2024

Neurosymbolic Customized and Compact CoPilots

Kaushik Roy
University of South Carolina - Columbia

Megha Chakraborty
University of South Carolina - Columbia

Yuxin Zi
University of South Carolina - Columbia

Manas Gaur
University of Maryland Baltimore County

Amit Sheth
University of South Carolina - Columbia

Publication Info

Preprint version 2024.

1 Content, Format, and Program

Title: Neurosymbolic Customized and Compact CoPilots

Abstract: Large Language Models (LLMs) are credible with open-domain interactions such as question answering, summarization, and explanation generation [1]. LLM reasoning is based on parametrized knowledge, and as a consequence, the models often produce absurdities and inconsistencies in outputs (e.g., hallucinations and confirmation biases) [2]. In essence, they are fundamentally hard to control to prevent off-the-rails behaviors, are hard to fine-tune, customize for tailored needs, prompt effectively (due to the "tug-of-war" between external and parametric memory), and extremely resource-hungry due to the enormous size of their extensive parametric configurations [3,4]. Thus, significant challenges arise when these models are required to perform in critical applications in domains such as healthcare and finance, that need better guarantees and in turn, need to support grounding, alignment, and instructibility. AI models for such critical applications should be customizable or tailored as appropriate for supporting user assistance in various tasks, compact to perform in real-world resource-constraint settings, and capable of controlled,robust, reliable, interpretable, and grounded reasoning (grounded in rules, guidelines, and protocols) [5]. This **special session** explores the development of compact, custom neurosymbolic AI models and their use through human-in-the-loop co-pilots for use in critical applications [6].

Motivation: The special session addresses critical needs in AI deployment by enhancing model customization and compactness, crucial for specialized tasks and low-resource environments [7]. This approach not only improves the interpretability and trustworthiness of AI systems, essential in high-stakes sectors like healthcare and finance, but also boosts their robustness and safety. Additionally, neurosymbolic AI enhances the adaptability of systems, allowing for dynamic learning and quick adjustment to new data without extensive retraining. We demonstrate these advancements through a mental health CoPilot deployed in School of Mental Health at University of South Carolina.

Format: The practical special session will be conducted in a lecture format, featuring interactive segments and demonstrations of the tool, totaling 100 minutes in duration.

Schedule:
Neurosymbolic AI and CoPilots (20 mins): Neurosymbolic AI represents a fusion of neural networks and symbolic AI, leveraging the strengths to create more robust and adaptable systems. Neural networks excel in pattern recognition and learning from large datasets, while symbolic AI offers explicit reasoning and manipulation of abstract concepts. This combination can be particularly powerful for developing CoPilots—AI assistants designed to support users in various tasks [8,9]. In the context of CoPilots, Neurosymbolic AI enables these assistants to understand and execute complex instructions and perform constrainable, interpretable, and scalable reasoning by leveraging knowledge graphs integrated with neural networks [10]. This integration enhances the capability of CoPilots,

particularly in making them accurate, safe, consistent, reliable, and interpretable [7]. **Learning Outcomes** Attendees will learn about the background of Neursoymbolic AI systems and their potential for developing CoPilots.

Customization of Copilots with Neurosymbolic AI (20 mins): Customization of CoPilots involves training them on curated data and specific knowledge sources, ensuring the AI aligns with particular needs and contexts [11]. Neurosymbolic AI aids this customization by enabling mechanisms for integrating domain-specific knowledge into the learning process [12]. This process ensures the CoPilots are safe and reliable, as they are trained on verified and relevant information, minimizing the risk of erroneous outputs. Customization also allows for incorporating ethical guidelines and industry-specific regulations, i.e., inculcating value systems and structures similar to that of the end-users, further enhancing the safety and appropriateness of the responses generated by CoPilots [13]. **Learning Outcomes:** Through various examples, attendees will learn about the potential of incorporating curated data and knowledge for training language models using neurosymbolic mechanisms, specifically toward minimizing adverse language model outputs.

