Adaptation of DealOrNoDeal (DND) dataset, Module III: Hierarchical encoder that encodes dialogues constructed from all the datasets in the same manner and is trained to generate a score for all the possible issues after each opponent utterance - used during inference to output the opponent priority order.

conversational AI, where it is useful to build an accurate opponent model early on in the conversation. However, the performance with the complete dialogue is also useful for other applications such as pedagogy (Johnson and Gratch, 2021). Hence, in our experiments, we consider metrics that measure the performance in both these cases (Section 6).

4 Methodology

We design a transformer-based ranker that leverages pretrained language models and builds contextual representations of the input utterances in a hierarchical manner. These representations are then used to predict the priority order of the opponent. We first describe this core model, followed by our data augmentation techniques. We provide an overview of our approach in Figure 1.

4.1 Hierarchical Ranking Model

Utterance Encoder: First, a sentence-level module (Level I) encodes each utterance $U_j = [w_1, w_2, \ldots, w_{L_j}]$ separately. We prepend the utterances with a special token to indicate the author: <self> or <opp>. To encode a contextually-rich representation, our level I encoder utilizes pretrained language models (Devlin et al., 2019; Liu et al., 2019), given their success across a wide range of NLP tasks, especially in low resource settings on similar NLU tasks (Balaraman et al., 2021). The pretrained model outputs d-dimensional word representations $W \in \mathbb{R}^{L_j \times d}$, which are then pooled to get the utterance representation $U_j \in \mathbb{R}^d$. The Level I output is essentially the conversation matrix $U \in \mathbb{R}^{N \times d}$, which is obtained after processing all the input utterances.

Dialogue Encoder: Here, we utilize a transformer block with masked self-attention (Vaswani et al., 2017). Self-attention enables efficient interactions for encoding partial conversations. A target utterance is only allowed to utilize the information from previously-seen utterances, which is accomplished by masking all the future utterances in the dialogue. In a single transformer layer, each target utterance *query* simultaneously assesses and encodes the information from all the unmasked *key* utterances, resulting in a contextualized representation of each utterance - the matrix $F \in \mathbb{R}^{N \times d}$.

Output Layers: Finally, a feed-forward network acts on F to output an m-dimensional representation for each utterance. This represents the scores for each of the issues that the model is trying to rank. We then apply the sigmoid operation to constrain each score between 0 and 1, resulting in the output $O \in \mathbb{R}^{N \times m}$.

Note that the value of m or the number of issues is *small* and *fixed* in the negotiations that we consider in this paper. This allows us to predict the scores for each of the issues together, unlike several text ranking tasks in the literature where each item is ranked separately (Yates et al., 2021).

Training: We employ the pairwise margin ranking loss to train our model in an end-to-end manner. The model is trained to make a prediction every time it encounters a new utterance from the opponent. The loss $\mathcal{L}_k$ after observing k utterances from the opponent is defined as:

$$\mathcal{L}_k = \sum_{q=(q_1,q_2)\in Q} L_k(o^k_{q_1}, o^k_{q_2}, y_q), \qquad (2)$$

where L_k is given by:

$$L_k(o^k_{q_1}, o^k_{q_2}, y_q) = max(0, -y_q(o^k_{q_1} - o^k_{q_2}) + c). \qquad (3)$$

Q represents the set of all possible pairs of issues. $o^k_{q_1}$ and $o^k_{q_2}$ are the predicted scores from the final layer of the hierarchical ranker after applying the sigmoid operation. y_q captures the ground truth ranking between q_1 and q_2. y_q is equal to $+1$ when q_1 should be ranked higher (has a larger score) than q_2 and it is kept as -1 otherwise. c is the margin.

The ranking loss trains the model to predict a *higher* score for the issue that is ranked *higher* by the ground truth priority order. A positive margin of c ensures a nonzero loss if the score for the higher ranked item is *not greater than or equal to its counterpart by* c, forcing the model to predict well-separated boundaries. We experimented with different values for c, concluding that a nonzero margin is necessary for any meaningful training.

Inference: Once the model is trained, the predicted scores can be used to output the desired ranking order for a given input dialogue. The model simply outputs the ranking of the issues by ordering them according to these predicted scores.

The pairwise ranking loss was chosen for its suitability and simplicity. We note, however, that other potential alternatives exist. Since the number of issues is limited, one can remodel the prediction task as classification over all the possible orderings. However, this trivially does not capture that although two orderings can be wrong, one can be *somewhat less* wrong than the other. Hence, a ranking loss is more suitable in giving a smoother signal to the model during training, leading to a better performance in our initial experiments. We also explored applying more complicated ranking loss functions and even leveraging a sequence-to-sequence model to directly generate the sequence of issues in their correct ranking order (Yates et al., 2021) - we instead found the pairwise ranking loss to be effective and simple for our approach in this paper that involves a limited set of issues and exploits partially-masked loss functions (next Section). Regardless, we encourage future work to explore these other formulations as well depending on the task at hand.

