

GEPA: REFLECTIVE PROMPT EVOLUTION CAN OUTPERFORM REINFORCEMENT LEARNING

Multi-TAG: 多工具ensemble的prompt method for TIR

Lakshya A Agrawal¹, Shangyin Tan¹, Dilara Soylu², Noah Ziemis⁴,
Rishi Khare¹, Krista Opsahl-Ong⁵, Arnav Singhvi^{2,5}, Herumb Shandilya²,
Michael J Ryan², Meng Jiang⁴, Christopher Potts², Koushik Sen¹,
Alexandros G. Dimakis^{1,3}, Ion Stoica¹, Dan Klein¹, Matei Zaharia^{1,5}, Omar Khattab⁶

¹UC Berkeley ²Stanford University ³BespokeLabs.ai ⁴Notre Dame ⁵Databricks ⁶MIT

未开源

简介

本文提出GEPA(Genetic-Pareto)，一种基于演化算法(evolutionary)和反思(reflective)的prompt优化框架，目的是通过优化prompt而非tuning llm来提升compound AI system的整体效果。我们从演化算法三要素(变异、评估、选择)角度来看下GEPA: 1) 变异阶段，通过llm结合prompt、 reasoning trace以及answer对reasoning trace进行反思，生成改进后的prompt；2)评估阶段，prompt会在小批量数据上执行得到answer以及效果，以比较不同prompt的表现；3)选择阶段，GEPA 引入基于单条数据级的Pareto最优的候选筛选机制，简单来说就是确保保留那些在至少一条数据上取得领先表现的prompt，从而兼顾exploration和exploitation。

背景

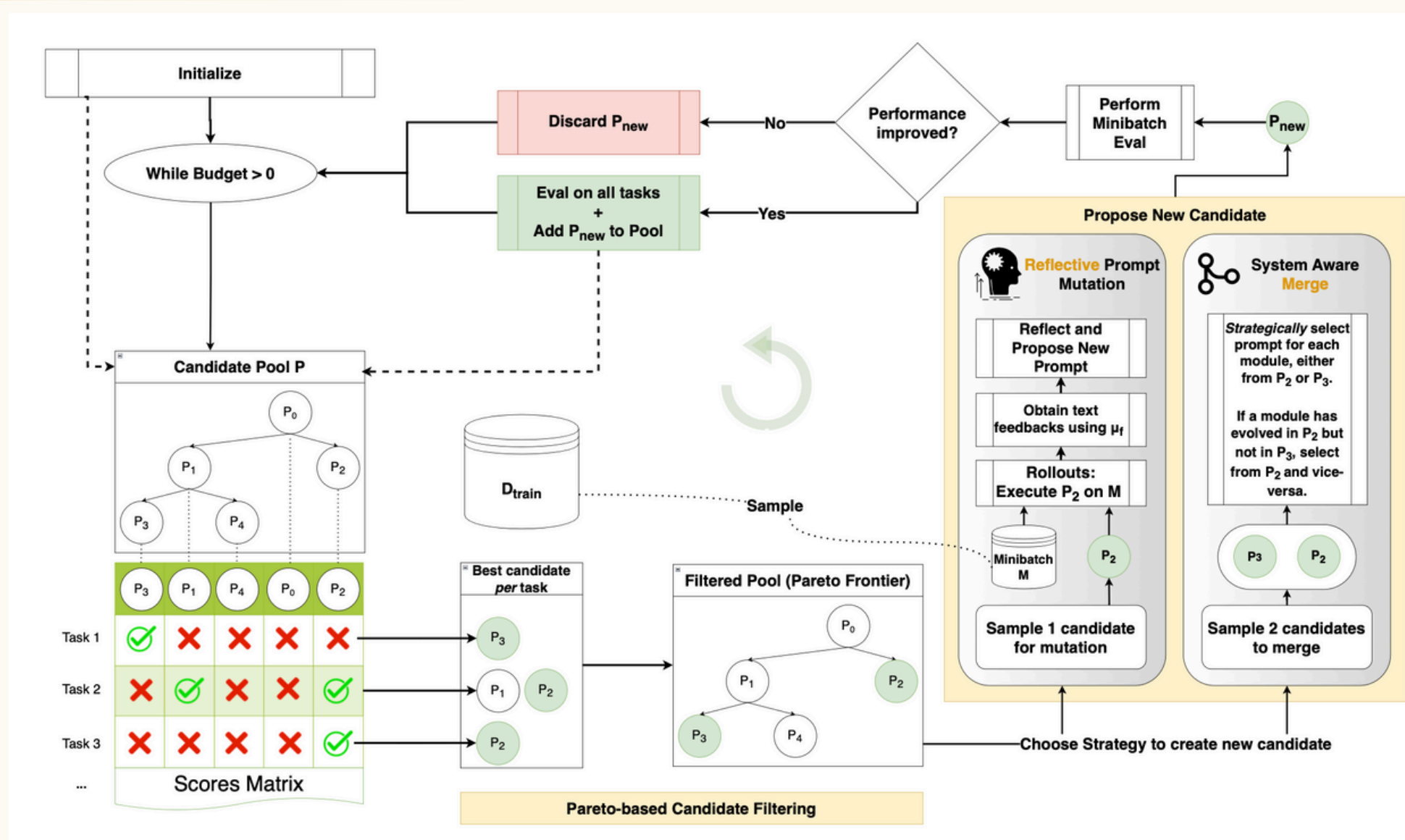
虽然使用RLHF/RLVR等方法对llm进行tuning是非常有效的，但我们也必须承认，在llm或者复合型智能系统(Compound AI Systems)中，prompt的设计依然至关重要。由此诞生了prompt optimization的研究方向，我们之前也读过几篇相关的论文，其中Stanford NLP 团队提出的DSPy无疑是其中的代表性工作。本文进一步将演化算法(Evolutionary Algorithms)引入prompt optimization，提出了一个全新的框架GEPA。

GEPA框架

实验设置

- 实验对象：Qwen3 8B 和 GPT-4.1 Mini

Model	HotpotQA	IFBench	Hover	PUPA	Aggregate	Improvement
Qwen3-8B						
Baseline	42.33	36.90	35.33	80.82	48.85	—
MIPROv2	55.33	36.22	47.33	81.55	55.11	+6.26
GRPO	43.33	35.88	38.67	86.66	51.14	+2.29
GEPA	62.33	38.61	52.33	91.85	61.28	+12.44
GEPA+Merge	64.33	28.23	51.67	86.26	57.62	+8.78
GPT-4.1 mini						
Baseline	38.00	47.79	46.33	78.57	52.67	—
MIPROv2	58.00	49.15	48.33	83.37	59.71	+7.04
GEPA	69.00	52.72	51.67	94.47	66.97	+14.29
GEPA+Merge	65.67	55.95	56.67	96.46	68.69	+16.02



思考

作者来自DSPy项目团队，虽然GEPA没有开源，但未来很大可能会作为模块集成进DSPy系统中。关于prompt optimization，认为目前有一种观点认为随着llm能力提升，prompt已经不那么重要了，我倒并不这样认为，对于chatbot用户来说，或许prompt不那么重要了，那是因为chatbot提供方/开发者们在负重前行，他们做了大量的prompt优化，包括memory、context engineering无非是prompt optimization的一种形式而已。