Compact Models for Copilots using Neurosymbolic AI (15 mins): Implementing neurosymbolic CoPilots can be quite complex, resulting in large systems in terms of resource consumption during both training and inference. Developing compact models for CoPilots is crucial for deployment in resource-constrained environments, such as mobile devices or embedded systems, a key requirement in critical application scenarios such as healthcare. Leveraging reasoning and learning over heterogeneous knowledge graphs at varying levels of abstraction in the neurosymbolic setup contributes to this compactness by reducing dependence on data and, consequently, the need for extensive computational resources [14,15]. Furthermore, complementary techniques for facilitating the compactness of the neural part of the CoPilots, such as model re-design (e.g., through reservoir computing and state space models), hardware acceleration (e.g., SRAM optimization), and memory optimization (e.g., quantization), can also be applied [16,17,18]. These features will result in CoPilots that maintain high levels of accuracy and functionality by balancing efficiency and capability, which is vital for the widespread adoption of CoPilots in various practical applications.**Learning Outcomes** Attendees will learn about neurosymbolic design choices that facilitate the construction of compact systems enabled by leveraging the robust and generalizable knowledge-guided reasoning of knowledge-infused AI and the adaptability, scalability, and optimization of neural networks.

Orchestration of Customized and Compact Copilots (15 mins): Orchestration involves multiple copilots that can work together in a coordinated manner to ensure seamless operation towards user-defined goals and objectives. Compactness and customization become crucial in this effort due to the need for high-quality performance while ensuring the efficient allocation and utilization of computational resources to balance the workload among different copilots. Successful orchestration ultimately results in an enhanced user experience, specifically efficient, relevant, and context-aware assistance, improving user satisfaction. **Learning Outcomes** Attendees will learn how Neurosymbolic Planning frameworks can effectively orchestrate Neurosymbolic CoPilots by working towards overall user goals by iterative progress via targetted subgoals.

Mental Health Use Case - A Demonstration (20 mins): We use mental health to exemplify a critical application. The motivation for applying CoPilots in the mental health domain stems from the increasing need for accessible and personalized mental health support [19]. With rising mental health concerns globally, there is a significant demand for tools that can provide timely assistance and support to individuals. The objectives of utilizing CoPilots in mental health include offering immediate, personalized interventions, facilitating continuous monitoring and support, and providing resources and guidance tailored to individual needs [20]. By customizing CoPilots with mental health expertise and ensuring they are compact enough to run on commonly available devices, we can enhance their accessibility and usability [21]. These CoPilots can offer various forms of support, from initial assessments to ongoing therapeutic interactions. CoPilots can even deliver cognitive behavioral therapy (CBT) techniques, mindfulness exercises, and other therapeutic interventions tailored to the user's needs [22]. A successful implementation that can operate in real-time and incorporate user feedback presents great potential for assistance in managing mental health conditions [23]. Moreover, the customization and compactness of these models ensure they are adaptable for operation in diverse and nuanced environments and contexts. This adaptability is crucial for addressing the varied and dynamic nature of mental health needs. **Learning Outcomes** Through a mental health support system that demonstrates personalized care and infrastructure-level support, attendees will learn about the operationalization of Neurosymbolic AI CoPilots within an orchestration framework.

2 Tutorial type and intended audience

Type: Specialised or advanced tutorial with domain-specific application and tool.

Level: Intermediate

Target audience: This special session aims to gather scholars from academia, industry, humanitarian organizations, and healthcare professionals interested in the intersection of neurosymbolic AI, knowledge graphs, reasoning, and knowledge-infused LLMs for edge devices. We expect 30-40 attendees as there has been a growing interest in topic, especially after AI Index Annual Report [24] and development of small models for resource-constrained settings [18].

Prior knowledge: The audience is required to have basic understanding of LLMs. This prerequisite ensures that the audience is adequately prepared to appreciate the content and insights shared during the special session.

3 Presenters information:

Kaushik Roy, is a Ph.D. student at the AI Institute, University of South Carolina (AIISC). His research encompasses statistical relational artificial intelligence, sequential decision-making, knowledge graphs, and reinforcement learning, with publications in esteemed conferences such as IEEE, KR, ECML-PKDD, and AAAI. He has delivered 5+ tutorials in semantic web conferences (e.g., KGSWC, KGC). He can reached at `kaushikr@email.sc.edu` and for more information, please check his Webpage.

