Introducing Dialogue-Act Framework for Multi-Agent LLM Negotiation

Takafumi Shimizu[1], Ryo Furukawa[2], Yusuke Muraki[2], Eisuke Sonomoto[2], Kento Yamamoto[2], Rikuho Takeda[2], Kazuma Fukumura[1], Yuta Konishi[1], Takayuki Ito[1], Yuya Kanemoto[2], Shiyao Ding[1]

[1]Graduate School of Informatics, Kyoto University, Kyoto, Japan
[2]Accenture Japan Ltd, Tokyo, Japan
shimizu.takafumi.65x@st.kyoto-u.ac.jp, ito@i.kyoto-u.ac.jp, yuya.kanemoto@accenture.com, ding@i.kyoto-u.ac.jp

Abstract

Negotiation presents a challenging benchmark for LLM-based multi-agent systems, as it requires not only local decision–making but also joint reasoning, Theory of Mind (ToM), and strategic adaptation across agents. However, existing LLM-based negotiation systems mainly rely on discrete offer exchanges, which constrain interaction richness and lead to inefficient or repetitive bargaining. To address this limitation, we propose a Dialogue-Act negotiation framework that represents each agent's next action using a finite and interpretable set of communicative acts (e.g., offer, inquire, inform). By separating strategic action selection from natural-language generation, our approach enables controllable behavior while preserving the expressiveness of free-form dialogue. Experiments in multi-issue, multi-party negotiation settings show that our framework reduces negotiation rounds, lowers redundant offers, and yields higher joint welfare compared with offer-only baselines. We also observe that agents sometimes converge prematurely due to social conformity, suggesting a trade-off between efficiency and decision reliability.

Introduction

Large language models (LLMs) are increasingly studied as autonomous agents in open-ended, collaborative tasks. LLM-based multi-agent negotiation is a testbed requiring strategic reasoning (Kwon et al. 2024; Gandhi, Sadigh, and Goodman 2023), and Theory-of-Mind (ToM) (Chan et al. 2024; Yu et al. 2025). While much prior work is bilateral (Davidson et al. 2024; Kwon et al. 2024; Xia et al. 2024), recent studies move to more complex multilateral, multi-issue settings. Notably, (Abdelnabi et al. 2025) proposed an environment where multiple LLM agents engage in multi-issue negotiation with scorable outcomes. However, the agent action space was confined to proposing offers, limiting information exchange and strategic exploration.

Our goal is to enable an efficient search for agreements in natural-language interaction among multiple LLM agents. Unlike traditional automated negotiation, which exchanges discrete or numeric offers (Rubinstein 1982; Aydoğan et al. 2017), natural-language interaction can help agents to reason, explain, and infer preferences more flexibly like humans-yet it also expands the action space dramatically. To leverage LLMs' linguistic competence and ToM-based inference while maintaining controllability, we introduce Dialogue-Act as a structured framework that defines a finite set of negotiation actions. Following prior decoupling approaches (He et al. 2018; Yang, Chen, and Narasimhan 2021) that separate strategic act selection from surface generation via a finite action set, we adapt this idea to LLM-based negotiation to keep interactions interpretable and expressive while bounding the action space.

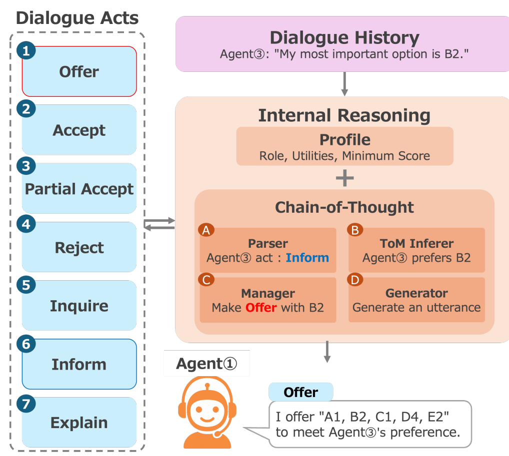

Figure 1: An agent's internal reasoning steps. Parser classifies peer utterances into Dialogue Acts; ToM Inferer updates counterparts' belief states; Manager selects the next Dialogue Act; Generator realizes the act in natural language.

The proposed approach reaches unanimous agreements more often and with fewer rounds than the baselines, achieving more balanced, high-utility outcomes. It efficiently explores feasible deals without redundant repetitions, though natural-language interaction occasionally caused conformity-driven erroneous accept decisions. This work contributes a Dialogue-Act framework that introduces natural-language interaction into multi-agent LLM negotiation in a controllable yet expressive form, enabling efficient and balanced agreement search. It also reveals conformity-driven failures as a key trade-off between coordination and robustness.

