HONDA

Honda Research Paper Accepted for Presentation at ICLR 2025 Workshop AgenticAI, a Top-tier International Conference in the Field of AI
－ Research Paper on Multi-Agent AI System Inspired by Honda Waigaya Culture

TOKYO, Japan, April 17, 2025 – Honda today announced that a research paper submitted by Honda associates has been accepted for presentation at the ICLR 2025 Workshop AgenticAI, an official workshop of the International Conference on Learning Representations (ICLR). ICLR is a top-tier international conference held annually in the field of machine learning and deep learning, and Honda is scheduled to present its paper at the ICLR 2025, which will be held in Singapore from Thursday, April 24 to Monday, April 28, 2025.

＜Background and Overview of the Research＞

In recent years, the rapid advancement of artificial intelligence (AI), particularly the advancement of large language models (LLMs)[1], has brought about a significant transformation in how products are developed. Especially in the field of automobile development, where the integration of interdisciplinary knowledge is required, there are growing expectations for the use of LLMs, which can generate appropriate responses and insights for a wide range of problems and challenges that need to be addressed.

However, it is difficult for a single LLM-based AI agent[2] to integrate and achieve mutual understanding among diverse and highly specialized knowledge, resulting in limitations in building consensus across different domains. In other words, it is difficult for a single-agent system to replicate the problem-solving process being practiced in real-world business and/or product development settings, where experts in various domains come together, engage in repeated discussions and exchanges of opinions, and work toward problem-solving.

The Honda research paper accepted this time proposes "multi-agent AI systems," where multiple LLM-based agents, each with expertise in different domains, work toward problem-solving through discussions and coordination, modeling after the real-world discussion process of Honda development teams. This study confirmed that multi-agent AI systems outperform single-agent models in the accuracy of generating content related to background information and problem-solving methods, as well as in the stability of the output.

The study also evaluated the effectiveness of each of four different styles of discussion among multiple agents – Decentralized, Centralized, Layered, and Shared Pool. The results confirmed that there was a tendency for diverse opinions to be naturally integrated with the Decentralized style, which was designed based on the process of Waigaya (pronounced Y-guy-a)[3], an element of Honda corporate culture to encourage open and vigorous discussions. The paper was accepted for presentation at the ICLR 2025 Workshop AgenticAI due to the high evaluation for the technical approach of applying the spirit of Waigaya to AI systems.

*1 Large Language Model (LLM) is the foundational technology model for generative AI, that learns from vast amounts of text data to understand and generate natural human language.

*2 AI agent is a type of AI software designed to autonomously plan, execute and improve given tasks by repeatedly inputting and outputting data to Large Language Models (LLMs).

*3 Waigaya is a unique Honda corporate culture that facilitates the generation of original ideas and innovative solutions through open, vigorous and freewheeling discussions.

■ Four different styles of discussion among multiple agents

| 1) Decentralized | 2) Centralized | 3) Layered | 4) Shared Pool |
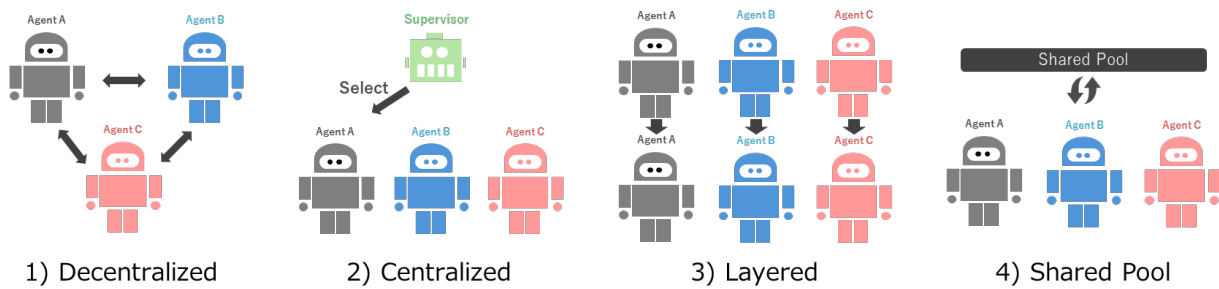
1) Decentralized: This structure enables all agents to exchange opinions freely, promoting the integration of diverse perspectives.

2) Centralized: Due to the centralized control of statements and/or information shared by each agent, the discussion progresses efficiently, but it is difficult to generate evolving dialogues.

3) Layered: The flow of information is organized in layers, and the points of discussion are easily organized in a structured manner.

4) Shared Pool: Since all agents have access to a shared conversation history, consistency of the content of the discussion and stability of the discussion are more easily maintained.

＜Future Initiatives and Prospects＞

In the development of various mobility products, "consensus-based development" plays an important role. In this development approach, in order to meet diverse performance requirements such as driving performance, safety, and comfort, engineers representing different areas of expertise come together to discuss their respective opinions and requirements to work toward the realization of optimal products. The multi-agent AI systems proposed in the paper will enable Honda to replicate/support discussions among development engineers with different areas of expertise through the use of AI agents. This approach – where each AI agent has knowledge in their respective area of expertise and collaborates with other agents specializing in different fields to build consensus – holds the potential to make the traditional consensus-building process more efficient and advanced.

Moving forward, Honda will continue making progress in designing and conducting real-world verification of AI agents that can serve more areas of expertise, and is planning to introduce them in stages within Honda operations which is expected to lead to the realization of faster and more innovative development processes.

Paper title: Dynamic Knowledge Integration in Multi-Agent Systems for Content Inference

Authors:
- First author: Atsushi Yamamoto,
- Corresponding author: Osamu Ito,
- Co-authors: Akihiko Katagiri, Yudai Koike
  All authors mentioned above belong to the Advanced AI Strategy Planning Department within the Digital Unit of Honda Motor Co., Ltd.

URL: https://openreview.net/forum?id=5XNYu4rBe4 (English)