Let's Negotiate! A Survey of Negotiation Dialogue Systems

Anonymous EACL submission

Abstract

Negotiation is a crucial ability in human communication. Recently, there has been a resurgent research interest in negotiation dialogue systems, whose goal is to create intelligent agents that can assist people in resolving conflicts or reaching agreements. Although there have been many explorations into negotiation dialogue systems, a systematic review of this task has not been performed to date. We aim to fill this gap by investigating recent studies in the field of negotiation dialogue systems, and covering benchmarks, evaluations and methodologies within the literature. We also discuss potential future directions, including multi-modal, multi-party and cross-cultural negotiation scenarios. Our goal is to provide the community with a systematic overview of negotiation dialogue systems and to inspire future research.

1 Introduction

Negotiation involves two or more individuals discussing goals and tactics to resolve conflicts, achieve mutual benefit, or find mutually acceptable solutions (Fershtman, 1990; Bazerman and Neale, 1993; Lewicki et al., 2011). It is commonly used to manage conflict and is the primary give-and-take process by which people try to reach an agreement (Fisher et al., 2011; Lewicki et al., 2011). Negotiations can be cooperative or competitive and are used in various social settings such as informal, peer to peer, organizational, and diplomatic country to country settings (Cano-Basave and He, 2016) and thus the implications for enhancing outcomes are vast. However, humans are naturally subject to various biases and can be swayed by emotion during negotiations, making them inclined to overlook useful implicit information from other participants in the negotiation process and hindering optimal outcomes. Negotiators also often lack the necessary skills, training and knowledge to achieve their desired goals (Walton and McKersie, 1991).

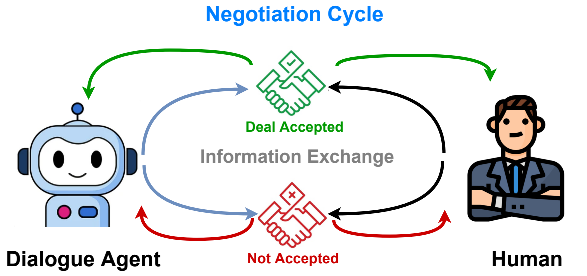

Figure 1: A typical negotiation dialogue involves a multi-turn interaction between agent and human. They exchange information about their deals and end up with accepting or declining deals.

To facilitate human negotiation processes, previous researchers (Lewandowska, 1982; Lambert and Carberry, 1992; Chawla et al., 2021b) have aimed to build intelligent negotiation agents that can aid humans or even directly negotiate with humans in multi-turn interactions (Figure 1). Effective agents could yield significant benefits in many real-world scenarios, ranging from bargaining prices in everyday life (He et al., 2018) to higher-stakes political or legal situations (Cano-Basave and He, 2016).

Research on negotiation has been conducted for almost 60 years in the field of psychology, political science, and communication. It has evolved over the past decades from exploring game theory (Walton and McKersie, 1991), behavior decisions driven by the cognitive revolution in psychology (Bazerman and Neale, 1993), to cultural differences in the 2000s (Bazerman et al., 2000). Negotiation research, however, is now forced to confront the implications of human/AI collaborations given recent advancements in machine learning (Bazerman et al., 2000; Ouali et al., 2017). Research has focused on establishing new benchmarks and testing environments for various negotiation dialogue scenarios, including product price bargaining (Lewis et al., 2017; Heddaya et al., 2023), multiple player strategic games (Asher et al., 2016) and job interviews (Zhou et al., 2019). Other research has attempted to propose new methodologies and frameworks to model the negotiation process, including various negotiation policy learning, negotiator mental status modeling and negotiation decision making. Converging efforts from social scientists and data scientists which incorporate insights from both fields will thus be fruitful in maximizing processes and outcomes in negotiations.

Despite the significant amount of research that has been conducted, we are not aware of a systematic review on the topic. In this work, we aim to fill this gap by reviewing contemporary research efforts in the field of negotiation dialogue systems from the dimensions of datasets, evaluation metrics and modeling approaches. We first briefly explore human negotiations and corresponding limitations, and propose how dialogue agents may supplement human negotiation processes. We then discuss the popular negotiation dialogue modeling methods, including *Strategy modeling*, *Negotiator modeling* and *Action modeling*. We further introduce existing datasets according to their negotiation scenarios. Finally, we give an overview for three major types of evaluation metrics, i.e., *goal-based metrics*, *game-based metrics* and *human evaluation*, used in negotiation dialogue systems.

In summary, our contributions are three-fold: (i) we point out human limitations in negotiation and systematically summarize the existing AI solutions aiming to address those limitations; (ii) we systematically categorize current negotiation dialogue benchmarks from a distributive and integrative perspective, and provide an overview of evaluation methods; (iii) we point out current limitations and promising future research directions.

2 Negotiations from a Social Science Perspective

In this section, we will first introduce a framework for human negotiation from social sciences, then discuss human limitations in negotiation, which motivates NLP researchers/practitioners to develop strong negotiation dialogue systems.

2.1 Understanding of Human Negotiations

Brett and Thompson (2016) propose a comprehensive framework for a two-party negotiation process, as shown in Figure 2. Preferences and strategies of the negotiators determine the potential outcomes and the interaction of the negotiation process. The preferences of both negotiators create the potential outcome that may be reached by them. The negotiators' strategies, defined as the goal-directed behaviors that are used in order to reach an agreement (Weingart et al., 1990), affect the interaction, ultimately determining how much of that potential outcome created by the negotiators' preferences is obtained.