4.2 Data Adaptations

Information about the opponent's priorities can primarily be gathered from their preference and offer statements. Sharing preferences by explicitly mentioning *'We need water'* or more implicitly - *'We like to go on runs'* can provide information that water is of high priority to the negotiator. Further, offers such as *'I would like all the food'* can imply that food is preferred. Instead of relying on additional annotations, we now describe an alternate way to better capture the preferences and offers in our hierarchical ranking model. We achieve this by adapting two additional data sources for this task, allowing the data to be directly added to the primary training dataset and enabling end-to-end parameter sharing between these related tasks.

Capturing preferences: In order to provide more direct supervision for the preferences, we leverage the metadata from CaSiNo. The dataset involves a preparation phase for all the participants before their actual negotiation. Each participant is randomly assigned a preference order (a permutation of {**H**igh, **M**edium, **L**ow}) for the three issues (Food, Water, Firewood). They are then asked to come up with arguments from their personal experiences as to why they would need or not need items of a particular issue. One such example is given in Figure 1. The participants came up with a variety of such arguments covering *Personal Care*, *Recreational*, *Group Needs* or *Emergency* requirements. We refer the readers to the dataset paper for more examples around these themes. The participants are then encouraged to leverage these arguments in their upcoming negotiations. As the authors state, this scenario mimics realistic negotiation settings where the participants engage in highly contextual conversations, based on their personal or domain-specific arguments to convince their partners.

This metadata can provide more direct feedback on which implicit preference statements can lead to a higher or a lower affinity towards a specific issue. To incorporate this, we create dummy dialogues using templates and add them to the training data for our opponent modeling task. Consider a set of arguments $A = (A_H, A_M, A_L)$, containing one argument for each priority. We extract two pairs[2] - (A_H, A_L) and (A_M, A_L). We construct the dummy dialogue as per Figure 1, ordering the arguments randomly to avoid any induced biases. To train the model, we partially mask the margin ranking loss from Equation 2 to only consider the loss from the pair for which the relation is known. Further, since a partial dialogue is not meaningful in this case, we only train the model with loss $\mathcal{L}_2$ using k=2. While we use metadata from CaSiNo in this case, such contextual data can be constructed based on domain-specific preferences and requirements for other realistic applications as well.

Capturing offers: We adapt DealOrNoDeal to better capture the offers. The dataset follows the same MIBT framework as CaSiNo, which enables our adaptation. Each dialogue in DealOrNoDeal concerns three *arbitrarily-defined* issues: *books*, *balls*, and *hats*. Due to the arbitrary nature of these issues, there is minimal context discussed in the dialogues, reducing it to essentially an exchange of offers from both sides (see example in Figure 1). Hence, we map these dialogues to our dataset by *randomly mapping the issues in this dataset to the issues in the target dataset*, in our case, CaSiNo. We modify the utterances by replacing all the occurrences of the issues with the corresponding issues in CaSiNo. For this purpose, we find that simple regular expressions prove to be effective (Appendix B.1). Once mapped, this adapted data is simply added to the training data for our opponent modeling task.

MIBT provides a generic framework for many useful negotiation tasks beyond DealOrNoDeal and CaSiNo such as salary negotiations or negotiations between art collectors distributing the items among each other. Hence, if the tasks follow the same MIBT structure, it is relatively straightforward to use such adaptations for other settings as well. This can be largely done with regular expressions but if not, this structural relatedness still paves the way for multi-task learning. We encourage researchers to explore this framework for future data collection procedures, especially with the current expensive data collection methodologies in this space.

[2]We skip the third pair due to an absence of a visible difference based on our qualitative analysis.

5 Experimental Design

We address the following questions. First, **how useful is the proposed methodology for opponent modeling?** We experiment with two pretrained language models and compare our ranker to standard baselines. To test the data augmentations, we analyze model ablations, including 0-shot and few-shot settings. We also observe if they lead to a better performance with a lower number of utterances. Second, **do preferences and offers contribute to the performance?** We look at average attention scores and analyze whether the performance varies by the integrative potential in the negotiation. We expected the performance to be higher in the cases with low integrative potential. In such cases, the negotiation is more competitive, which usually leads to a higher expression of preferences and offers. Third, **what is the scope for improvement?** We compare our model to a human expert and recognize some of the errors that the model makes, discussing potential directions for future work.

