Sentimental Agents: Combining Sentiment Analysis and Non-Bayesian Updating for Cooperative Decision-Making

Extended Abstract

Daniele Orner*
Brave Venture Labs
London, United Kingdom
daniele@braveventurelabs.com

Elizabeth Akinyi Ondula*
University of Southern California
Los Angeles, USA
ondula@usc.edu

Nick Mumero Mwangi
Brave Venture Labs
Nairobi, Kenya
nick@braveventurelabs.com

Richa Goyal
University of Southern California
Los Angeles, CA, USA
richagoy@usc.edu

ABSTRACT

With ongoing exploration of Large Language Model(LLM)-based multi-agent systems, it is becoming increasingly important to understand and interpret the dynamics of agent interactions and their beliefs, particularly when designed to emulate diverse roles and perspectives or to engage in debates. At present, there are no unified solutions that can systematically interpret and analyze the beliefs and interactions of these agents. This study introduces *Sentimental Agents*, a framework designed to support decision-making by providing multiple perspectives on a topic. Agents within this framework are equipped with a mental model of self, articulated in natural language. We have integrated sentiment analysis with a non-Bayesian updating mechanism to interpret and analyze the agents' beliefs and interactions systematically. A collective viewpoint is achieved when the update is marginal. We have adapted this framework for a simulated scenario in the Human Resource domain, implementing a conceptual tool known as the *Artificial Board of Advisors (ABA)*. A key focus of this simulation is the application of ABA in the assessment of candidates for roles, showcasing its potential application in a theoretical HR recruiting environment.

KEYWORDS

Multi-Agent Systems; Large Language Models; Sentiment Analysis; Opinion Dynamics; Non-Bayesian Updating; Cooperative Engagement; Decision Science

ACM Reference Format:
Daniele Orner*, Elizabeth Akinyi Ondula*, Nick Mumero Mwangi, and Richa Goyal. 2024. Sentimental Agents: Combining Sentiment Analysis and Non-Bayesian Updating for Cooperative Decision-Making: Extended Abstract. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

* indicates equal contribution.

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

1 INTRODUCTION

This paper investigates how Large Language Models (LLMs) such as GPT can enhance decision-making processes, especially when multiple perspectives on a topic are indispensable. We present *Sentimental Agents*, an LLM-based multi-agent system. Its design supports decision-making without making decisions itself. The focus is on gathering diverse opinions and perspectives. Central to this system is a structured, systematic conversation in a controlled environment. We apply the framework to a prototype tool, *Artificial Board of Advisors* (ABA), demonstrating its practical applicability in a simulated HR scenario. ABA is built around 'expert' agents, each providing unique opinions and engaging in a conversation. These interactions are focused on evaluating a job candidate, culminating in a collective stance. We present preliminary outcomes and an evaluation based on this case study.

LLMs have shown remarkable capabilities in generating text that embodies sentiment, and in executing sentiment analysis tasks [4]. Yet, the influence of sentiment on the dynamics of opinion formation within artificial societies of agents remains underexplored [3], [18], [2], [14], [10], [16], [5], [15] and [6]. Our work adopts a nuanced approach to understand how the output of LLM agents affects each other within a multi-agent framework.

This work makes the following contributions:

- We develop and apply the framework, *Sentimental Agents*, to explore and study deliberation processes in a society of conversational agents.
- We propose using sentiment analysis as a method to quantify content generated by LLM-based agents for evaluation and recommendation tasks.
- We apply a non-Bayesian updating method, as a non-intrusive, non-strategic descriptive tool to observe changes in agents' opinions as they interact and potentially influence each other.

2 SYSTEM ARCHITECTURE

The main components of the framework include:

(1) **System Initialization**, which takes as an input a *Brief* describing the task in natural language. It also includes initializing the expert agents, assigning them a unique role and a

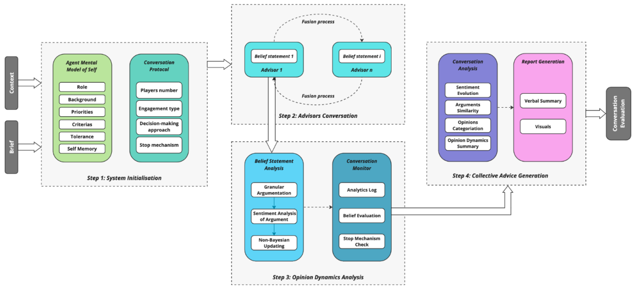

Figure 1: *System Architecture showing a phased implementation of the ABA tool.*

Mental Model of Self (MMS) [13] [11]. The MMS is designed as a series of prompts that gives the agents a distinct, specialized function, as well as priorities and evaluation criteria.

(2) **Conversation Protocol** establishes a controlled environment for structured agent interactions, with cooperative engagement type set to ordered. It features a stopping mechanism where conversations end upon minimal changes or after a maximum number of rounds. Agents follow a non-interventionist, non-strategic approach, not weighting opinions or aiming for consensus.