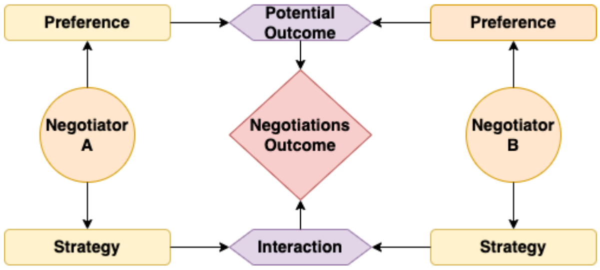

Figure 2: Negotiation Framework for two negotiator scenario from Brett and Thompson (2016).

2.2 Human limitations in Negotiation

Although negotiations are commonly found in daily life (e.g., price bargaining), it is still a challenging task. Without professional training, people often lack the negotiation skills to achieve their desirable goals. They may not know what *strategies* to be used and how to implement these *strategies*. It is also challenging to identify and process implicit information about other negotiators' interests and preferences in the negotiation. Often times, people view negotiation as a competition and may not even be motivated to seek or express this information (Brett and Thompson, 2016). Finally, human cognitive heuristics, biases and emotionality may prove a hindrance in negotiation scenarios. For example, people view themselves, the world and the future as being more positive than in reality (Taylor, 1989), which may lead to overestimation and optimism in negotiations (Crocker, 1982). The negotiation could also lead participants to be emotionally engaged and make it more difficult to process information rationally (Pinkley and Northcraft, 1994). Thus, developing effective negotiation conversational dialogue agents can be beneficial for understanding and controlling for these various factors, and optimizing the negotiation.

3 Methodology Overviews

In negotiation dialogues, negotiators interact with each other in a strategic discussion to reach a final goal. As discussed above, *strategies* and *preferences* significantly affect the negotiation outcomes.

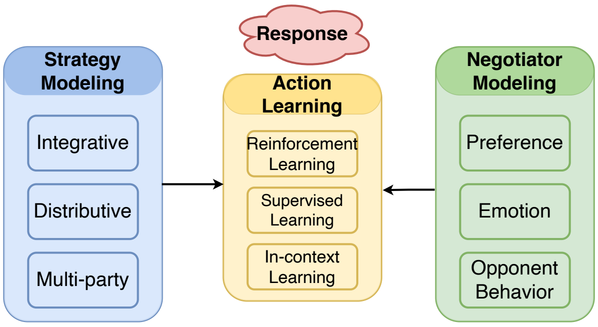

Figure 3: An overview architecture of method section. The *strategy* and *negotiator* modules collect information from the negotiation dialogue, and the *action learning* module conditions on the information and produce responses to push the negotiation forward.

To effectively assist people in this process, as shown in Figure 3, existing research on negotiation dialogues can be categorized into *a) Negotiator Modeling*; *b) Strategy Modeling*; *c) Action Learning*. Herein, *Negotiator Modeling* aims to infer the *explicit information* from other negotiators based on a dialogue context. *Strategy Modeling* learns to select strategies to use given the current dialogue context. Finally, the *Action Learning* incorporates the above negotiation information to map strategies into observable actions or responses, e.g. utterances, by developing dialogue models within the existing machine learning frameworks.

3.1 Problem Formulation

Formally, a negotiation dialogue process can be formally characterized as a tuple $(n, \mathcal{K}, \mathcal{S}, \mathcal{U}, \pi, g)$. Herein, n refers to the number of negotiation party ($n \geq 2$), $\mathcal{K}$ refers to the background information for a negotiation dialogue, such as negotiator's preferences and demands towards items. This information may not be transparent to others in a dialogue. $\mathcal{S}$ denotes a strategy trajectory $\{s_1, s_2, ...\}$ used during the negotiation process. $\mathcal{U} = \{u_1, u_2, ...\}$ is a sequence of dialogue utterances or actions in a negotiation process. A policy $\pi_\theta(\mathcal{K}, \mathcal{S}, \mathcal{U})$ is a distribution of actions or a mapping to determine which actions or utterances to produce in order to reach the final negotiation goal g.

3.2 Strategy Modeling

In negotiations, people use a wide range of tactics and approaches to achieve their goals. Many previous research efforts have focused on modeling these strategies. They can be categorized into three aspects: *integrative* (win-win), such as maximizing unilateral interests (Bazerman and Neale, 1993), and *distributive* (win-lost), such as bargaining (Fershtman, 1990), and *multi-party* (Li et al., 2021).

3.2.1 Integrative Strategy

Integrative strategy (known as *win-win*) modeling aims to achieve mutual gains among participants. For instance, Zhao et al. (2019) propose to model the discourse-level strategy using a latent action reinforcement learning (LaRL) framework. LaRL can model strategy transition within a latent space. However, due to the lack of explicit strategy labels, LaRL can only analyze strategies in implicit space. To resolve this problem, Chawla et al. (2021b) define a series of explicit strategies such as *Elicit-Preference*, *Coordination* and *Empathy*. While *Elicit-Preference* is a strategy attempting to discover the preference of an opponent, *Coordination* promotes mutual benefits through an explicit offer or implicit suggestion. In order to capture user's preference, Chawla et al. (2022) utilize those strategies using a hierarchical neural model. Yamaguchi et al. (2021) also present another collaborative strategy set to negotiate workload and salaries during the interview, whose goal is to reach an agreement between an employer and employee, recommending, for example, to communicate politely, address concerns, and provide side offers.

3.2.2 Distributive Strategy

Distributive strategy (known as *win-loss*) modeling focuses on achieving one's own goals and maximizing unilateral interests over mutual benefits. Distributive strategy is used when one insists on their own position or resists the opponent's deal (Zhou et al., 2019). For example, to persuade others to donate to a charity, Wang et al. (2019) propose a set of persuasion strategies containing 10 different strategies, including logical appeal, emotional appeal, source-related inquiry and others. Further exploration on the role of structure (e.g., facing act, emotion) (Li et al., 2020a; Dutt et al., 2020) helps utilize strategy modeling between asymmetrical roles. Another line of research focuses on the adversarial attack strategy. Dutt et al. (2021a) investigate four resisting categories, namely contesting, empowerment, biased processing, and avoidance (Fransen et al., 2015). Each individual category contains fine-grained strategic behaviors. For example, contesting refers to attacking the message source, and empowerment implies reinforcing personal preference to contradict a claim (*Attitude*

Bolstering) or attempting to arouse guilt in the opponent (*Self Pity*).