Datasets: Our primary focus is on the **C**aSiNo **D**ataset (**CD**). Each CaSiNo dialogue results in *two* dialogues for our analysis based on the two negotiator *perspectives* (Section 3). We report results on 5-fold cross validation for this dataset. We further leave out 100 dialogues from the training data for hyperparameter tuning, resulting in 1548 dialogues for training, 100 for tuning, and 412 for evaluation - for each cross fold. The arguments data is based on the metadata of CaSiNo. We extract the arguments from the training data of **CD**, leaving out 200 constructed dialogues for validation. This data is referred to as **CA**, for **C**aSiNo **A**rguments. The **DND** data is adapted from **D**ealOr**N**o**D**eal dataset where we only select the dialogues with at least 4 total utterances and unique priority values for meaningful training. We end up with 4074 dialogues for training and 444 for validation. All the models are primarily validated and tested on the corresponding subsets of **CD** (except for some additional analysis in the next Section).

Evaluation Metrics: Our metrics are inspired by the negotiation literature, along with related research in Dialog State Tracking (DST) and Learning-to-Rank(LTR) tasks in NLP. Our primary metric is Exact Match Accuracy (EMA): the percentage of cases where the predicted priority order is entirely correct. This is analogous to the popular Joint Goal Accuracy in DST which captures the cases where all the slots are correctly identified (Balaraman et al., 2021). For negotiation tasks, even knowing the topmost priority can be useful. Hence, we also report Top-1 Accuracy: the percentage of cases where the highest priority issue is correctly predicted. Finally, we report the Normalized Discounted Cumulative Gain (NDCG@3). NDCG has been widely used in LTR tasks with distinct relevance values (Yates et al., 2021), which is also true for the setting that we consider. In our case, we use the relevance values as 5, 4, and 3 for the most, second, and least ranked issue respectively, following the incentive design structure of CaSiNo. We compute these metrics for all k from 1 to 5, varying the number of opponent utterances seen by the model. We present the results at k=5 to analyze the performance after seeing almost all of the opponent utterances in CaSiNo. To capture the performance with partial dialogues, we report corresponding k-penalty versions that take a weighted average of the performance for different values of k, while giving a linearly higher weight to the performance at a lower k.

Methods: We call the complete model from Figure 1 as **CD + CA + DND** that combines all the three datasets for training. We compare it with its ablations, including 0-shot and few-shot scenarios. We further develop two standard baselines. The **Random** baseline chooses the final ranking at random, from all the possible orderings. **BoW-Ranker** is based on the Bag-of-Words paradigm. The input features are based on the normalized frequencies of the 500 most frequent words in the training dataset, except stopwords. Instead of contextualized hierarchical representations, this method directly uses a feed-forward network on the input BoW features to predict the ranking. The model is trained on partial dialogues using the same margin ranking loss.

Training Details: The embedding dimension throughout is 768 for transformer-based models. These models use base variant of either BERT (Devlin et al., 2019) or RoBERTa (Liu et al., 2019) for Level I encoder. The Level II encoder uses one transformer layer. The feed-forward network contains two fully connected layers with a final sigmoid activation. We train the model with Adam optimizer using a learning rate of $2e^{-5}$ for transformer-based methods and $2e^{-3}$ for **BoW-Ranker**. The margin c is kept as 0.3. We use a dropout of 0.1 to prevent overfitting. We further employ a loss-specific dropout of 0.15, in order to backpropagate the loss from fewer ks simultaneously. The models were trained for 20 epochs with a batch size of 25. We checkpoint after every epoch and the one with the highest EMA at k=5 on the held out **CD** dataset is chosen for evaluation. We provide the details on the computing infrastructure, hyper-parameter tuning, and validation performance in Appendix A. We will release our code on acceptance.

6 Results and Discussion

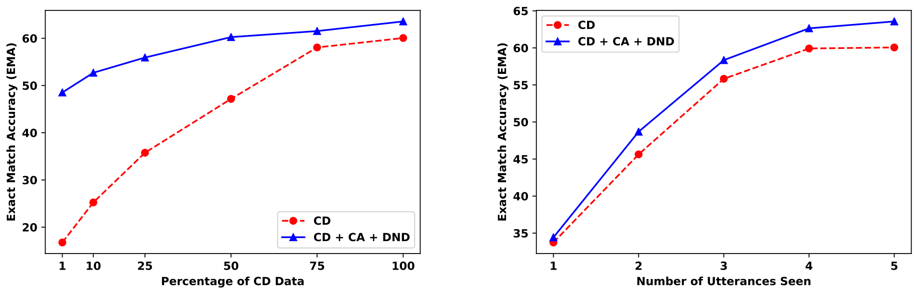
We summarize our results in Table 1. Our proposed ranking-based models beat the **Random** and **BoW-Ranker** baselines by a huge margin across all metrics. This is true even for 0-shot **DND** and for **CA + DND**, attesting the utility of the proposed ranking methodology and data adaptations[3]. We observe that RoBERTa-based models outperform BERT-based models on this task. The best performing configuration is the RoBERTa **CD + CA + DND** that combines all the three data sources. In Figure 2a, we plot the performance for different percentages of **CD** data. We only show RoBERTa-based models due to their superior performance. The plot highlights the advantage of adapting the related data sources, especially in few-shot settings, with **CD + CA + DND** at 50% matching the performance of **CD** at 100%. We also look at how the performance varies with the number of utterances seen in Figure 2b. We find that the performance gains are visible across all values of k. The data augmentations allow the model to perform better with a fewer number of observed utterances, making the model more useful in realistic scenarios.

