

## **Compositional AI - Advanced Abilities Emerge When All Necessary Basic Abilities Are Strong**

**Nan DUAN**

**Microsoft Research Asia, China**

**Email: [nanduan@microsoft.com](mailto:nanduan@microsoft.com)**

Large language models (LLMs) have achieved rapid development recently and demonstrated emergent abilities (e.g., in-context learning and chain-of-thought) that are present in large-scale models only, instead of the small ones. In this talk, the speaker will discuss the possible reason for such emergent abilities: an advanced ability emerges in an LLM when all required components of this ability are sufficiently strong. Some related work [1] also attempted to analyze such novel capabilities from a theoretical perspective and demonstrate that different emergent abilities may be influenced and activated by different latent modules within an LLM. However, since such latent modules are not accessible because of the black-box nature of LLMs, the emergent abilities are not easy to control and train under the current LLM framework. Inspired by this, the speaker proposes Compositional AI as a potential new AI paradigm, which aims to build AI agents that explicitly combine and coordinate multiple AI modules with different skills to tackle complex problems. By putting together multiple modules with specific functions, a compositional AI system can display new abilities that go beyond what any single module can do. The speaker will also introduce two of our recent work, TaskMatrix.AI [2, 3] and ReasonFormer [4], as case studies to demonstrate the advantages of compositional AI and show how they can enable new abilities on completing diverse tasks and solving complex reasoning problems.

### References:

- [1] Hui Jiang. A Latent Space Theory for Emergent Abilities in Large Language Models. arXiv (2023).
- [2] Yaobo Liang, Chenfei Wu, Ting Song, Wenshan Wu, Yan Xia, Yu Liu, Yang Ou, Shuai Lu, Lei Ji, Shaoguang Mao, Yun Wang, Linjun Shou, Ming Gong, Nan Duan. TaskMatrix.AI: Completing Tasks by Connecting Foundation Models with Millions of APIs. arXiv (2023).
- [3] Chenfei Wu, Shengming Yin, Weizhen Qi, Xiaodong Wang, Zecheng Tang, Nan Duan. Visual ChatGPT: Talking, Drawing and Editing with Visual Foundation Models. arXiv (2023).
- [4] Wanjun Zhong, Tingting Ma, Jiahai Wang, Jian Yin, Tiejun Zhao, Chin-Yew Lin, Nan Duan. Disentangling Reasoning Capabilities from Language Models with Compositional Reasoning Transformers. ACL (2023).