3.2.3 Multi-party Strategy

While the previously mentioned work on integrative and distributive strategy modeling mainly relates to two-party negotiations, multi-party strategy modeling is slightly different. In multi-party situations, strategy modeling needs to consider different attitudes and complex relationships among individual participants, whole groups, and subgroups (Traum et al., 2008). Georgila et al. (2014) attempt to model multi-party negotiation using a multi-agent RL framework. Furthermore, Shi and Huang (2019) propose to construct a discourse dependency tree to predict relation dependency among multi-parties. Li et al. (2021) disclose relations between multi-parties using a graph neural network. However, research in multi-party strategies is currently hindered by limited relevant datasets and benchmarks.

3.3 Negotiator Modeling

Negotiation dialogues are affected by various features of negotiators. There is psychological evidence showing that, for example, a negotiation process is affected by personality (Sharma et al., 2013), relationships (Olekalns and Smith, 2003), social status (Blader and Chen, 2012) and cultural background (Leung and Cohen, 2011). We thus summarize the existing works on modeling negotiators from following three perspectives: *Preference*, *Emotion*, and *Opponent Behavior*.

3.3.1 Preference Modeling

Preference estimation helps an agent infer the intention of their opponents and guess how their own utterances would affect the opponents' preference. Nazari et al. (2015) propose a simple heuristic frequency-based method to estimate the negotiator's preference. However, a critical challenge for preference modeling in negotiation is that it usually requires complete dialogues, so it is difficult to predict those preferences precisely from a partial dialogue. Therefore, Langlet and Clavel (2018) consider a rule-based system to carefully analyze linguistic features from partial dialogue to identify user's preference. In further, to enhance preference modeling in those partial dialogues, which widely exist in real-world applications, Chawla et al. (2022) formulate preference estimation as a ranking task and propose a transformer-based model that can be trained directly on partial dialogues.

3.3.2 Emotion Modeling

Emotion modeling refers to recognizing emotions or emotional changes of negotiators. Explicit modeling of emotions throughout a conversation is crucial to capture and estimate reactions from opponents. To study emotional feelings and expressions in negotiation dialogues, Chawla et al. (2021a) explore the prediction of two important subjective goals, including outcome satisfaction and partner perception. Liu et al. (2021) provide explicit modeling on emotion transition engaged using pre-trained language models (e.g., DialoGPT), to support patients. Further, Dutt et al. (2020) propose a novel set of dialogue acts modeling *face*, which refers to the public self-image of an individual, in persuasive discussion scenarios. Mishra et al. (2022) utilize a reinforcement learning framework to elicit emotions in persuasive messages.

3.3.3 Opponent Behavior Modeling

Opponent behavior modeling refers to detecting and predicting opponents' behaviors during a negotiation process. For example, fine-grained dialogue act labels are provided in the Craigslist dataset (He et al., 2018), to help track the behaviors of buyers and sellers. Based on this information, Zhang et al. (2020) propose an opposite behavior modeling framework to estimate opposite action using DQN-based policy learning. Tran et al. (2022) leverage dialogue acts to identify optimal strategies for persuading people to donate. He et al. (2018) firstly propose a framework to decouple the opponent behavior modeling with utterance generation, which allows negotiation systems to manage opponent modeling in a precise manner. Yang et al. (2021) further improve the negotiation system with a first-order model based on the theory of Mind (Frith and Frith, 2005), which allows agents to compute an expected value for each mental state. They provided two variants of ToM-based dialogue agents: explicit and implicit, which can fit both pipeline and end-to-end systems.

3.4 Action Learning

Action learning empowers negotiation dialogue systems to properly incorporate previous strategies and other negotiator information to generate high-quality responses. Existing research on policy learning can be broadly categorized into *reinforce-*

ment learning, *supervised learning* and *in-context learning*.

3.4.1 Reinforcement Learning

English and Heeman (2005) pioneer applying reinforcement learning (RL) techniques to negotiation dialogue systems. They propose a single-agent RL framework that learns the policy of two participants individually. However, the single-agent framework is not feasible for situations where two agents interact frequently in a continuously changing environment. Georgila et al. (2014) further propose to use multi-agent RL techniques and provide a way to deal with multi-issue negotiation scenarios. Furthermore, Keizer et al. (2017) propose to learn about the actions of negotiators with a Q-learning reward function. They use a Random Forest model trained on a large human negotiation corpus from (Afantenos et al., 2012).

Most recent works have tried to build negotiation dialogue models using RL techniques with deep learning. Zhang et al. (2020) propose OPPA, which utilizes the system actions to estimate how a target agent behaves. The system actions are predicted based on the target agent's actions. The reward of the executed actions is obtained by predicting a structured output given a whole dialogue. Additionally, Shi et al. (2021) use a modular framework containing a language model to generate responses. A response detector would automatically annotate the response with a negotiation strategy and an RL-based reward function to assign a score to the strategy. However, this modular framework separates policy learning from response generation. Gao et al. (2021) propose an integrated framework with deep Q-learning, which includes multiple channel negotiation skills. It allows agents to leverage parameterized DQN to learn a comprehensive negotiation strategy that integrates linguistic communication skills and bidding strategies.