Performance on the adapted datasets: We analyze if our joint learning also improves the performance on the validation sets of **CA** and **DND** datasets, showing advantages across multiple tasks. For **CA** dataset, we measure argument ranking accuracy: for a given input dialogue based on a pair of arguments, we consider a prediction as correct if the scores predicted by the model correctly rank the arguments. For **DND**, we analyze **EMA** at k=2 for opponent modeling, similar to our setup for CaSiNo. As evident from Tables 2a and 2b,

[3] Training with just the **CA** data only was not useful due to the lack of training with any partial dialogues.

Model	k=5			k-penalty		
	EMA	Top-1	NDCG@3	EMA	Top-1	NDCG@3
Random	16.46 (1.47)	32.49 (1.58)	48.49 (1.16)	16.59 (1.22)	33.99 (1.13)	49.76 (0.75)
BoW-Ranker	28.49 (1.3)	53.38 (2.21)	65.51 (0.62)	27.71 (1.24)	52.98 (1.97)	64.31 (1.67)
Bert-based						
DND	41.12 (3.06)	64.69 (2.94)	73.88 (1.57)	34.5 (1.12)	58.75 (1.35)	68.48 (0.77)
CA+DND	41.9 (2.93)	66.98 (3.17)	75.91 (2.28)	36.01 (1.25)	61.09 (1.9)	70.09 (1.49)
CD	53.97 (3.02)	77.7 (2.85)	83.75 (1.96)	42.3 (1.53)	66.8 (1.78)	74.39 (1.45)
CD+CA	57.24 (3.09)	79.74 (2.37)	84.99 (1.87)	44.39 (1.17)	67.88 (1.16)	75.31 (1.1)
CD+DND	56.12 (4.07)	79.16 (2.57)	84.66 (1.84)	43.79 (2.07)	68.18 (1.55)	75.38 (1.6)
CD+CA+DND	56.56 (2.07)	80.13 (1.07)	85.49 (1.09)	44.22 (1.82)	69.21 (2.05)	76.03 (1.6)
RoBerta-based						
DND	45.21 (3.07)	68.1 (2.8)	77.01 (1.76)	37.66 (1.41)	61.41 (2.3)	70.44 (1.5)
CA+DND	46.76 (1.89)	68.73 (1.22)	77.65 (0.9)	39.43 (1.67)	62.87 (2.5)	71.7 (1.83)
CD	60.06 (3.01)	81.98 (1.75)	86.54 (1.31)	46.57 (1.6)	69.26 (1.69)	76.17 (1.22)
CD+CA	60.01 (2.23)	80.23 (2.11)	85.85 (1.41)	46.96 (2.1)	68.59 (1.93)	76.05 (1.14)
CD+DND	62.54 (3.3)	82.56 (1.24)	**87.57** (1.18)	47.69 (2.52)	69.98 (1.96)	76.71 (1.55)
CD+CA+DND	**63.57** (3.44)	**82.76** (2.47)	87.55 (1.58)	**48.72** (2.03)	**70.03** (1.63)	**77.14** (1.38)

Table 1: Performance on the opponent modeling task, showing the utility of the proposed methods. EMA and Top-1 represent the accuracy in percentage. We also scaled NDCG@3 to 0-100. For all the metrics, higher is better. The numbers represent Mean (Std.) over 5-cross folds of the **CD** data.



(a) (b)

Figure 2: Mean performance for two RoBERTa-based models: (a) on different percentages of **CD** data. The Y-Axis represents EMA at k=5, (b) on different values of k.

Model	Accuracy
Random	52.4 (4.14)
AD	63.8 (9.33)
AD+DND	73.4 (6.19)
CD+AD	78.9 (1.39)
CD+AD+DND	76.7 (3.52)

(a)

Model	EMA
Random	16.04 (0.92)
DND	60.68 (2.05)
AD+DND	60.9 (1.87)
CD+DND	63.11 (1.77)
CD+AD+DND	63.56 (0.94)

(b)

Table 2: Performance for RoBERTa-based models: (a) argument classification accuracy on the validation set of **CA**, (b) EMA at k=2 for opponent modeling on the validation set of **DND**. The numbers represent Mean (Std.) over 5-cross folds.

we find support that joint learning improves the performance on **CA** and **DND** datasets as well.