Method

Problem Settings

We consider a multi-party, multi-issue negotiation game. Agents negotiate over several issues, each with a finite set of options; an offer is a complete deal that selects one option per issue. Each agent has a private utility and a minimum acceptable threshold. Agents take turns speaking and proposing deals toward consensus. A session ends early when a proposed deal receives unanimous Accept from all remaining agents; otherwise it stops at a fixed round limit. After termination, we score the final deal for each agent, and the negotiation is deemed successful if every agent's score meets its own threshold.

Introducing Dialogue Acts

We map free-form utterances to a finite set of Dialogue Acts to keep actions bounded and interactions interpretable. We retain Offer, Accept, and Reject as the transactional core following alternating-offer protocols (Aydoğan et al. 2017; Rubinstein 1982). Partial-Accept captures multi-issue structure by allowing agents to lock in agreed subsets under asymmetric valuations and deadlines (Fatima, Wooldridge, and Jennings 2006). To enable preference elicitation and justification without numeric utilities, we add Inquire, Inform, and Explain (Bunt et al. 2020). Together, these acts standardize communication and reasoning, reducing redundant search and improving outcome quality.

Internal Reasoning of LLM Agents

As an agent's internal reasoning, we implement a four-stage pipeline—**Parser**, **ToM Inferer**, **Manager**, and **Generator**—, following prior approaches (He et al. 2018; Yang, Chen, and Narasimhan 2021). This pipeline is realized with a lightweight chain-of-thought (CoT) prompting (Wei et al. 2022) scheme that separates understanding, belief updating, decision, and generation, as shown in Figure 1. **Parser** maps each peer utterance to a Dialogue Act. **ToM Inferer** estimates each counterpart's belief state, which represents what issues/options they value, how satisfied they are, and what changes or information they seek. It maintains this inferred knowledge in a JSON-like format and updates it each round based on parsed Dialogue Acts and observed offers. **Manager** selects the next act based on the ToM inference. **Generator** realizes the selected act in natural language.

Experiments

Setup and Metrics

We evaluate three-party negotiation with a 42-round limit, running 100 trials and reporting mean values across all sessions. We measure four metrics. *Agreement rate (Agr)* is the proportion of sessions that end in a unanimous Accept. *Success rate (Suc)* counts sessions where the final deal satisfies all agents' minimum thresholds. For successful negotiations in each method, we measure *Rounds to agreement (Rnd)* as the number of turns until consensus and *Nash product (Nash)* as the joint welfare, calculated as the product of agents' utilities for the final deal. We use GPT-4.1 as the LLM for agents.

Table 1: Mean results of each method. Bold indicates the best performance. For Baseline1, explicit Accept is unavailable and all sessions run to the round limit.

Method	Agr ↑	Suc ↑	Rnd ↓	Nash ↑
Baseline1	–	**0.790**	–	2273.00
Baseline2	0.270	0.750	36.56	1001.65
Proposed	**0.980**	0.768	**17.26**	**3169.70**

Methods Compared

We compare three methods; Baseline 1 follows (Abdelnabi et al. 2025)'s offer-only protocol, whereas the other two adopt our protocol with unanimous Accept termination.

Baseline 1 (Offer-only): Agents' Dialogue Act is limited only to Offer. There is no Accept act; the session terminates at the final round of the schedule produced by a lightweight planning scratchpad, and metrics are computed on the standing deal at that time. **Baseline 2 (Offer and Accept only):** Agents' Dialogue Acts are restricted to Offer and Accept only. This variant is based on existing research in automated negotiation (Aydoğan et al. 2017). **Proposed (Dialogue-Act Framework):** Following the internal reasoning steps, agents select their next move from seven Dialogue Acts.

Results

Table 1 shows that the Proposed agent reached unanimous agreements more often and required fewer rounds than the baselines. It achieved the highest Nash product among successful cases, indicating more balanced, high-utility outcomes. While Baseline 1 and 2 often repeated identical deals, the Proposed method efficiently explored feasible agreements without redundancy. However, natural-language interaction occasionally led to erroneous accept decisions, as agents sometimes recognized sub-threshold deals yet conformed to others' decisions.

Conclusions

The results suggest that defining a finite set of Dialogue Acts promotes more frequent unanimous agreements and leads to faster and more balanced outcomes than offer-only baselines. However, there remains room for improvement in success rate, and we have not yet verified whether the proposed Dialogue-Act taxonomy is the most general or effective formulation in other game settings. Examining different configurations of Dialogue Acts and reasoning modules reveals insights for designing more versatile multi-agent frameworks. We will vary act definitions and set size to find effective inventories, and adjust the pipeline to identify which reasoning modules drive gains in convergence speed and outcome quality, informing more versatile multi-agent negotiation frameworks. Moreover, we aim to implement more sophisticated ToM reasoning to enable deeper strategic negotiation and to conduct human-in-the-loop evaluations focusing on interpretability and interaction quality, toward realizing practical, real-world applications of LLM-based negotiation agents using natural language.