3.4.2 Supervised Learning

Supervised learning (SL) is another popular paradigm for policy learning. Lewis et al. (2017) adopt a Seq2Seq model to learn what action should be taken by maximizing the likelihood of the training data. However, supervised learning only aims to mimic the average human behavior, so He et al. (2018) propose to apply a supervised model to directly optimize a particular dialogue reward function, which is characterized by i) the utility function of the final price for the buyer and seller ii) the differences between two agents' utilities iii) the number of utterances in the dialogue. Zhou et al. (2020) first train a strategy predictor to predict whether a certain negotiation strategy occurred in the next utterance using supervised training. Then, the response generation conditions on the predicted negotiation strategy, as well as user utterance and dialogue context. In addition, Joshi et al. (2021) incorporate a pragmatic strategies graph network with the seq2seq model to create an interpretable policy learning paradigm. Recently, Dutt et al. (2021b) propose a generalized framework for identifying resisting strategies in persuasive negotiations using a pre-trained BERT model (Devlin et al., 2019). In addition, there are also research attempts to jointly train several sub-tasks simultaneously. Li et al. (2020b) propose an end-to-end framework that integrates several sub-tasks, including intent and semantic slot classification, response generation and filtering tasks in a Transformer-based pre-trained model. Zhou et al. (2020) propose jointly modelling semantic and strategy history using finite state transducers (FSTs) with hierarchical neural models. Chawla et al. (2022) integrate a preference-guided response generation model with a ranking module to identify opponents' priority.

3.4.3 In-context Learning

With the recent emergence of large language models (LLMs) such as GPT-3.5 and GPT-4, there have been a few studies that apply zero-shot and few-shot in-context learning. These techniques leverage the inherent knowledge of LLMs to predict agent behaviors more accurately. Fu et al. (2023) utilize LLMs in the context of bargaining, while Xu et al. (2023) employ them for the popular game "Werewolf." In both tasks, the LLMs were tasked with negotiating with other LLMs under specific scenarios.

4 Negotiation Datasets

In this section, we summarize the existing negotiation datasets and resources. Table 1 shows all of the 14 collected benchmarks, along with their negotiation types, scenarios, data scale and modality. We categorize these benchmarks based on their negotiation types, namely, *integrative* negotiation and *distributive* negotiation.

4.1 Integrative Negotiation Datasets

In integrative negotiations, there is normally more than one issue being negotiated. To achieve optimal

DataSet	Negotiation Type	Scenario	# Dialogue	# Avg. Turns	# Party	# Modality
InitiativeTaking (Nouri and Traum (2014))	Integrative	Fruit Assignment	41	-	Multi	-
STAC (Asher et al. (2016))	Integrative	Strategy Games	1081	8.5	Two	-
DealorNoDeal (Lewis et al. (2017))	Integrative	Item Assignment	5808	6.6	Two	-
Craigslist (He et al. (2018))	Distributive	Price Bargain	6682	9.2	Two	-
M3 (Kontogiorgos et al. (2018))	Integrative	Object Moving	15	-	Multi	MultiModal
Niki & Julie (Artstein et al. (2018))	Integrative	Item Ranking	600	-	Two	MultiModal
NegoCoach (Zhou et al. (2019))	Distributive	Price Bargain	300	-	Two	-
PersuasionforGood (Wang et al. (2019))	Distributive	Donation	1017	10.43	Two	-
FaceAct (Dutt et al. (2020))	Distributive	Donation	299	35.8	Two	-
AntiScam (Li et al. (2020b))	Distributive	Privacy Protection	220	12.45	Two	-
CaSiNo (Chawla et al. (2021b))	Integrative	Item Assignment	1030	11.6	Two	-
JobInterview (Yamaguchi et al. (2021))	Integrative	Job Interview	2639	12.7	Two	-
DeliData (Karadzhov et al. (2021))	Integrative	Puzzle Game	500	28	Multi	-
DinG (Boritchev and Amblard (2022))	Integrative	Strategy Game	10	2357.5	Multi	-
NegoBar (Heddaya et al. (2023))	Distributive	Price Bargain	408	35.85	Two	-

Table 1: Negotiation dialogues benchmarks are sorted by their publication time. For each dataset, we present the negotiation type, scenario, the number of dialogues and corresponding average turns, and party attributes.

negotiation goals, the involved players should make trade-offs for these multiple issues.

Multi-player Strategy Games Strategy video games provide ideal platforms for people to verbally communicate with other players to accomplish their missions and goals. Asher et al. (2016) propose the STAC benchmark, which is based on the game of Catan. In this game, players need to gather resources, including wood, wheat, sheep, and more, with each other to purchase settlements, roads and cities. As each player only has access to their own resources, they have to communicate with each other. To investigate the linguistic strategies used in this situation, STAC also includes an SDRT-styled discourse structure. Boritchev and Amblard (2022) also collect a *DinG* dataset from French-speaking players in this game. The participants are instructed to focus on the game, rather than talk about themselves. As a result, the collected dialogues can better reflect the negotiation strategy used in the game process.

Negotiation for Item Assignment Item assignment scenarios involve a fixed set of items as well as a predefined priority for each player in the dialogue. As the players only have access to their own priority, they need to negotiate with each other to exchange the items they prefer. Nouri and Traum (2014) propose *InitiativeTalking*, occurring between the owners of two restaurants. They discuss how to distribute the fruits (i.e., apples, bananas, and strawberries) and try to reach an agreement. Lewis et al. (2017) propose *DealorNoDeal*, a similar two-party negotiation dialogue benchmark where both participants are only shown their own sets of items with a value for each and both of them are asked to maximize their total score after negotiation. Chawla et al. (2021b) propose *CaSiNo*, a dataset on campsite scenarios involving campsite neighbors negotiating for additional food, water, and firewood packages. Both parties have different priorities over different items.

