Research Statement: Natural Language Guided Reasoning

Niket Tandon

June 2023

I am broadly interested in the computational foundations of never-ending learning through natural language guidance. The overarching theme of my research is to develop algorithms that **make NLP systems more explainable, interactive and teachable over time**. Output from current models is generally of a good form and syntax but can be senseless in places. End users can spot these commonsense and consistency problems or identify when their instructions and expectations are not met, and expect an intelligent behavior that follows their guidance and improves over time, but large models are prohibitive to finetune or to improve after deployment. On the other hand, *theory of recursive reminding* from Psychology suggests that humans record error context and correction received in their episodic memory to learn from their mistakes and avoid repeating them for the future. Can machines behave in this way?

Inspired by the theory of recursive reminding, our goal is to "effectively provide natural language guidance for improved reasoning, and understanding." as a step towards never ending learning. Such an intelligent behavior is endowed with the following key abilities, and this summarizes my career's research contributions:

- who to seek guidance from (e.g., humans, knowledge bases, with self-reflection, or, other supervised agents)
- when to seek guidance (e.g., during inconsistency, or when a similar problem was encountered in the past),
- what to seek guidance on (e.g., on the model's explanations or justification, or on a structured output), and
- how to apply guidance (e.g., algorithms to change model behavior by changing the input, weights, or output).

This guidance can take up different forms such as feedback, corrections, preferences, knowledge, self-reflection, or guidance by established theories e.g. from psychology. The overall design space is presented in this figure:

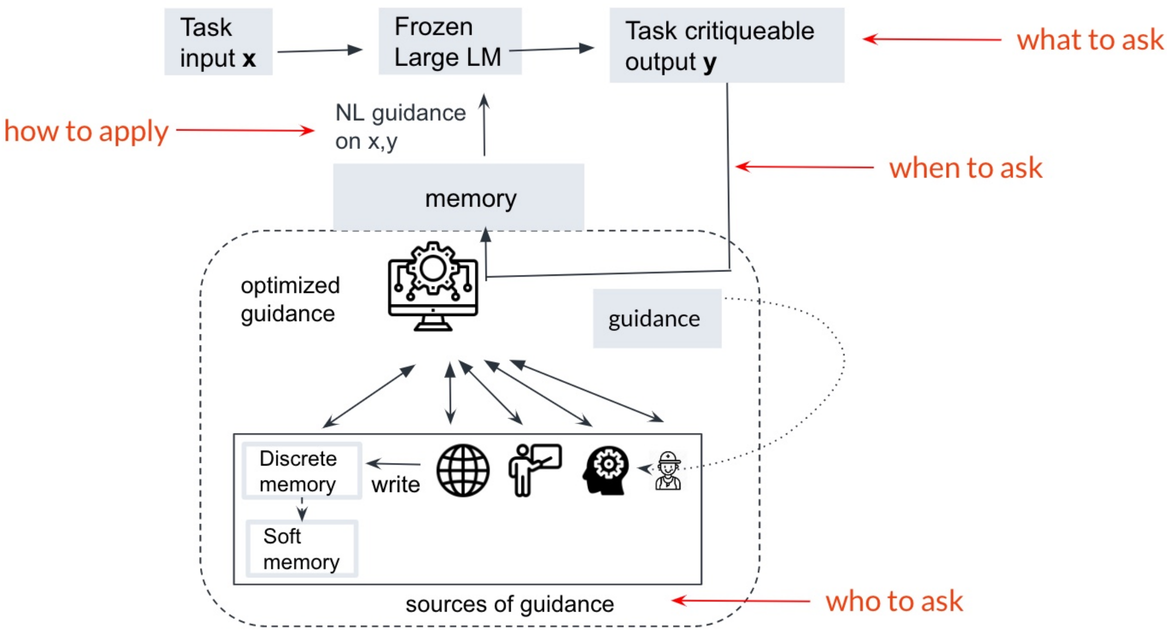

Figure 1: Overall design space of my research.

