

外部ツール連携と反復修正で長期タスクを自律完遂

事後学習スケーリングと実務ツールの強み

事後学習の強化 (RL)



外部ツール連携と
反復修正で安更ける
長期タスクを自律完遂

金融分析 (OfficeQA Pro v2)
で世界1位の64.4%を記録。

事後学習の強化で
世界第4位 (55点) に浮上

モデル構造を変更せず強化学習の
投入のみで性能を大幅更新。

能力の非対称性 (偏科) とコスト構造の課題



ツール非依存の知識推論や
ワンショット完成度に課題

ツールなしのHLEテストでは
51.1%へ低下し、順位が後退。



API単価33%値下げ

推論量の増大

API単価33%値下げも
推論量の増大で実コストは相殺

常時思考や試行錯誤ループにより出力トークン数が14%増加。