References

Abdelnabi, S.; Gomaa, A.; Sivaprasad, S.; Schönherr, L.; and Fritz, M. 2025. Cooperation, competition, and maliciousness: LLM-stakeholders interactive negotiation. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, NIPS '24. Red Hook, NY, USA: Curran Associates Inc. ISBN 9798331314385.

Aydoğan, R.; Festen, D.; Hindriks, K. V.; and Jonker, C. M. 2017. Alternating offers protocols for multilateral negotiation. In *Modern approaches to agent-based complex automated negotiation*, 153–167. Springer.

Bunt, H.; Petukhova, V.; Gilmartin, E.; Pelachaud, C.; Fang, A.; Keizer, S.; and Prévot, L. 2020. The ISO Standard for Dialogue Act Annotation, Second Edition. In Calzolari, N.; Béchet, F.; Blache, P.; Choukri, K.; Cieri, C.; Declerck, T.; Goggi, S.; Isahara, H.; Maegaard, B.; Mariani, J.; Mazo, H.; Moreno, A.; Odijk, J.; and Piperidis, S., eds., *Proceedings of the Twelfth Language Resources and Evaluation Conference*, 549–558. Marseille, France: European Language Resources Association. ISBN 979-10-95546-34-4.

Chan, C.; Jiayang, C.; Yim, Y.; Deng, Z.; Fan, W.; Li, H.; Liu, X.; Zhang, H.; Wang, W.; and Song, Y. 2024. NegotiationToM: A Benchmark for Stress-testing Machine Theory of Mind on Negotiation Surrounding. In Al-Onaizan, Y.; Bansal, M.; and Chen, Y.-N., eds., *Findings of the Association for Computational Linguistics: EMNLP 2024*, 4211–4241. Miami, Florida, USA: Association for Computational Linguistics.

Davidson, T. R.; Veselovsky, V.; Kosinski, M.; and West, R. 2024. Evaluating Language Model Agency Through Negotiations. In *The Twelfth International Conference on Learning Representations*.

Fatima, S. S.; Wooldridge, M.; and Jennings, N. R. 2006. Multi-issue negotiation with deadlines. *J. Artif. Int. Res.*, 27(1): 381–417.

Gandhi, K.; Sadigh, D.; and Goodman, N. 2023. Strategic Reasoning with Language Models. In *NeurIPS 2023 Foundation Models for Decision Making Workshop*.

He, H.; Chen, D.; Balakrishnan, A.; and Liang, P. 2018. Decoupling Strategy and Generation in Negotiation Dialogues. In Riloff, E.; Chiang, D.; Hockenmaier, J.; and Tsujii, J., eds., *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2333–2343. Brussels, Belgium: Association for Computational Linguistics.

Kwon, D.; Weiss, E.; Kulshrestha, T.; Chawla, K.; Lucas, G.; and Gratch, J. 2024. Are LLMs Effective Negotiators? Systematic Evaluation of the Multifaceted Capabilities of LLMs in Negotiation Dialogues. In Al-Onaizan, Y.; Bansal, M.; and Chen, Y.-N., eds., *Findings of the Association for Computational Linguistics: EMNLP 2024*, 5391–5413. Miami, Florida, USA: Association for Computational Linguistics.

Rubinstein, A. 1982. Perfect Equilibrium in a Bargaining Model. *Econometrica*, 50(1): 97–109.

Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Ichter, B.; Xia, F.; Chi, E. H.; Le, Q. V.; and Zhou, D. 2022. Chain-of-thought prompting elicits reasoning in large language models. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22. Red Hook, NY, USA: Curran Associates Inc. ISBN 9781713871088.

Xia, T.; He, Z.; Ren, T.; Miao, Y.; Zhang, Z.; Yang, Y.; and Wang, R. 2024. Measuring Bargaining Abilities of LLMs: A Benchmark and A Buyer-Enhancement Method. In Ku, L.-W.; Martins, A.; and Srikumar, V., eds., *Findings of the Association for Computational Linguistics: ACL 2024*, 3579–3602. Bangkok, Thailand: Association for Computational Linguistics.

Yang, R.; Chen, J.; and Narasimhan, K. 2021. Improving Dialog Systems for Negotiation with Personality Modeling. In Zong, C.; Xia, F.; Li, W.; and Navigli, R., eds., *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)*, 681–693. Online: Association for Computational Linguistics.

Yu, F.; Jiang, L.; Huang, S.; Wu, Z.; and Dai, X. 2025. PersuasiveToM: A Benchmark for Evaluating Machine Theory of Mind in Persuasive Dialogues. arXiv:2502.21017.