Average attention: We recognize the utterances with preference statements by utilizing strategy annotations in CaSiNo (Chawla et al., 2021). We assume that an utterance contains a preference if it was annotated with at least one of **Self-Need**, **Other-Need**, or **No-Need** strategies. For identifying offers, we use regular expressions following prior work (He et al., 2018) (refer Appendix B.2). We consider any utterance that is not labeled with a preference or an offer as *Other*. Then, we observed the average attention put by the best-performing model on these categories in the Level II encoder. Preferences received an average of 0.3, offers received 0.27, and other utterances received 0.08 attention scores, without any explicit indication about these categories during model training. We consider this as preliminary evidence that the learning process matches our intuition, with preferences and offers contributing to the performance.

Performance across integrative potential: For more concrete evidence on the utility of preferences and offers, we look at how the performance varies between scenarios with low and high integrative potential. This basically captures how aligned the preferences of the two negotiators are in a negotiation. In a scenario with low integrative potential, the negotiations are more competitive, leading to a higher expression of preferences and offers and providing a better signal to our ranking models. For our best-performing model, we find EMA at k=5 to be 68.75 (4.58) for scenarios with low integrative potential against 60.31 (2.67) for those with high potential. This provides stronger evidence that the learning process sensibly takes into account the preference and offer statements in the data.

Scope for Improvement? Similar to the trained models, we asked a human expert (an author of this work) to guess the priority order of the opponent by accessing partial dialogues. The expert was allowed to make multiple guesses if she is unsure, in which case the final ranking was chosen randomly from all the guesses. We compare the expert to our best-performing model on 100 dialogues from the evaluation set. The expert achieved 75% mean EMA at k=5 against 66% for the model, while performing better on other metrics as well. We show the comparison by varying the parameter k in Appendix C. While the model performs reasonably, there is a scope for improvement. We performed a qualitative analysis of the errors made by the model and the expert. In many cases, it is infeasible to predict accurately, especially when negotiators engage in small talk early on - indicating a limited scope for improvement with fewer utterances. In some cases, there is more focus on the highest priority issue, giving less explicit signals of the entire ranking. This might work for some applications but in other cases, the agent design can be modified to discuss the complete ranking more explicitly. Integrating other datasets that follow the same MIBT structure (such as (DeVault et al., 2015)) via data adaptation or multi-task learning is another potential direction. We also observed errors in the cases that included longer contextually-dense utterances, where preferences are shared indirectly as a response to the partner, and when the negotiators give away their higher priority issues out of empathy towards their partner. These cases are easier for the expert but can be confusing to the model. Better modeling of the prior context and handling of longer utterances are also avenues for improvements in the future.

7 Related Work

Opponent modeling encompasses several tasks in negotiations such as priority estimation, predicting opponent limits like BATNA (Sebenius, 2017), and classifying them into various categories like personality (Albrecht and Stone, 2018; Baarslag et al., 2016). We focus only on inferring their priorities but in a more challenging domain involving chat-based interactions, instead of structured communication channels popular in prior work (Williams et al., 2012; Mell and Gratch, 2017; Johnson and Gratch, 2021). A realistic interface like natural language fundamentally alters the negotiation dynamics in terms of the exchange of information, and hence, requires a separate investigation.

For chat-based negotiations, Nazari et al. (2015) relied on heuristics and utterance-level annotations to infer the opponent priorities using frequency-based methods. Langlet and Clavel (2018) explored a symbolic rule-based system to parse the utterances collected from a multimodal interaction. Instead, our focus is on modeling the priorities directly from partial dialogues as input. Research in negotiation dialogue systems has mainly focused on end-to-end modeling of the agent, without any explicit opponent modeling (Lewis et al., 2017; He et al., 2018; Zhou et al., 2019; Cheng et al., 2019; Parvaneh et al., 2019). However, there is evidence that even end-to-end systems can benefit from being more opponent-aware, such as recent work that uses dialogue acts to estimate opponent's behavior (Zhang et al., 2020; Yang et al., 2021).

A number of related data augmentation strategies have been explored in Computer Vision and NLP (Shorten and Khoshgoftaar, 2019; Feng et al., 2021). Most methods use rules or models to transform the available data or create synthetic data to avoid overfitting while training. This especially helps in low-resource languages (Li et al., 2020) and few-shot scenarios (Kumar et al., 2019).

