Tutorial on Generalist Agent AI

Ending Remarks

Naoki Wake

Microsoft, 6/18/2024

Thank the speakers and the audience!

Invited Speakers

Juan Carlos Niebles

Yong Jae Lee

Katsushi Ikeuchi

Stanford

WISCONSIN
UNIVERSITY OF WISCONSIN-MADISON

Microsoft

Announcement: the presenters' slides will be available

Timetable Schedule

Time Slot	Talk Scheduling	Talk title	Tutorial Materials
08:30 - 08:40	Jianfeng Gao	Opening Remarks	Slides will be available
08:40 - 09:30	Talk1: Juan Carlos Niebles	Language-based AI Agents and Large Action Models (LAMs)	Slides will be available
09:30 - 09:50	Coffee Break		
09:50 - 10:40	Talk2: Yong Jae Lee	Generalist Multimodal Models	Slides will be available
10:40 - 11:30	Talk3: Katsushi Ikeuchi	Agent Robotics	Slides will be available
11:30 - 11:40	Naoki Wake	Ending Remarks	Slides will be available

https://multimodalagentai.github.io/

Tutorial goal

Generalist Agent AI (GAA) is a family of systems that perform effective actions in an environment based on the understanding of multimodal sensory input

Large language models (LLMs) and large multimodal models (LMMs) have empowered GAA systems, making GAAs applicable in both basic research and practical applications

This tutorial aims to provide an overview of LLM/LMM-empowered GAAs – covering methodologies, evaluation methods, ethical considerations, and future challenges

Generalist Agent AI: Takeaways

	Talk title	Topics
 Juan Carlos Niebles	Language-based AI Agents and Large Action Models (LAMs)	Methodology, Benchmarking, Future development, Open sourcing
 Yong Jae Lee	Generalist Multimodal Models	Methodology, Benchmarking, Open sourcing, Future development
 Katsushi Ikeuchi	Agent Robotics: Learning-from-observation	Methodology, Ethical considerations, Future development

Generalist Agent AI: Takeaways

GAA Applications:

- Large Action Models, Multi-Agent orchestration
- ViP-LLaVA
- Robot Task Planning (Learning-from-Observation)

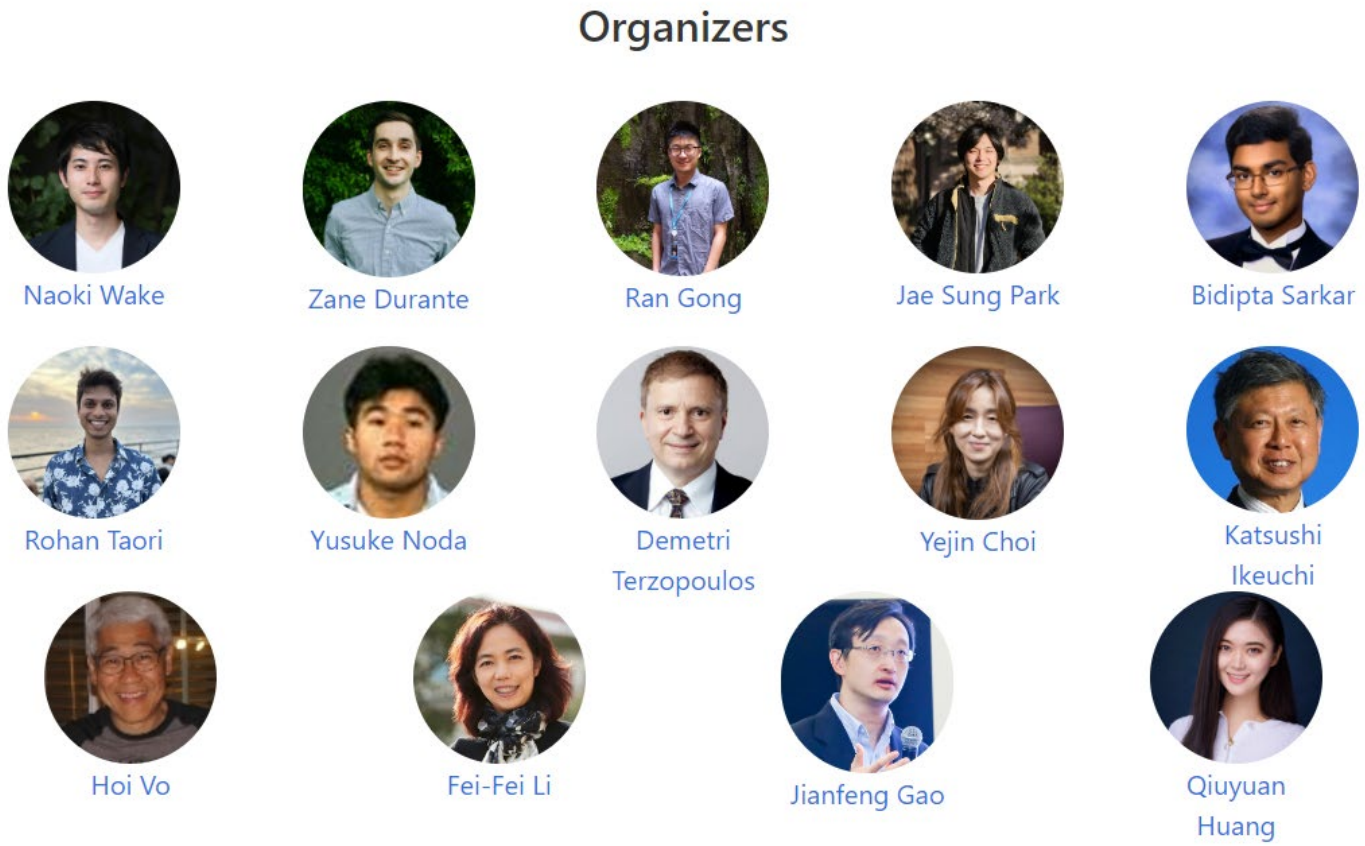
Techniques to Enhance GAAs

- In-context examples, prompting reasoning, scoring result, reflection mechanism
- Instruction tuning, leveraging combinational generalization, personalization
- Incorporation of human-in-the-loop

Continuous development of GAAs

- Open sourcing, continuous development within the community

Organizers



Contact Naoki Wake (naoki.wake@microsoft.com) for any questions

Thank you!