Simplicity is Enough

ReAct Agents for Prompt Optimization

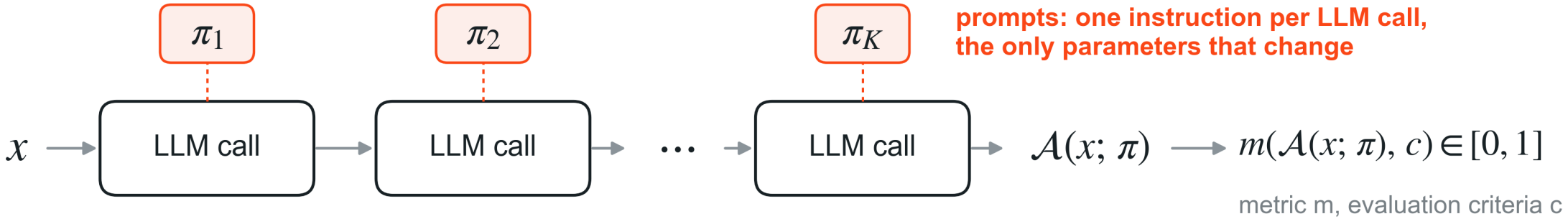
Andrei Dumitrescu[†], Andrei C. Rusu[†*]
Adrian Catalin Badea, Bai Li, Cosmin Maria,
Ion Anghelina, Tomasz Lukasz Religa, Dragos H Bobolea

UiPath · NeurIPS 2026

† Equal contribution * Corresponding author: rusu.c.andrei@gmail.com

Prompt optimization for compound LLM systems

A compound system chains several LLM calls, each steered by its own prompt.



$$\pi^* = \arg\max_{\pi} \mathbb{E}_{(x,\,c)\sim\mathcal{D}_{\text{train}}} \left[m(\mathcal{A}(x;\pi), c) \right]$$

Recent optimizers, such as APO, MIPROv2 and GEPA, search this space with evolutionary operators, Pareto fronts and mutation heuristics.

The price is outer-loop complexity, large rollout budgets and, at times, hand-crafted feedback.

Prompt optimization as a four-stage loop

GEPA and ReAPO implement the same stages with different mechanisms.

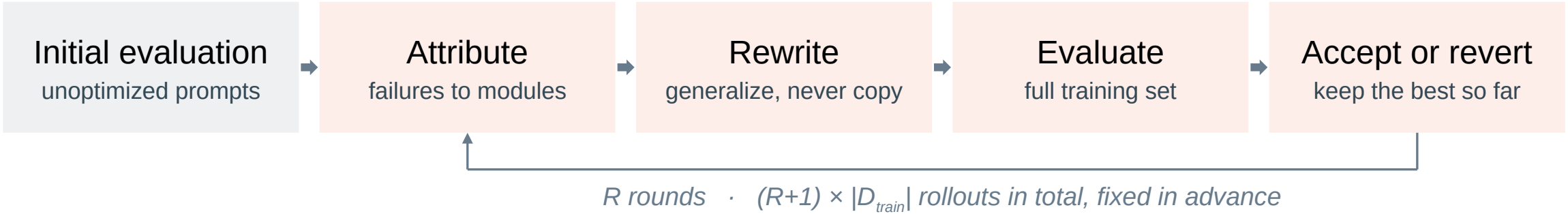
Selection	Evaluation	Proposal	Acceptance
GEPA · Pareto-based evolutionary search			
Pareto sampling	Minibatch screening	Round-robin order	Pareto archive
ReAPO · a single ReAct agent			
Single lineage	Full training set	Failure attribution	Accept or revert

Sufficiency: competitive performance does not require outer-loop machinery.

ReAPO: one ReAct agent runs the whole loop

Tools and protocol are reused unchanged across tasks.