8 Conclusion

We presented and evaluated an approach for the task of opponent modeling in negotiation dialogues. Our comparison to baselines and ablations attest to the utility of our method. We found that the proposed data adaptations can be especially beneficial in 0-shot and few-shot scenarios. In the future, we will explore two primary directions: first, improving the model performance on opponent modeling by leveraging other related available datasets and by better incorporating the negotiation dialogue context, and secondly, using effective opponent modeling techniques towards the design of automated negotiation systems for applications in pedagogy and conversational AI.

9 Broader Impact and Ethical Considerations

Datasets Used: Both the datasets used in this work had been completely anonymized before their release by the respective authors. Moreover, we carefully verified the licensing details and ensured that the datasets were only used within the scope of their intended usage. Finally, we note that both the datasets are in English. Although this means that our experiments were limited to one language, our approach makes no such assumptions and should be broadly applicable to other settings as well. We encourage researchers to extend this work and study human-machine negotiations for other languages as well, provided suitable datasets are available. This would open up exciting avenues for future research, given the well-documented differences in how humans negotiate across cultures (Andersen et al., 2018; Luo, 2008).

Human Annotations: Human annotations were used to estimate the expert performance on this task. This did not involve any additional crowdsourcing effort. Instead, the dialogues were annotated by an author of this work.

Opponent Modeling For Negotiation Dialogues: Negotiations are typically non-collaborative in nature, where the goals of the negotiating parties may not align with each other. Hence, the negotiators may not always feel comfortable in revealing their preferences for fear of being exploited. Even if they do, inferring them from natural language is challenging as preferences might be implied, and resolving these implications involves domain-specific knowledge and prior dialogue context. Regardless, incorporating such realistic communication channels is critical for designing practical and robust AI systems for downstream applications. However, most of the prior efforts in negotiations use restrictive menu-driven systems based on button clicks. Our work is a step towards bridging this gap.

Our efforts are part of our broader objectives towards building automated negotiation systems, that are trained either in an end-to-end manner or based on a modular design. For conversational AI applications, opponent modeling systems that can predict the priorities of the opponent reliably based on a partial dialogue can inform the strategy of the agent in the latter parts of the conversation. From the perspective of pedagogical applications, even the systems that can predict the priorities of a negotiator at the end of the negotiation can be helpful. For instance, consider a negotiation between two students, A and B who are asked to guess the opponent's priorities at the end of their negotiation. If the pedagogical agent is able to accurately guess the priorities of student B, while student A fails to guess correctly, this can be used to give concrete feedback to students who fail to recognize these strategies even if the information in the conversation was enough for the model to make these predictions accurately.

Ethical Recommendations: Finally, we briefly discuss the ethical considerations around the design of automated negotiation systems. A considerable amount of research in negotiations has focused on ethics. Primary concerns revolve around the acts of emotion manipulation, bias, deception, and misinterpretation (Lewicki et al., 2016). Consequently, these issues can also emerge in the systems that are developed on human-human negotiation dialogue datasets. Our central recommendation in mitigating the impact of these issues for negotiation dialogue systems or other conversational AI assistants is transparency - around the identity, capabilities, and any known undesirable behaviors of the system. Further, any data collected during the deployment phase should be properly anonymized and the users of the system should be well-informed. In particular, we recommend extra precautions for systems that are adaptive towards their opponents or users such as having regular monitoring for any unexpected behaviors, to ensure that the systems are not offensive or discriminatory.

References

Stefano V Albrecht and Peter Stone. 2018. Autonomous agents modelling other agents: A comprehensive survey and open problems. *Artificial Intelligence*, 258:66–95.

Steffen Andersen, Seda Ertac, Uri Gneezy, John A List, and Sandra Maximiano. 2018. On the cultural basis of gender differences in negotiation. *Experimental Economics*, 21(4):757–778.

Tim Baarslag, Mark Hendrikx, Koen Hindriks, and Catholijn Jonker. 2013. Predicting the performance of opponent models in automated negotiation. In *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, volume 2, pages 59–66. IEEE.

Tim Baarslag, Mark JC Hendrikx, Koen V Hindriks, and Catholijn M Jonker. 2016. Learning about the opponent in automated bilateral negotiation: a comprehensive survey of opponent modeling techniques. *Autonomous Agents and Multi-Agent Systems*, 30(5):849–898.

Vevake Balaraman, Seyedmostafa Sheikhalishahi, and Bernardo Magnini. 2021. Recent neural methods on dialogue state tracking for task-oriented dialogue systems: A survey. In *Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 239–251.

Kushal Chawla, Jaysa Ramirez, Rene Clever, Gale Lucas, Jonathan May, and Jonathan Gratch. 2021. Casino: A corpus of campsite negotiation dialogues for automatic negotiation systems. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 3167–3185.

Minhao Cheng, Wei Wei, and Cho-Jui Hsieh. 2019. Evaluating and enhancing the robustness of dialogue systems: A case study on a negotiation agent. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 3325–3335.