Memory based architecture

One instantiation of this general design space is our recent work, MemPrompt [7] which is a memory based architecture where input = question, output = critiqueable output and the model receives guidance from various agents, potentially iteratively. To avoid repeating similar mistakes, mistakes with context is recorded in a memory (analogous to a human's episodic memory). MemPrompt is making semi-parametric methods a feasible and effective approach to engage non-expert end users in meaningful ways (this work was also covered and interview in an interview on a popular machine learning YouTube channel). Going forward, another impact of MemPrompt could be when this memory is per user, as it can lead to personalized large models which can be taught to behave consistently with a user's belief system or language dialects (especially helpful in personalizing models for a diverse cultural setting like India).

Who to seek guidance from?

In this memory architecture, to answer the first research question "who to seek guidance from". Table 1 list various genres of guidance. Knowledge is the most obvious genre of guidance such as from KBs. Much of my Ph.D. dissertation was in pre-constructing the largest KB of commonsense at the time, WebChild [1] and a series of papers we showed that automatic compilation of a large-scale commonsense KB is possible from small chunks of text (N-grams), movie-scripts, and other sources. The impact of this work has been that it inspired more automatically compiled KBs since, and has been applied in applications in NLP, and computer vision. We also showed that KGs constructed on-the-fly can be quite effective in reasoning applications [4, 5]. Indeed, a memory of user feedback containing commonsense statements in our recent work [6], can be seen as a KB of commonsense generated on the fly. We are working on such knowledge being gathered in a memory via interactions in an environment (e.g., in ScienceWorld or AlfWorld and the memory contains preconditions). This in-progress work will help us with an automatic compilation of knowledge by exploration.

Genre of guidance	Guidance from	Most representative paper
Knowledge as guidance	Pre-constructed KG	WebChild-KB [1]
	On-the-fly KG	Inference graphs [5]
Feedback as guidance	Human feedback	Interscript [11]
	Supervised feedback	Learning-to-repair [6]
	RL feedback	RL4F [9]
	Self feedback	Self-Refine [8]
Schemas as guidance	Dyadic theory of harm	Moral bottleneck models [in progress]
	State tracking schema	[EMNLP 2023 submission]
Preference as guidance	User preference	[EMNLP 2023 submission]
	User opinions	[EMNLP 2023 submission]
Environment as guidance	Simulated environment	Planning in simulated env [in progress]

Table 1: Different sources of guidance studied in my research. Pre-prints available (href) for in-progress works.

The other form of guidance is feedback. In [11], we effectively get guidance from humans, and in [5, 6] we trained a supervised model on human feedback. In [9], we train a Reinforcement learning model to generate feedback that is optimized towards end task or a custom reward and this setup were more effective than even a supervised setup. Very recently, in [8] we show for the first time that it is possible for a large language model to generate very effective feedback itself. This work received much attention and was the most popular arXiv paper for a few days when it was published. There are several open problems and some that we are currently engaged in making a truly self-refine models that adapt by introspection, and in designing a multi-aspect reward for RL agents for detailed introspection. The impact of this line of research has been that feedback is no longer seen as coming from a single-agent source but in recent papers, starting to be seen in a multi-agent setup.

More recently, we have worked on user preferences and user opinions as a source of guidance. We showed that language models can model population level statistics better than demographic level and still struggle with personal opinions and preferences. We are working on an interpretable model for user level personalization that accounts for past opinions from the user to explain their persona, and these past opinions are obtained by interaction with a human-in-the-loop and are stored in a memory. We also finding in our recent work that it is possible to rely on inter-disciplinary theories e.g., from psychology that suggest the mental picture a human undergoes when dealing with moral queries, and we find that when models are guided through such theories that improves performance and makes them more faithful. Finally, in a work in progress for the planning task, we are studying if a simulated environment can be an effective source of feedback, and how to store these interactions in a memory which can distill reflections from these interactions. The impact of these different lines of research is that we are looking at new ways of guidance that is now enabled by advancements in models and simulations and due to an interdisciplinary collaboration established with a professor of psychology.

What, when to ask and how to apply guidance?