Protocol · a natural-language skill prompt



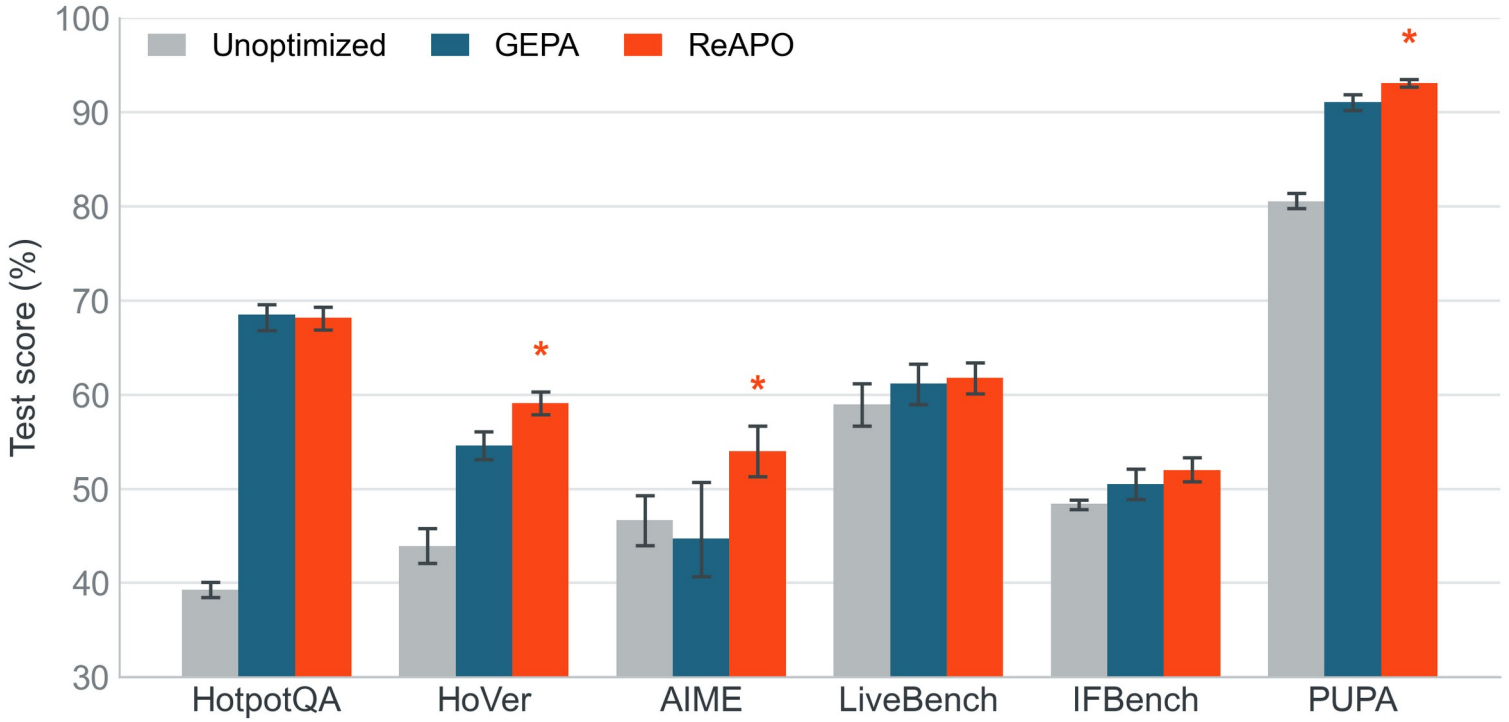
Agent and environment · tools exposed through an MCP server



Competitive scores at lower optimization cost

Six benchmarks, GEPA's splits and programs. System model: GPT-4.1 Mini.