David DeVault, Johnathan Mell, and Jonathan Gratch. 2015. Toward natural turn-taking in a virtual human negotiation agent. In *AAAI Spring Symposia*. Citeseer.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186.

Steven Y Feng, Varun Gangal, Jason Wei, Sarath Chandar, Soroush Vosoughi, Teruko Mitamura, and Eduard Hovy. 2021. A survey of data augmentation approaches for nlp. *arXiv preprint arXiv:2105.03075*.

Chaim Fershtman. 1990. The importance of the agenda in bargaining. *Games and Economic Behavior*, 2(3):224–238.

Jonathan Gratch, David DeVault, Gale M Lucas, and Stacy Marsella. 2015. Negotiation as a challenge problem for virtual humans. In *International Conference on Intelligent Virtual Agents*, pages 201–215. Springer.

He He, Derek Chen, Anusha Balakrishnan, and Percy Liang. 2018. Decoupling strategy and generation in negotiation dialogues. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2333–2343.

Emmanuel Johnson and Jonathan Gratch. 2021. Comparing the accuracy of frequentist and bayesian models in human-agent negotiation. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*, pages 139–144.

Emmanuel Johnson, Gale Lucas, Peter Kim, and Jonathan Gratch. 2019a. Intelligent tutoring system for negotiation skills training. In *International Conference on Artificial Intelligence in Education*, pages 122–127. Springer.

Emmanuel Johnson, Sarah Roediger, Gale Lucas, and Jonathan Gratch. 2019b. Assessing common errors students make when negotiating. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, pages 30–37.

Varun Kumar, Hadrien Glaude, Cyprien de Lichy, and Wlliam Campbell. 2019. A closer look at feature space data augmentation for few-shot intent classification. In *Proceedings of the 2nd Workshop on Deep Learning Approaches for Low-Resource NLP (DeepLo 2019)*, pages 1–10.

Caroline Langlet and Chloé Clavel. 2018. Detecting user's likes and dislikes for a virtual negotiating agent. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction*, pages 103–110.

Yaniv Leviathan and Yossi Matias. 2018. Google duplex: An ai system for accomplishing real-world tasks over the phone. *URL https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html*, 3.

Roy J Lewicki, Bruce Barry, and David M Saunders. 2016. *Essentials of negotiation*. McGraw-Hill.

Mike Lewis, Denis Yarats, Yann Dauphin, Devi Parikh, and Dhruv Batra. 2017. Deal or no deal? end-to-end learning of negotiation dialogues. In *EMNLP*.

Yu Li, Xiao Li, Yating Yang, and Rui Dong. 2020. A diverse data augmentation strategy for low-resource neural machine translation. *Information*, 11(5):255.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Peng Luo. 2008. Analysis of cultural differences between west and east in international business negotiation. *International Journal of Business and Management*, 3(11):103–106.

Johnathan Mell and Jonathan Gratch. 2017. Grumpy & pinocchio: answering human-agent negotiation questions through realistic agent design. In *Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems*, pages 401–409. International Foundation for Autonomous Agents and Multiagent Systems.

Zahra Nazari, Gale M Lucas, and Jonathan Gratch. 2015. Opponent modeling for virtual human negotiators. In *International Conference on Intelligent Virtual Agents*, pages 39–49. Springer.

Amin Parvaneh, Ehsan Abbasnejad, Qi Wu, and Javen Shi. 2019. Show, price and negotiate: A hierarchical attention recurrent visual negotiator. *arXiv preprint arXiv:1905.03721*.

James K Sebenius. 2017. Batna s in negotiation: Common errors and three kinds of “no”. *Negotiation Journal*, 33(2):89–99.

Connor Shorten and Taghi M Khoshgoftaar. 2019. A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1):1–48.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.

Colin R Williams, Valentin Robu, Enrico H Gerding, and Nicholas R Jennings. 2012. Iamhaggler: A negotiation agent for complex environments. In *New Trends in Agent-based Complex Automated Negotiations*, pages 151–158. Springer.

Runzhe Yang, Jingxiao Chen, and Karthik Narasimhan. 2021. Improving dialog systems for negotiation with personality modeling. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 681–693.

Andrew Yates, Rodrigo Nogueira, and Jimmy Lin. 2021. Pretrained transformers for text ranking: Bert and beyond. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, pages 1154–1156.

Zheng Zhang, Lizi Liao, Xiaoyan Zhu, Tat-Seng Chua, Zitao Liu, Yan Huang, and Minlie Huang. 2020. Learning goal-oriented dialogue policy with opposite agent awareness. *arXiv preprint arXiv:2004.09731*.

Yiheng Zhou, Yulia Tsvetkov, Alan W Black, and Zhou Yu. 2019. Augmenting non-collaborative dialog systems with explicit semantic and strategic dialog history. In *International Conference on Learning Representations*.