(3) **Opinion Dynamics Analysis** begins with the belief analysis process. It involves analyzing advisors' opinions by breaking them down into arguments, assessing each argument's sentiment, and calculating an average sentiment across the entire opinion. This process uses non-Bayesian updating, inspired by DeGroot's model [8], to track sentiment changes and update agents' sentiments per round through a weighted average. Additionally, it assesses argument quality using metrics like the *Platitudinal Score* to measure how original an agent's arguments are, compared with its interlocutors.

(4) **Collective Advice Generation**, which is activated once the conversation concludes. It involves a *Conversation Analysis*, in which arguments are grouped based on criteria like similarity, agreement levels among agents, strength of sentiments on particular topics, or intensity of debate. This steps concludes with a report generation in textual format.

3 PRELIMINARY RESULTS

The experiment employed OpenAI's GPT 3.5 Turbo [1], with a 16K token limit, and integrated a Non-Bayesian update mechanism in the dialogues. Key parameters included an alpha of 0.7 for the Non-Bayesian Update and a Tolerance level set at 0.01, with sentiment values ranging between -1 and 1. Advisor Description, Advisor Priorities, and Evaluation Criteria prompts were limited to 10 words or 5 bullet points, and the GPT model temperature was fixed at 0.5 for generating Advisor profiles. Figure 2 illustrates the inter-agent similarities heatmap. This shows a contrast in sentiment alignment among agents, leading to a lower overall platitudinal score. The observation from the non-Bayesian updating data collected during the simulation run, as shown in Figure 4 and Figure 3 reveal fluctuations in sentiment among the agents. The volatility, evident in both sentiment and change metrics, highlights the dynamic nature of opinion formation in multi-agent conversations. The agents[1] initially presented ideas specific to their roles, priorities, and evaluation criteria. While consensus was rare, they sometimes changed their opinions and sentiments influenced by other agents' arguments. Notably, after multiple rounds of interaction influencing beliefs, advisors often shifted their MMS, losing track of their original roles.

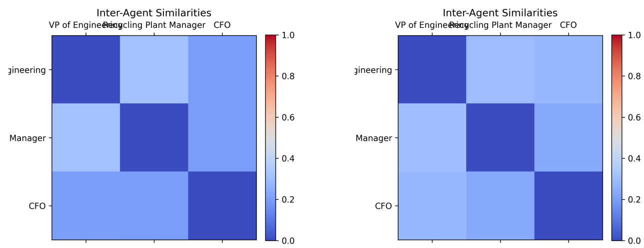

Figure 2: The uniqueness of outcomes in the conversation rounds among agents (a lower score indicates a more original contribution)

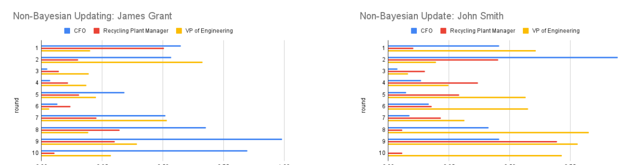

Figure 3: Opinion Change in conversations about two different candidates.

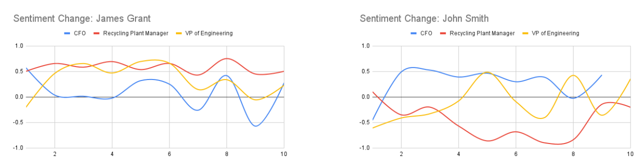

Figure 4: Sentiment Change in conversations about two different candidates.

4 CONCLUSION

The Sentimental Agents framework serves as a foundational observational tool to analyze LLM-based agent outputs in a multi-agent setting. Our roadmap includes the exploration of diverse conversational protocols, decision-making approaches, and report structures, complemented by extensive experiments on bias testing and real-world applications.

[1] In our experiment, agents on the Artificial Board of Advisors included a CFO, a VP of Engineering and a Recycling Plant Manager. They reviewed a series of candidates for the role of Head of Tech in a recycling plant.

5 ETHICS STATEMENT

It is important to acknowledge the inherent biases in Large Language Models (LLMs), including but not limited to demographic, cultural, linguistic, and temporal biases. LLMs may exhibit social biases and toxicities during the generation process, leading to biased outputs [17], [7], [9] and [12]. This acknowledgment is critical in understanding the limitations and ethical considerations of deploying such technology. The application of the framework, exemplified in the simulation scenario in the Human Resource domain is not intended to replace human judgement but to augment it. We advocate for a balanced approach where AI-generated insights complement, rather than dictate, human decision-making processes in organizations. Consequently, our framework is not designed for commercial application but rather serves as a tool for research and exploration. Its primary objective is to evaluate and potentially mitigate the pitfalls of LLM-based multi-agent applications in decision-making processes. By doing so, we aim to contribute to the responsible and ethical development of LLM-based agents, particularly in domains like Human Resources, where the implications of bias and ethical considerations are significant.