Negotiation for Job Interview Another commonly encountered negotiation scenario is job offer negotiation with recruiters. Yamaguchi et al. (2021) fill this gap and propose the *JobInterview* dataset. *JobInterview* includes recruiter-applicant interactions over salary, day off, position, and workplace. Participants are informed with opposite's preferences and the corresponding issues. Feedback from the opposites will be forwarded to participants during the negotiation process.

4.2 Distributive Negotiation Datasets

Distributive negotiation is a discussion over a fixed amount of value (i.e., slicing up the pie). In such negotiation, the involved people normally talk about a single issue (e.g., item price) and therefore, there are hardly trade-offs between multiple issues in such a negotiation.

Persuasion For Donation Persuasion, convincing others to take specific actions, is a necessary required skill for negotiation dialogue (Sycara, 1990; Sierra et al., 1997). Wang et al. (2019) focus on persuasion and propose *PersuasionforGood*, two-party persuasion conversations about charity donations. In the data annotation process, the persuaders are provided some persuasion tips and example sentences, while the persuaders are only told that this conversation is about charity. The annotators are required to complete at least ten utterances in a dialogue and are encouraged to reach an agreement at the end of the conversations. Dutt et al. (2020) further extend *PersuasionforGood* by adding the utterance-level annotations that change the positive and/or the negative face acts of the participants in a conversation. A face act can either raise or attack the positive or negative face of opponents in the conversation.

Negotiation For Product Price Negotiations over product prices can be observed on a daily basis. He et al. (2018) propose *CraigslistBargain*, a negotiation benchmark based on a realistic item price bargaining scenario. In *CraigslistBargain*, two agents, a buyer and a seller, are required to negotiate the price of a given item. The listing price is available to both sides, but the buyer has a private price. Two agents chat freely to decide on a final price. The conversation is completed when both agents agree on a price or one of the agents quits. Zhou et al. (2019) propose *NegoCoach* benchmark on similar scenarios, but with an additional negotiation coach who monitors messages between the two annotators and recommends tactics in real-time to the seller to get a better deal.

User Privacy Protection Privacy protection of negotiators has become more and more vital. Participant (e.g., attackers and defenders) goals are also conflicting. Li et al. (2020b) propose *Anti-Scam* benchmark which focuses on online customer service. In *Anti-Scam*, users try to defend themselves by identifying whether their components are attackers who try to steal sensitive personal information. *Anti-Scam* provides an opportunity to study human elicitation strategies in this scenario.

5 Evaluation

We categorize the methods for evaluating the negotiation dialogue systems into three types: *goal-oriented* evaluation, *game-based* evaluation and *human* evaluation. Table 2 summarizes the evaluation metrics that are introduced in our survey.

5.1 Goal-based Metrics

Goal-oriented metrics mainly refer to the quantifiable measures on evaluating agent's proximity to the negotiation goals from the perspective of strategy modeling, task fulfillment, and sentence realization. *Success Rate (SR)* (Zhao et al., 2019) is the most widely used metric to measure how frequently an agent completes the task within their goals. Meanwhile, *Prediction Accuracy (PA)* and *macro/average F1 score* are also employed to evaluate the accuracy of agent's strategy predictions (Nouri and Traum, 2014; Wang et al., 2019; Dutt et al., 2020; Chawla et al., 2021b). Specifically, Yamaguchi et al. (2021) present a task where the model is required to label the human-human negotiation outcomes as either a success or a breakdown, and use following metrics: *area under the curve* (ROC-AUC), *confusion matrix* (CM), and *average precision* (AP) to evaluate the model. Moreover, Kornilova et al. (2022) introduce Item Response Theory (IRT) to analyze the effectiveness of persuasion on the audience.

Goal-based Metrics	SR (2019); PA (2014; 2019; 2020); Average F1 score (2021b); Macro F1 score (2019; 2020); ROC-AUC, CM, AP (2021); IRT (2022); Naturalness (2015); PPL, BLEU-2, ROUGE-L, Extrema (2017)
Game-based Metrics	WinRate, AvgVPs (2017); Utility, Fairness, Length (2018); Avg. Sale-to-list Ratio, Task Completion Rate (2019); Robustness (2019)
Human Evaluation	Customer satisfaction, Purchase decision, Correct response rate (2015); Achieved agreement rate, Pareto optimality rate (2017); Likert score (2018)

Table 2: Various Metrics used in the existing negotiation dialogues benchmarks.

In terms of language realization for negotiation dialogue, Hiraoka et al. (2015) employ a predefined naturalness metric (i.e., a bi-gram overlap between the prediction and ground-truth) as part of the reward to evaluate policies in negotiation dialogues. Other classical metrics for evaluating the quality of response are also used, i.e., perplexity (PPL), BLEU-2, ROUGE-L, and BOW Embedding-based Extrema matching score (Lewis et al., 2017).

5.2 Game-based Metrics

Different from the goal-oriented metrics that focus on measuring how successful an agent achieves the negotiation goals, game-based evaluation provides a user-centric perspective to evaluate systems. Keizer et al. (2017) measure agent's ability on negotiation strategy prediction within the online game "*Settlers of Catan*". They propose the metrics *WinRate* and *AvgVPs* to evaluate the success of human and agent seperately. He et al. (2018) present a task where two agents bargain to get the best deal using natural language. They use task-specific scores to test the performance of the agents, including: *utility*, *fairness*, and *length*. Zhou et al. (2019) design a task where a seller and a buyer try to achieve a mutually acceptable price through a natural language negotiation. They adopt *average sale-to-list ratio* and *task completion rate* to evaluate agent performance. Besides, Cheng et al. (2019) propose an adversarial attacking evaluation approach to test the *robustness* of negotiation systems.