A Experiments

A.1 Computing Infrastructure

All experiments were performed on a single Tesla V100 GPU. The complete model (**CD + CA + DND**) takes around 10 hours for training with 32-bit precision on a single cross-validation fold with a batch size of 25.

A.2 Training Details

We used a combination of randomized and manual search to tune the hyperparameters. For each cross fold, we kept 50 dialogues from the **CD** training data for parameter tuning. This amounts to 100 data points, considering the two perspectives extracted from each dialogue. The metric for choosing the best hyperparameters is EMA at k=5, averaged over the 5 cross-validation folds. We tuned the parameters on the performance of the BERT-based model with **CD + CA + DND** configuration.

We vary the learning rate in $\{1e^{-5}, 2e^{-5}, 3e^{-5}\}$, dropout in $\{0.0, 0.1, 0.2\}$, and loss-specific dropout in $\{0.0, 0.15, 0.25\}$. We also varied the number of transformer layers in Level II encoder from Figure 1 in the set $\{1, 2, 3\}$. For **DND**, we also varied the number of instances that were chosen for adaptation but found that using all the instances that passed our filtering gave the best performance. We further varied the margin for ranking loss in $\{0.0, 0.3, 0.5\}$. Finally, for the models trained on combined datasets, we tried with a higher weightage (2x) for the loss contribution of CA-adapted instances due to their lower total count but found no visible improvements in the performance. The rest of the hyper-parameters were fixed based on the available computational and space resources. We report the best performing hyper-parameters in the main paper.

The models used in the paper have nearly 171 million trainable parameters. We report the mean performance on the validation set in Table 3.

A.3 External Packages and Frameworks

The models were developed in PyTorch Lightning[4] and relied on the HuggingFace Transformers library[5] for using the pretrained models and their corresponding tokenizers. We used a number of external packages such as Python Scikit Learn[6] library for implementing the evaluation metrics, and NLTK[7] for tokenization for the Bag-of-Words model.

[4]https://www.pytorchlightning.ai/

[5]https://github.com/huggingface/transformers

Model	EMA
Random	17.8 (4.87)
BoW-Ranker	35 (3.35)
Bert-based	
DND	51 (1.67)
CA + DND	51.2 (3.12)
CD	63.6 (4.84)
CD + CA	65.8 (1.94)
CD + DND	69 (2.28)
CD + CA + DND	70 (2.61)
RoBerta-based	
DND	54.6 (5.43)
CA + DND	55 (5.55)
CD	70.2 (3.19)
CD + CA	70 (3.95)
CD + DND	75.6 (2.15)
CD + CA + DND	**77.8** (2.32)

Table 3: Validation performance for opponent modeling on **CD** dataset. The reported EMA is at k=5. The numbers represent Mean (Std.) over 5-cross folds of the **CD** data.

B Regular Expression Usage

B.1 Adapting DealOrNoDeal data

We randomly mapped *book* from DealOrNoDeal to *food*, replacing all occurrences of ‘book’ and ‘books’ with ‘food’ in the utterances. Similarly, *hat* was mapped to *water*, and *ball* was mapped to *firewood*. Since the dialogues only involve minimal context about the issues, we found these replacements to be sufficient.

B.2 Identifying Offer statements

The offer statements were also recognized by regular expressions for the purpose of computing average attention scores. Specifically, an utterance is classified as having an offer, if it contains 3 or more of the following phrases - {’0’, ’1’, ’2’, ’3’, ’one’, ’two’, ’three’, ’all the’, ’food’, ’water’, ’firewood’, ’i get’, ’you get’, ’what if’, ’i take’, ’you can take’, ’can do’}. The threshold 3 and these phrases were chosen heuristically via qualitative analysis.

[6]https://scikit-learn.org/stable/modules/model_evaluation.html

[7]https://www.nltk.org/api/nltk.tokenize.html

C Comparison with Human Performance

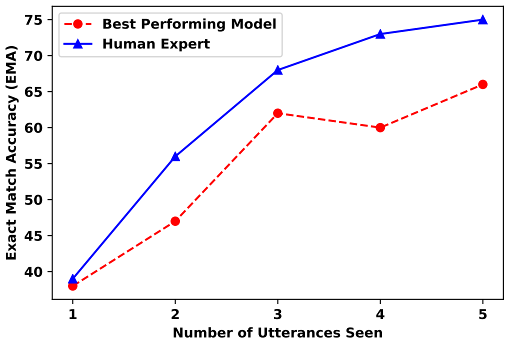


Figure 3: Mean performance comparison for the best performing model with the human expert for different values of k.

We present the performance for our best performing model with the human expert across different values of k in Figure 3.