A Bird Did Not Sit in Aerodynamics Classes: Structural Priors for Non-Parametric Personal Memory in Language Models

General world knowledge and personal knowledge in large language models are represented together on the same parameteric base, so persistent personalization would be fundamentally different from the editing of an isolated memory: adding one's personal fact could require changing the same representations used to store the general world knowledge. There is a more profound architectural confusion of two different learning tasks: the learning of the concrete content and its representation. The developmental cognitive science provides a nice conceptual difference, where the prior knowledge about the structure allows learning something, but does not prescribe the content that is learned.

This architecture features the use of a static pretrained language model in conjunction with a non-parametric, typed relational memory. Extracted information from the language model is regarded as a candidate instead of directly storing it as fact. Whether or not the extracted information qualifies as a fact depends on the semantic categories and constraints associated with the relations. The language model cannot extract any ambiguous or structurally inconsistent information. Once the information passes through this stage, it is then converted into a relation and stored as a valid graph relation that may later undergo updates through memory operations without changing the language model.

Personal-fact question-answering evaluation of the architecture involves comparing it against retrieval-augmented and parameter-efficient fine-tuning paradigms based on factual error, hallucination rate, update speed, stability of interactions repeated multiple times, and resilience to contradictions or ambiguity. The study considers both strengths and weaknesses of fixed structural constraints, including situations when predefined categories are not enough to represent the context-based facts. The evidence suggests a more specific version of the hypothesis, namely, separating the structure of experience and language model expressing it could be beneficial to make persistent personal knowledge more localized, inspectable, and updatable, transferring the open problem to learning structural categories.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.