2025 START Program
CFP Brief

THEME: **03. Artificial Intelligence**

SUB-THEME: **3.2. Multimodal Reasoning**

Context/ Overview

Multimodal reasoning is a core pillar for achieving AGI by developing systems that integrate and reason over diverse data modalities such as text, images, audio, video, knowledge graphs among others. This will pave the way for augmenting human creativity and capability by building task-agnostic intelligence ranging from AI radiologists to humanoid robots that can learn to play a new sport by simply reading its rules.

Problem Statement

Language models have recently evolved from quick-response generator to deliberative, thoughtful “System 2” models with OpenAI o series models and the open-source DeepSeek-v3 models attractive a lot of attention. These systems can effectively solve complex cognitive tasks such as coding and deep mathematical research that were difficult for the previous generation of LLMs. We encourage the research community to build upon early success in the language domain and invent new methods for multimodal reasoning. The integration of different modalities with varying complexities, frequency and entropy poses a challenge both in terms of extracting actionable modality-specific insights as well as integrating it with other modalities.

Objectives & Scope

We are particularly interested in system, methods and frameworks that demonstrate generalizability of reasoning abilities beyond the language domain. Furthermore, the capabilities should lead to practical systems that can be deployed with realistic compute, latency and form-factor.

Specific Topics & focus areas*

1. Theoretical and empirical evaluations of deductive, inductive, abductive, analogical reasoning in multimodal systems
2. Training time techniques for improving reasoning
3. Process and outcome reward models (PRM/ORM) and other test-time techniques beyond language modality
4. 3D and physics-enriched models for reasoning

※ The topics are not limited to the above examples and the participants are encouraged to propose other original ideas.

END OF DOCUMENT