REFERENCES

[1] 2023. https://platform.openai.com/docs/models/gpt-3-5. Accessed: 2023-10-9.
[2] Gregor Betz. 2021. Natural-language multi-agent simulations of argumentative opinion dynamics. *arXiv preprint arXiv:2104.06737* (2021).
[3] Chi-Min Chan, Weize Chen, Yusheng Su, Jianxuan Yu, Wei Xue, Shanghang Zhang, Jie Fu, and Zhiyuan Liu. 2023. Chateval: Towards better llm-based evaluators through multi-agent debate. *arXiv preprint arXiv:2308.07201* (2023).
[4] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Kaijie Zhu, Hao Chen, Linyi Yang, Xiaoyuan Yi, Cunxiang Wang, and Yidong Wang. 2023. A survey on evaluation of large language models. *arXiv preprint arXiv:2307.03109* (2023). https://arxiv.org/abs/2307.03109
[5] Weize Chen, Yusheng Su, Jingwei Zuo, Cheng Yang, Chenfei Yuan, Chen Qian, Chi-Min Chan, Yujia Qin, Yaxi Lu, Ruobing Xie, et al. 2023. Agentverse: Facilitating multi-agent collaboration and exploring emergent behaviors in agents. *arXiv preprint arXiv:2308.10848* (2023).
[6] Yun-Shiuan Chuang, Agam Goyal, Nikunj Harlalka, Siddharth Suresh, Robert Hawkins, Sijia Yang, Dhavan Shah, Junjie Hu, and Timothy T. Rogers. 2023. Simulating Opinion Dynamics with Networks of LLM-based Agents. http://arxiv.org/abs/2311.09618 arXiv:2311.09618 [physics].
[7] Maria De-Arteaga, Alexey Romanov, Hanna Wallach, Jennifer Chayes, Christian Borgs, Alexandra Chouldechova, Sahin Geyik, Krishnaram Kenthapadi, and Adam Tauman Kalai. 2019. Bias in bios: A case study of semantic representation bias in a high-stakes setting. In *proceedings of the Conference on Fairness, Accountability, and Transparency*. 120–128.
[8] Morris H DeGroot. 1974. Reaching a consensus. *Journal of the American Statistical association* 69, 345 (1974), 118–121.
[9] Jwala Dhamala, Tony Sun, Varun Kumar, Satyapriya Krishna, Yada Pruksachatkun, Kai-Wei Chang, and Rahul Gupta. 2021. Bold: Dataset and metrics for measuring biases in open-ended language generation. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 862–872.
[10] Yilun Du, Shuang Li, Antonio Torralba, Joshua B Tenenbaum, and Igor Mordatch. 2023. Improving Factuality and Reasoning in Language Models through Multiagent Debate. *arXiv preprint arXiv:2305.14325* (2023).
[11] Larry G Epstein and Tan Wang. 1996. " Beliefs about beliefs" without probabilities. *Econometrica: Journal of the Econometric Society* (1996), 1343–1373.
[12] Emilio Ferrara. 2023. Should chatgpt be biased? challenges and risks of bias in large language models. *arXiv preprint arXiv:2304.03738* (2023).
[13] Daniel Hart and Suzanne Fegley. 1994. Social imitation and the emergence of a mental model of self. (1994).
[14] Yuan Li, Yixuan Zhang, and Lichao Sun. 2023. Metaagents: Simulating interactions of human behaviors for llm-based task-oriented coordination via collaborative generative agents. *arXiv preprint arXiv:2310.06500* (2023).
[15] Yuan Li, Yixuan Zhang, and Lichao Sun. 2023. MetaAgents: Simulating Interactions of Human Behaviors for LLM-based Task-oriented Coordination via Collaborative Generative Agents. http://arxiv.org/abs/2310.06500 arXiv:2310.06500 [cs].
[16] Tian Liang, Zhiwei He, Wenxiang Jiao, Xing Wang, Yan Wang, Rui Wang, Yujiu Yang, Zhaopeng Tu, and Shuming Shi. 2023. Encouraging Divergent Thinking in Large Language Models through Multi-Agent Debate. *arXiv preprint arXiv:2305.19118* (2023).
[17] Alicia Parrish, Angelica Chen, Nikita Nangia, Vishakh Padmakumar, Jason Phang, Jana Thompson, Phu Mon Htut, and Samuel R Bowman. 2021. BBQ: A hand-built bias benchmark for question answering. *arXiv preprint arXiv:2110.08193* (2021).
[18] Qingyun Wu, Gagan Bansal, Jieyu Zhang, Yiran Wu, Shaokun Zhang, Erkang Zhu, Beibin Li, Li Jiang, Xiaoyun Zhang, and Chi Wang. 2023. Autogen: Enabling next-gen llm applications via multi-agent conversation framework. *arXiv preprint arXiv:2308.08155* (2023).