5.3 Human Evaluation

To evaluate the users' satisfaction with the dialogue systems, human judgment is employed as a subjective evaluation of agent performance. Hiraoka et al. (2015) use a user simulator as the salesperson to bargain with customers in real and have the users annotate subjective *customer satisfaction* (a five-level score), the final decision of making a purchase (a binary number indicating whether persuasion is successful), and the *correct response rate* in the dialogues. Lewis et al. (2017) employ crowd-sourcing workers to highlight that essential information when bargaining with negotiation systems, covering the percentage of dialogues where both interlocutors finally achieve an agreement, and *Pareto optimality*, i.e., the percentage of the Pareto optimal solutions in all the agreed deals. He et al. (2018) propose human likeness as a metric in evaluating how well the dialogue system is doing in a bargain. They ask workers to manually score the dialogue agent using a *Likert* metric to judge whether the agent acts like a real human or not.

6 New Frontiers and Challenges

The previous sections summarize the prominent achievements of previous work in negotiation dialogue, including benchmarks, evaluation metrics, and methodology. In this section, we will discuss some new frontiers that allow negotiation dialogue systems to be fit to actual application needs and to be applied in real-world scenarios.

Multi-modal Negotiation Dialogue Existing research works in negotiation dialogue rarely consider multi-modality. However, humans tend to perceive the world in multi-modal patterns, not limited to text but also including audio and visual information. For example, the facial expressions and emotions of participants in a negotiation dialogue could be important cues for making negotiation decisions. Further work can consider adding this non-text-based information into the negotiation.

Multi-Party Negotiation Dialogue Although some work sheds light on multi-party negotiation, most current negotiation dialogue benchmarks and methods predominantly focus on two-party settings. Therefore, multi-party negotiation dialogues are underexplored. Future work can consider collecting dialogues in multi-party negotiation scenarios, including *General multi-party negotiation* and *Team negotiation*. Specifically, *General multi-party negotiation* is a type of bargaining where more than two parties negotiate toward an agreement. For example, next-year budget discussion with multiple department leaders in a large company. *Team negotiation* is a team of people with different relationships and roles. It is normally associated with large business deals and highlights the significance of relationships between multi-parties. There could be several roles, including leader, recorder, and examiner, in a negotiation team (Halevy, 2008).

Cross-Culture & Multi-lingual Negotiation Dialogue Existing negotiation dialogue benchmarks overwhelmingly focus on English while leaving other languages and cultures underexplored. With the acceleration of globalization, a dialogue involving individuals from different cultural backgrounds becomes increasingly important and necessary. There is an urgent need to provide people with a negotiation dialogue system that is multicultural and multi-lingual. Further works can consider incorporating multi-lingual utterances and social norms among different countries into negotiation dialogue benchmarks.

Negotiation Dialogue in Real-world Scenarios As discussed in Section 4, previous works have already proposed many negotiation dialogue benchmarks in various scenarios. However, we notice that most of these benchmarks are created through human crowd-sourcing. Participants are often invited to play specific roles in the negotiation dialogue. The resulting dialogues may not perfectly reflect the negotiations in real-world scenarios (e.g., politics, business). Therefore, it could be a promising research direction to collect real-world negotiation dialogues. For example, one could collect recorded business meetings or phone calls.

7 Conclusion

This paper presents the first systematic review on the progress of negotiation dialogue systems. We firstly provide an understanding of negotiation between humans from a social science perspective. Then we thoroughly summarize the existing works, which covers various domains and highlight their challenges, respectively. We additionally summarize currently available methodologies, benchmarks, and evaluation methods. We also shed light on some new trends in this research field. We hope this survey inspires and facilitates future research on negotiation dialogue systems.

Limitations

This survey briefly introduced the motivation and limitation of human negotiation from social science perspectives, and summarized methodology, dataset and evaluation methods in the field of computational linguistics. The limitation relays on that we only have brief investigation on the human negotiation. Further, we will conduct a comprehensive investigation from the social science perspectives and then motivate our future work in the dialogue research. In further, we will summarize the details of each paper and illustrate the difference between these papers. Nevertheless, we hope that our survey will inspire and facilitate future research as a good foundation.

References

Stergos Afantenos, Nicholas Asher, Farah Benamara, Anaıs Cadilhac, Cedric Dégremont, Pascal Denis, Markus Guhe, Simon Keizer, Alex Lascarides, Oliver Lemon, et al. 2012. Modelling strategic conversation: model, annotation design and corpus. In *Proceedings of the 16th Workshop on the Semantics and Pragmatics of Dialogue (Seinedial), Paris*.

Ron Artstein, Jill Boberg, Alesia Gainer, Jonathan Gratch, Emmanuel Johnson, Anton Leuski, Gale Lucas, and David Traum. 2018. The niki and julie corpus: collaborative multimodal dialogues between humans, robots, and virtual agents. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

Nicholas Asher, Julie Hunter, Mathieu Morey, Benamara Farah, and Stergos Afantenos. 2016. Discourse structure and dialogue acts in multiparty dialogue: the STAC corpus. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 2721–2727, Portorož, Slovenia. European Language Resources Association (ELRA).

Max H Bazerman, Jared R Curhan, Don A Moore, and Kathleen L Valley. 2000. Negotiation. *Annual review of psychology*, 51(1):279–314.

Max H Bazerman and Margaret Ann Neale. 1993. *Negotiating rationally*. Simon and Schuster.

Steven L Blader and Ya-Ru Chen. 2012. Differentiating the effects of status and power: a justice perspective. *Journal of personality and social psychology*, 102(5):994.

Maria Boritchev and Maxime Amblard. 2022. A multi-party dialogue ressource in French. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 814–823, Marseille, France. European Language Resources Association.

Jeanne Brett and Leigh Thompson. 2016. Negotiation. *Organizational Behavior and Human Decision Processes*, 136:68–79.

Amparo Elizabeth Cano-Basave and Yulan He. 2016. A study of the impact of persuasive argumentation in political debates. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1405–1413, San Diego, California. Association for Computational Linguistics.