Megha Chakraborty is a Ph.D. student at AIISC. Her research encompases LLMs and Mental Health. She has given multiple tutorials and presentations in summer school and technical symposiums organized at USC. Her research has been accepted at EMNLP and ACL. She can reached at `MEGHAC@email.sc.edu` and for more information, please check her Webpage.

Yuxin Zi is a Ph.D. student at AIISC. Her research encompases neurosymbolic AI, LLMs, grounded and instructable LLMs, and Mental Health. Her research has been accepted by AAAI, IEEE, and Sensors. She can reached at `YZI@email.sc.edu` and for more information, please check her Webpage.

Manas Gaur, an Assistant Professor at University of Maryland Baltimore County (UMBC). His interests lie in Neurosymbolic AI, Knowledge graphs, Knowledge-infused Learning, Grounding and Instructable LLMs, Conversational Systems, and Mental Health, with publications in Web Conference, ISWC, AAAI, ECIR, CIKM, IEEE Semantic Computing, LREC-COLING, and others. He has delivered 20+ tutorials in major conferences in ACM and IEEE. She can reached at `manas@umbc.edu` and for more information, please check his Webpage.

Amit Sheth: is the founding director of AIISC and NCR Chair at USC. He is a Fellow of IEEE, AAAI, ACM, and AAAS. He has organized >100 international events (general/program chair, organization committee chair), >70 keynotes, given >45 many well-attended tutorials and is among the well-cited computer scientists. He has founded three companies including the first Semantic Web company in 1999 that pioneered technology similar to what is found today in Google Semantic Search and Knowledge Graph. He can reached at `amit@sc.edu` and for more information, please check his Webpage.

4 Materials

This session provides comprehensive, self-contained material that is easily accessible through the following link: Tutorial Neurosymbolic Customized and Compact-CoPilots. Additionally, we will establish a page akin to `https://aiisc.ai/neurone/` to pre-share the session's content. For materials from our previous tutorials on similar topics, please refer to the Keynotes, Talks, and Tutorials section at https://wiki.aiisc.ai.

5 Information about previous editions

Several NeuroSymbolic AI and LLMs tutorials have been given at related conferences. For example, *Proactive Conversational Agents in the Post-ChatGPT World* [25], *Neuro-Symbolic Representations for Information Retrieval* [26] in SIGIR'23 and *Some Useful Things to Know When Combining IR and NLP: the Easy, the Hard and the Ugly* in CIKM '23 [27]. However, these tutorials have not explicitly focused on grounding, instructability, and alignment, not as independent aspects, but as a single braid for LLM, achievable through neurosymbolic design. This session arrives at a crucial moment, responding to the pressing demand for enhancing conversational agents (chatbots to copilots) in various practical settings. Its subject matter is poised to captivate scholars and professionals active in Semantic Web, Knowledge Representation and Reasoning, LLMs, and Trusworthy AI.