What should the model seek guidance on? In our work, we have found that primarily explanations and structured output and to some extent unstructured output is to guidance. Towards this, I have worked on several papers that generate effective explanations in language models e.g., [3]. Another important question our research has tried to address is whether the output requires a critique? As calibrating language models is generally hard, we instead develop some automatic measure to detect inconsistency in the explanation structure of a model. The final question is how to update the model behavior by using the guidance (see Table 2)? We have effectively done at the input (enhancing the context [7]), at the output (with a supervised corrector [2, 6]) or update the model parameters (for relatively smaller models or when models weights are accessible [12]). We show in [12] that is possible for an unfrozen model to take binary feedback from a human and update model parameters via causal tracing (thus performing model editing which can work instantly even for one instance rather than traditional finetuning). An interesting open question is whether natural language user feedback can be used to effectively update a model's parameters by model editing.

Where to apply	How to apply	Most representative paper
at the input	Input context	MemPrompt [7]
at the output	Decoder, corrector	Learning-to-repair [6]
at the parameters	Loss function	ProStruct [2]
	Causal tracing	[EMNLP 2023 submission]

Table 2: How to apply the obtained guidance, that we studied. Pre-prints available (href) for in-progress works.

More future work

LLMs have several important aspects with open research questions that we abbreviate as "The CRITERIUM". LLMs must be controllable, responsible, interactive, trustworthy, environmental friendly, reasonable, be able to induce facts, and work in a multimodal+multilingual setting. The proposed topic of natural language guided reasoning broadly covers two of these aspects *interactive* systems that learn from errors and are teachable and personalized; and *reasonable* systems that exhibit for instance commonsense and moral reasoning. We also strive for the models to be trustworthy and faithful in their reasoning. Within each aspect of the CRITERIUM, there are several open research questions that we will continue to work on, with a special focus on interactive, reasoning (commonsense and moral reasoning), trustworthy AI and controllable models.

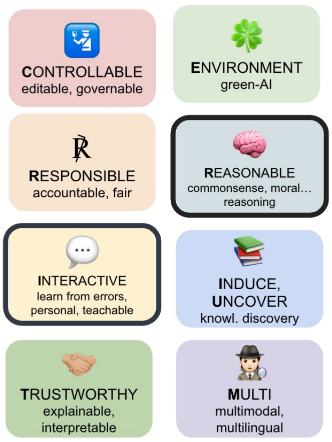

Figure 2: My research in the bigger landscape of NLP (highlighted with a boundary)

Selected publications

1. Niket, Gerard, Fabian, Gerhard Weikum. "WebChild: harvesting and organizing commonsense knowledge from the web." WSDM 2014

2. Niket, Bhavana, et. al. "Reasoning about Actions and State Changes by Injecting Commonsense Knowledge." EMNLP 2018

3. Dheeraj, Niket, Peter Clark et. al. "What-if I ask you to explain: Explaining the effects of perturbations in procedural text." EMNLP Findings 2020

4. Aman, Niket, Dheeraj, et. al. "Think about it! Improving defeasible reasoning by first modeling the question scenario." EMNLP 2021

5. Aman, Dheeraj, Niket et. al. “Could you give me a hint? Generating inference graphs for defeasible reasoning.” EMNLP Findings 2021

6. Niket, Aman, Peter, Yiming. “Learning to repair: Repairing model output errors after deployment using a dynamic memory of feedback.” NAACL 2022.

7. Aman, Niket, Peter, Yiming. “Memory-assisted prompt editing to improve GPT-3 after deployment.” EMNLP 2022.

8. Aman, Niket, et. al. “Self-Refine: Iterative Refinement with Self-Feedback.” ArXiv abs/2303.17651 (2023)

9. Feyza, Ekin, Ashwin, Peter, Derry, Niket. “RL4F: Generating Natural Language Feedback with Reinforcement Learning for Repairing Model Outputs.” ACL 2023

10. Yash Kumar Lal, Niket, Tanvi Aggarwal, Horace Liu, Nathanael Chambers, Raymond Mooney, and Niranjan Balasubramanian. “Using Commonsense Knowledge to Answer Why-Questions.” EMNLP 2022

11. Niket, et al. “Interscript: A dataset for interactive learning of scripts through error feedback.” ACL 2021 workshop

12. Gupta A, Mondal D, Sheshadri AK, Zhao W, Li XL, Wiegreffe S, Tandon N. “Editing Commonsense Knowledge in GPT”. arXiv preprint arXiv:2305.14956. 2023 May 24.