Kushal Chawla, Rene Clever, Jaysa Ramirez, Gale Lucas, and Jonathan Gratch. 2021a. Towards emotion-aware agents for negotiation dialogues. In *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 1–8. IEEE.

Kushal Chawla, Gale Lucas, Jonathan May, and Jonathan Gratch. 2022. Opponent modeling in negotiation dialogues by related data adaptation. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 661–674, Seattle, United States. Association for Computational Linguistics.

Kushal Chawla, Jaysa Ramirez, Rene Clever, Gale Lucas, Jonathan May, and Jonathan Gratch. 2021b. 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, Online. Association for Computational Linguistics.

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, Minneapolis, Minnesota. Association for Computational Linguistics.

Jennifer Crocker. 1982. Biased questions in judgment of covariation studies. *Personality and Social Psychology Bulletin*, 8(2):214–220.

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, Minneapolis, Minnesota. Association for Computational Linguistics.

Ritam Dutt, Rishabh Joshi, and Carolyn Rose. 2020. Keeping up appearances: Computational modeling of face acts in persuasion oriented discussions. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7473–7485, Online. Association for Computational Linguistics.

Ritam Dutt, Sayan Sinha, Rishabh Joshi, Surya Shekhar Chakraborty, Meredith Riggs, Xinru Yan, Haogang Bao, and Carolyn Rose. 2021a. ResPer: Computationally modelling resisting strategies in persuasive conversations. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 78–90, Online. Association for Computational Linguistics.

Ritam Dutt, Sayan Sinha, Rishabh Joshi, Surya Shekhar Chakraborty, Meredith Riggs, Xinru Yan, Haogang Bao, and Carolyn Rose. 2021b. ResPer: Computationally modelling resisting strategies in persuasive conversations. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 78–90, Online. Association for Computational Linguistics.

Michael English and Peter Heeman. 2005. Learning mixed initiative dialog strategies by using reinforcement learning on both conversants. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 1011–1018, Vancouver, British Columbia, Canada. Association for Computational Linguistics.

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

Roger Fisher, William L Ury, and Bruce Patton. 2011. *Getting to yes: Negotiating agreement without giving in*. Penguin.

Marieke L Fransen, Edith G Smit, and Peeter WJ Verlegh. 2015. Strategies and motives for resistance to persuasion: An integrative framework. *Frontiers in psychology*, 6:1201.

Chris Frith and Uta Frith. 2005. Theory of mind. *Current biology*, 15(17):R644–R645.

Yao Fu, Hao Peng, Tushar Khot, and Mirella Lapata. 2023. Improving language model negotiation with self-play and in-context learning from ai feedback. *arXiv preprint arXiv:2305.10142*.

Xiaoyang Gao, Siqi Chen, Yan Zheng, and Jianye Hao. 2021. A deep reinforcement learning-based agent for negotiation with multiple communication channels. In *2021 IEEE 33rd International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 868–872. IEEE.

Kallirroi Georgila, Claire Nelson, and David Traum. 2014. Single-agent vs. multi-agent techniques for concurrent reinforcement learning of negotiation dialogue policies. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 500–510, Baltimore, Maryland. Association for Computational Linguistics.

Nir Halevy. 2008. Team negotiation: Social, epistemic, economic, and psychological consequences of subgroup conflict. *Personality and Social Psychology Bulletin*, 34(12):1687–1702.

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, Brussels, Belgium. Association for Computational Linguistics.

Mourad Heddaya, Solomon Dworkin, Chenhao Tan, Rob Voigt, and Alexander Zentefis. 2023. Language of bargaining. *arXiv preprint arXiv:2306.07117*.

Takuya Hiraoka, Graham Neubig, Sakriani Sakti, Tomoki Toda, and Satoshi Nakamura. 2015. Evaluation of a fully automatic cooperative persuasive dialogue system. In *Natural Language Dialog Systems and Intelligent Assistants, 6th International Workshop on Spoken Dialogue Systems, IWSDS 2015, Busan, Korea, January 11-13, 2015*, pages 153–167. Springer.

Rishabh Joshi, Vidhisha Balachandran, Shikhar Vashishth, Alan W. Black, and Yulia Tsvetkov. 2021. Dialograph: Incorporating interpretable strategy-graph networks into negotiation dialogues. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net.

Georgi Karadzhov, Tom Stafford, and Andreas Vlachos. 2021. Delidata: A dataset for deliberation in multi-party problem solving. *arXiv preprint arXiv:2108.05271*.

Simon Keizer, Markus Guhe, Heriberto Cuayáhuitl, Ioannis Efstathiou, Klaus-Peter Engelbrecht, Mihai Dobre, Alex Lascarides, and Oliver Lemon. 2017. Evaluating persuasion strategies and deep reinforcement learning methods for negotiation dialogue agents. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 480–484, Valencia, Spain. Association for Computational Linguistics.

Dimosthenis Kontogiorgos, Vanya Avramova, Simon Alexanderson, Patrik Jonell, Catharine Oertel, Jonas Beskow, Gabriel Skantze, and Joakim Gustafson. 2018. A multimodal corpus for mutual gaze and joint attention in multiparty situated interaction. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

Anastassia Kornilova, Vladimir Eidelman, and Daniel Douglass. 2022. An item response theory framework for persuasion. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 77–86, Seattle, United States. Association for Computational Linguistics.

Lynn Lambert and Sandra Carberry. 1992. Modeling negotiation subdialogues. In *30th Annual Meeting of the Association for Computational Linguistics*, pages 193–200, Newark, Delaware, USA. Association for Computational Linguistics.

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.

Angela K-Y Leung and Dov Cohen. 2011. Within-and between-culture variation: individual differences and the cultural logics of honor, face, and dignity cultures. *Journal of personality and social psychology*, 100(3):507.