Test score · GPT-5 as optimizer, five runs, 95% CIs * non-overlapping



Totals over the six benchmarks

Wall-clock time

GPT-5 as optimizer, as in the chart

ReAPO 5.3 h

GEPA 20.5 h

Rollout budgets

Fixed in advance, same for either optimizer

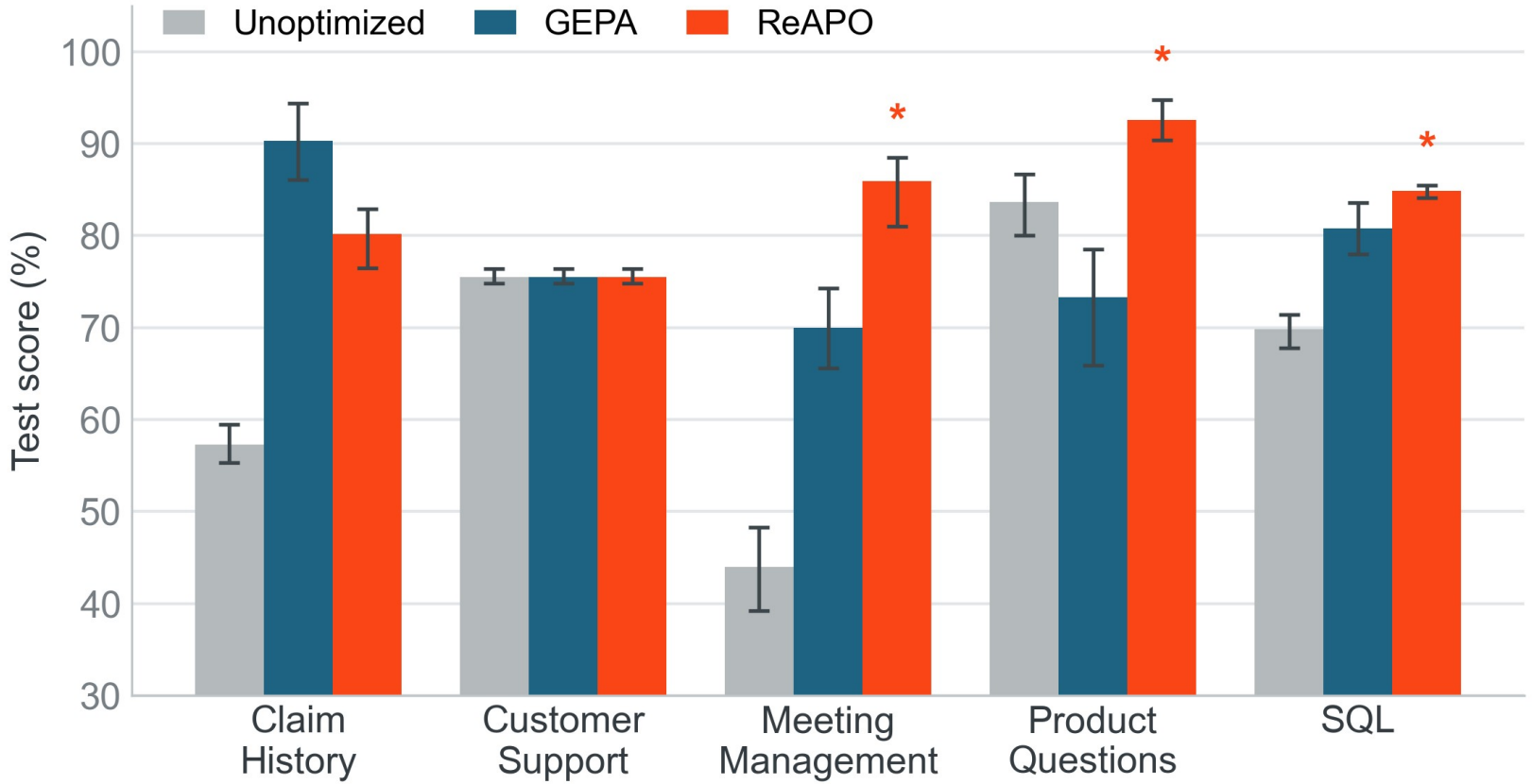
ReAPO 4,512

GEPA 23,619

The gains carry over to enterprise agents

Five tool-calling ReAct agents. GPT-5 on both sides, similar rollout budgets.

Test score · five trials, 95% CIs * non-overlapping



Customer Support

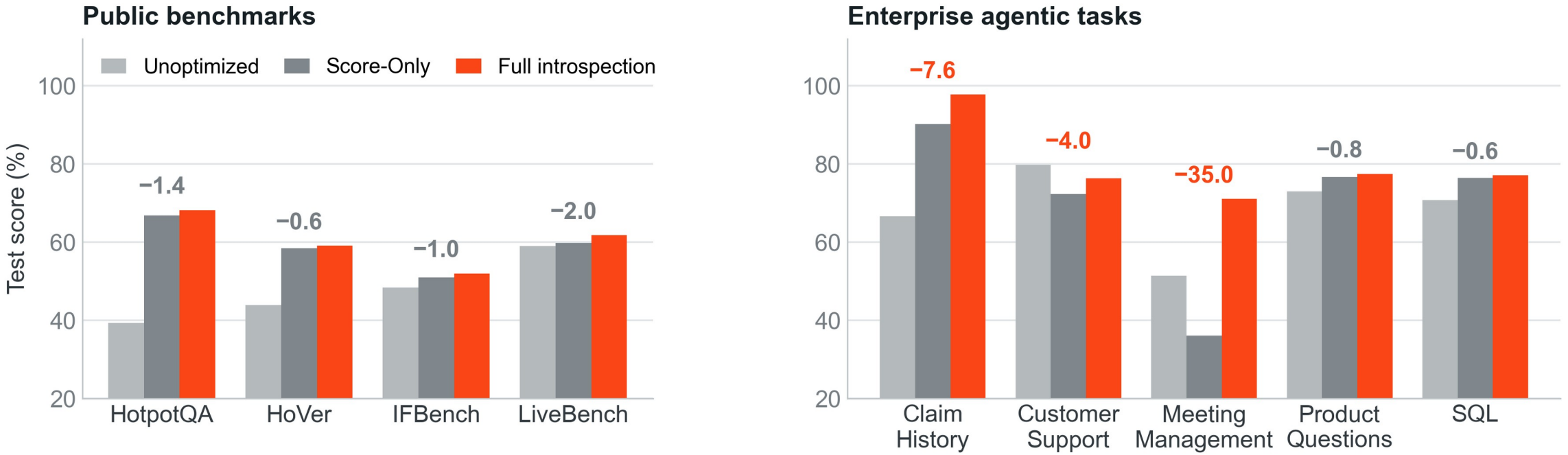
Neither method changed the unoptimized prompt.

Claim History

An early erroneous ReAPO revision propagated to every later round.

Benchmark saturation is measurable

Without the analysis tools, the optimizer sees only the aggregate score.



GPT-5 optimizer, GPT-4.1 Mini system model.

Aggregate scores suffice on public benchmarks, not on specialized workflows.



Simplicity is enough

1. A single ReAct agent matches or surpasses state-of-the-art optimizers at lower cost.
2. The optimizer's familiarity with a task can compensate for the absence of introspection.
3. A reliable comparison requires confidence intervals and less saturated benchmarks.

Thank you!

Contact: Andrei C. Rusu · rusu.c.andrei@gmail.com

Code and prompts: github.com/UiPath/agents-gym-core

UiPath