Enoch Levandovsky

Department of Computer Science
Boise State University
Boise, Idaho, USA

enoch.levandovsky@u.boisestate.edu
https://enochlev.com

1 Research interests

My main research interests lies in the application of Reinforcement Learning (RL) alignment of LLMs in human robot dialogue. More specifically, my latest research aims to use RL alignment as an efficient training regime to train a newly initialized tiny LM to behave like a toddler. Previous research expresses the difficulty of building a robust tiny LM with an educated adult level understanding. Our hypothesis is that the cognitive barrier to train a tiny LM to at-least behave as a child is achievable with a very small number of parameters especially if training efficiently using RL LLM training regime. My interests also extend to apply RL to LLM training for dialogue management and planning.

1.1 Training Language Models using RL

Like most core concepts of computer science, research takes real world concepts and applies them to computer science. The application Reinforcement Learning to Language Model (LM) training regimes is an excellent example of that application. Initial RL and LLM training research shows how RL can be used to make LM training explainable and more efficient Rafailov et al. (2024). RL has especially gained traction from DeepSeek-AI et al. (2025) where the researchers have demonstrated splitting LLM training goals into modular tasks that are computed by a well defined reward models or functions, then alignment is achievable and observable. This has opened up a new sphere of research that if we are able to create and modularize stTrong reward functions, then we can also efficiently align an LLM to the relative tasks.

1.2 Using RL for Dialogue Management and Planning

As noted in the DeepSeek paper, the robustness of the Reward Models chosen to align a LM is vital to the success of the training run via GRPO. The reward model must carefully created to correctly score and provide guidance during training in order for it to reach its goals. This is especially difficult if a BERT reward model is chosen due to its black box nature of deep neural networks. This is the driving reason DeepSeek chose Mathematics and Syntax because the tasks are easily structured, explainable and assessable. I believe dialogue management also falls within the structured domain, and if done right I believe RL can be successfully applied to LM training of a dialogue management task.

1.3 Training Language Models with child behavior

Another of my research area, aims is to further extend the application of RL to an SDS pipeline in a child-robot human interaction setting. I believe traditional training regimes are not optimized for LMs with childish behavior. For example, the data requirements to train an LM as you reduce the number of parameters decreases according to SmollLM2, a pretraining takes 2 trillion tokens for a 135M parameter model Allal et al. (2025). This falls short of the 10 million tokens available for even large datasets such as Chides parent child interactions MacWhinney (2025). Although, this data is enough to finetune a pretrained model, but interestingly the finetuning task reduces cognitive understanding of the model making training unstable. For example, after a successful finetune of a model to conversationally behave like a toddler, we were surprised to be informed by the model that it knows more than 10 languages. This is why I am convinced LMs designed to behave as a child must be trained from initialized weights to develop language understanding as a toddler would despite the limited availability of pretraining tokens. Thus my research interest lies in using RL to efficiently conduct training.

1.4 Incremental SDS Processing

My master's research has primarily revolved around the incrementalization of an SDS pipeline. Typically, an SDS processes input only after detecting an end-of-utterance from the human speaker, which introduces latency and limits the system's responsiveness in dynamic interactions. By contrast, an incremental SDS processes audio and text as it unfolds, allowing the system to react to partial utterances in real time. Each new segment of incoming audio is immediately processed and integrated, thereby enabling more fluid and timely feedback. This shift toward word or Incremental Unit (IU) level processing can make human-robot interactions to be more natural. As of now the ASR and the NLU have been successfully incrementalized, but there is yet more research to be done to make make the Dialogue Manger incremental.

The 21th Annual Meeting of the Young Researchers' Roundtable on Spoken Dialogue Systems, pages 4–5
Aug 28–29, 2025.

Its success is a vital step to complete a fully incremental SDS pipeline.

2 Spoken dialogue system (SDS) research

I think the progress of SDS has dramatically been driven forward since the rise of LLMs and has lots of potential of where it can go. For the most part this is because we have alot of text on the internet in which we can process easily unlike other domains. However, I believe the amount of speech and visual data is also within the domain.

As for the future of research, firstly, I think there will be significant progress in multi-model LLMs allowing SDS to encompass visual and speech integration much more fluently. Secondly, I think speech and audio will be embedded into a transformer model allows for LLMs to natively conduct STT and emotion recognition taking a step forward in LLMs behaving more closely to humans.

Industrial research is a lot more financial and goal driven, but only selects research on areas where short-term ROI can be relatively guaranteed such as an agentic airline agent. Where as academia research is rather driven by the interest of younger thinkers. Thus the future of SDS has a lot to do with how the younger generation is introduced to SDS and how they envisions its future. I think students envision mimicking a broader human to human interactions and environmental interaction rather then simpler agentic airline bots. Thus I think SDS research will be financially driven in areas more where the industry can see a short-term ROI in areas such as stronger LLMs, STT, TTS, and vision models.

In this section, the authors will briefly present their opinions and point of views of current and future SDSs research in general. These could include answers to questions like:

3 Suggested topics for discussion

Here, authors will suggest three topics for discussion in the discussion panels during the event. As an example, here are some of the discussion topics discussed in previous workshops:

- Statistical methods and user simulations: Best practices for building and evaluation user simulations, challenges in using reinforcement learning for building systems.
- Speech driven web browsing, in-car systems, mobile devices, health-care, interface to robots.
- Affect and other paralinguistic phenomena: Emotion detection and handling, turn taking, response timing.

References

Loubna Ben Allal, Anton Lozhkov, and Elie Bakouch. 2025. SmolLM2: When smol goes big – data-centric training of a small language model. https://doi.org/10.48550/arXiv.2502.02737.

DeepSeek-AI, Daya Guo, Dejian Yang, and Haowei Zhang. 2025. DeepSeek-r1: Incentivizing reasoning capability in LLMs via reinforcement learning. https://doi.org/10.48550/arXiv.2501.12948.

Dale Hamilton, Enoch Levandovsky, and Nicholas Hamilton. 2020. Mapping burn extent of large wildland fires from satellite imagery using machine learning trained from localized hyperspatial imagery 12(24). https://doi.org/10.3390/rs12244097.

Brian MacWhinney. 2025. CHILDES. https://talkbank.org/childes/.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Stefano Ermon, Christopher D. Manning, and Chelsea Finn. 2024. Direct preference optimization: Your language model is secretly a reward model. https://doi.org/10.48550/arXiv.2305.18290.

Ryan Whetten, Enoch Levandovsky, Mir Tahsin Imtiaz, and Casey Kennington. 2023. Evaluating automatic speech recognition and natural language understanding in an incremental setting. https://www.semdial.org/anthology/papers/Z/Z23/Z23-3012/.

Biographical sketch

Enoch Levandovsky is a 2nd year Ph.D. student as well as AI Data Scientist in the industry. He is currently studying at Boise State University under Ph.D. Casey Kennington. He demonstrates a strong skills in fien-tuning and deployment of Deep Learning models. He has a passion for applied AI and actively seeks for ways to make human and AI interaction more natural. His personal hobbies include routine and home automation using LLMs. He has publishings at the proceedings from Whetten et al. (2023) and Hamilton et al. (2020).