References

1. Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al. A survey of large language models. *arXiv preprint arXiv:2303.18223*, 2023.
2. Vipula Rawte, Amit Sheth, and Amitava Das. A survey of hallucination in large foundation models. *arXiv preprint arXiv:2309.05922*, 2023.
3. Kevin Wu, Eric Wu, and James Zou. How faithful are rag models? quantifying the tug-of-war between rag and llms' internal prior. *arXiv preprint arXiv:2404.10198*, 2024.
4. Zhuoran Jin, Pengfei Cao, Yubo Chen, Kang Liu, Xiaojian Jiang, Jiexin Xu, Qiuxia Li, and Jun Zhao. Tug-of-war between knowledge: Exploring and resolving knowledge conflicts in retrieval-augmented language models. *arXiv preprint arXiv:2402.14409*, 2024.
5. Amit Sheth, Manas Gaur, Kaushik Roy, Revathy Venkataraman, and Vedant Khandelwal. Process knowledge-infused ai: Toward user-level explainability, interpretability, and safety. *IEEE Internet Computing*, 2022.
6. Amit Sheth, Kaushik Roy, and Manas Gaur. Neurosymbolic artificial intelligence (why, what, and how). *IEEE Intelligent Systems*, 38(3):56–62, 2023.
7. Manas Gaur and Amit Sheth. Building trustworthy neurosymbolic ai systems: Consistency, reliability, explainability, and safety. *AI Magazine*, 2024.
8. Qingyun Wu, Gagan Bansal, Jieyu Zhang, Yiran Wu, Shaokun Zhang, Erkang Zhu, Beibin Li, Li Jiang, Xiaoyun Zhang, and Chi Wang. Autogen: Enabling next-gen llm applications via multi-agent conversation framework. *arXiv preprint arXiv:2308.08155*, 2023.
9. Robert Feldt, Sungmin Kang, Juyeon Yoon, and Shin Yoo. Towards autonomous testing agents via conversational large language models. In *2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 1688–1693. IEEE, 2023.
10. Nova Spivack, Sam Douglas, Michelle Crames, and Tim Connors. Cognition is all you need-the next layer of ai above large language models. *arXiv preprint arXiv:2403.02164*, 2024.
11. Tim Valicenti, Justice Vidal, and Ritik Patnaik. Mini-gpts: Efficient large language models through contextual pruning. *arXiv preprint arXiv:2312.12682*, 2023.
12. Yu Gu, Robert Tinn, Hao Cheng, Michael Lucas, Naoto Usuyama, Xiaodong Liu, Tristan Naumann, Jianfeng Gao, and Hoifung Poon. Domain-specific language model pretraining for biomedical natural language processing. *ACM Transactions on Computing for Healthcare (HEALTH)*, 3(1):1–23, 2021.
13. Amit Sheth and Kaushik Roy. Neurosymbolic value-inspired artificial intelligence (why, what, and how). *IEEE Intelligent Systems*, 39(1):5–11, 2024.
14. Amit Sheth, Manas Gaur, Kaushik Roy, et al. Knowledge-intensive language understanding for explainable ai. *IEEE Internet Computing*, 2021.
15. Amit Sheth and Krishnaprasad Thirunarayan. the duality of data and knowledge across the three waves of ai. *IT professional*, 23(3):35–45, 2021.
16. Wei Miao, Vignesh Narayanan, and Jr-Shin Li. Interpretable design of reservoir computing networks using realization theory. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.
17. Albert Gu and Tri Dao. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*, 2023.

18. Edward J Hu, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, et al. Lora: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2021.
19. Tracy Butryn, Leah Bryant, Christine Marchionni, and Farhad Sholevar. The shortage of psychiatrists and other mental health providers: causes, current state, and potential solutions. *International Journal of Academic Medicine*, 3(1):5–9, 2017.
20. Amelie Gyrard, Manas Gaur, Saeedeh Shekarpour, Krishnaprasad Thirunarayan, and Amit Sheth. Personalized health knowledge graph. In *CEUR workshop proceedings*, volume 2317. NIH Public Access, 2018.
21. Kaushik Roy, Yuxin Zi, Manas Gaur, Jinendra Malekar, Qi Zhang, Vignesh Narayanan, and Amit Sheth. Process knowledge-infused learning for clinician-friendly explanations. In *AAAI Spring Symposium Series*, 2023.
22. Kaushik Roy, Manas Gaur, Misagh Soltani, Vipula Rawte, Ashwin Kalyan, and Amit Sheth. Proknow: Process knowledge for safety constrained and explainable question generation for mental health diagnostic assistance. *Frontiers in big Data*, 5:1056728, 2023.
23. Kaushik Roy, Vedant Khandelwal, Raxit Goswami, Nathan Dolbir, Jinendra Malekar, and Amit Sheth. Demo alleviate: Demonstrating artificial intelligence enabled virtual assistance for telehealth: The mental health case. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 16479–16481, 2023.
24. Human Centered Artificial Intelligence. 2024 ai index report, April 2024. Accessed: 2024-04-27.
25. Lizi Liao, Grace Hui Yang, and Chirag Shah. Proactive conversational agents in the post-chatgpt world. In *SIGIR*, 2023.
26. Laura Dietz, Hannah Bast, Shubham Chatterjee, Jeffrey Dalton, Jian-Yun Nie, and Rodrigo Nogueira. Neuro-symbolic representations for information retrieval. In *SIGIR*, 2023.
27. Omar Alonso and Kenneth Church. Some useful things to know when combining ir and nlp: The easy, the hard and the ugly. In *WSDM*, 2024.