Barbara Lewandowska. 1982. Meaning negotiation in dialogue. In *Coling 1982 Abstracts: Proceedings of the Ninth International Conference on Computational Linguistics Abstracts*.

Roy J Lewicki, David M Saunders, John W Minton, J Roy, and Negotiation Lewicki. 2011. *Essentials of negotiation*. McGraw-Hill/Irwin Boston, MA, USA:.

Mike Lewis, Denis Yarats, Yann Dauphin, Devi Parikh, and Dhruv Batra. 2017. Deal or no deal? end-to-end learning of negotiation dialogues. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2443–2453, Copenhagen, Denmark. Association for Computational Linguistics.

Jialu Li, Esin Durmus, and Claire Cardie. 2020a. Exploring the role of argument structure in online debate persuasion. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8905–8912, Online. Association for Computational Linguistics.

Jiaqi Li, Ming Liu, Zihao Zheng, Heng Zhang, Bing Qin, Min-Yen Kan, and Ting Liu. 2021. Dadgraph: A discourse-aware dialogue graph neural network for multiparty dialogue machine reading comprehension. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

Yu Li, Kun Qian, Weiyan Shi, and Zhou Yu. 2020b. End-to-end trainable non-collaborative dialog system. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 8293–8302. AAAI Press.

Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang, and Minlie Huang. 2021. Towards emotional support dialog systems. 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 3469–3483, Online. Association for Computational Linguistics.

Kshitij Mishra, Azlaan Mustafa Samad, Palak Totala, and Asif Ekbal. 2022. PEPDS: A polite and empathetic persuasive dialogue system for charity donation. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 424–440, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

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.

Elnaz Nouri and David Traum. 2014. Initiative taking in negotiation. In *Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 186–193, Philadelphia, PA, U.S.A. Association for Computational Linguistics.

Mara Olekalns and Philip L Smith. 2003. Testing the relationships among negotiators' motivational orientations, strategy choices, and outcomes. *Journal of experimental social psychology*, 39(2):101–117.

Lydia Ould Ouali, Nicolas Sabouret, and Charles Rich. 2017. A computational model of power in collaborative negotiation dialogues. In *International Conference on Intelligent Virtual Agents*, pages 259–272. Springer.

Robin L Pinkley and Gregory B Northcraft. 1994. Conflict frames of reference: Implications for dispute processes and outcomes. *Academy of management journal*, 37(1):193–205.

Sudeep Sharma, William P Bottom, and Hillary Anger Elfenbein. 2013. On the role of personality, cognitive ability, and emotional intelligence in predicting negotiation outcomes: A meta-analysis. *Organizational Psychology Review*, 3(4):293–336.

Weiyan Shi, Yu Li, Saurav Sahay, and Zhou Yu. 2021. Refine and imitate: Reducing repetition and inconsistency in persuasion dialogues via reinforcement learning and human demonstration. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 3478–3492, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Zhouxing Shi and Minlie Huang. 2019. A deep sequential model for discourse parsing on multi-party dialogues. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*, pages 7007–7014. AAAI Press.

Carles Sierra, Nick R Jennings, Pablo Noriega, and Simon Parsons. 1997. A framework for argumentation-based negotiation. In *International Workshop on*

Agent Theories, Architectures, and Languages, pages 177–192. Springer.

Katia P Sycara. 1990. Persuasive argumentation in negotiation. *Theory and decision*, 28(3):203–242.

Shelley E Taylor. 1989. *Positive illusions: Creative self-deception and the healthy mind.* Basic Books/Hachette Book Group.

Nhat Tran, Malihe Alikhani, and Diane Litman. 2022. How to ask for donations? learning user-specific persuasive dialogue policies through online interactions. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, pages 12–22.

David Traum, Stacy C Marsella, Jonathan Gratch, Jina Lee, and Arno Hartholt. 2008. Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents. In *International workshop on intelligent virtual agents*, pages 117–130. Springer.

Richard E Walton and Robert B McKersie. 1991. *A behavioral theory of labor negotiations: An analysis of a social interaction system*. Cornell University Press.

Xuewei Wang, Weiyan Shi, Richard Kim, Yoojung Oh, Sijia Yang, Jingwen Zhang, and Zhou Yu. 2019. Persuasion for good: Towards a personalized persuasive dialogue system for social good. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5635–5649, Florence, Italy. Association for Computational Linguistics.

Laurie R Weingart, Leigh L Thompson, Max H Bazerman, and John S Carroll. 1990. Tactical behavior and negotiation outcomes. *International Journal of Conflict Management*.

Yuzhuang Xu, Shuo Wang, Peng Li, Fuwen Luo, Xiaolong Wang, Weidong Liu, and Yang Liu. 2023. Exploring large language models for communication games: An empirical study on werewolf. *arXiv preprint arXiv:2309.04658*.

Atsuki Yamaguchi, Kosui Iwasa, and Katsuhide Fujita. 2021. Dialogue act-based breakdown detection in negotiation dialogues. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 745–757, Online. Association for Computational Linguistics.

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, Online. Association for Computational Linguistics.

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. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, pages 122–132, Suzhou, China. Association for Computational Linguistics.

Tiancheng Zhao, Kaige Xie, and Maxine Eskenazi. 2019. Rethinking action spaces for reinforcement learning in end-to-end dialog agents with latent variable models. 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 1208–1218, Minneapolis, Minnesota. Association for Computational Linguistics.

Yiheng Zhou, He He, Alan W Black, and Yulia Tsvetkov. 2019. A dynamic strategy coach for effective negotiation. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*, pages 367–378, Stockholm, Sweden. Association for Computational Linguistics.

Yiheng Zhou, Yulia Tsvetkov, Alan W. Black, and Zhou Yu. 2020. Augmenting non-collaborative dialog systems with explicit semantic and strategic dialog history. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net.