



A Minimal Agent for Automated Theorem Proving

Borja Requena, Austin Letson, Krystian Nowakowski, Izan Beltran, Leopoldo Sarra
[name]@axiomatic-ai.com

A minimal agent with feedback, memory and search achieves competitive theorem proving results at a fraction of the cost of SOTA systems

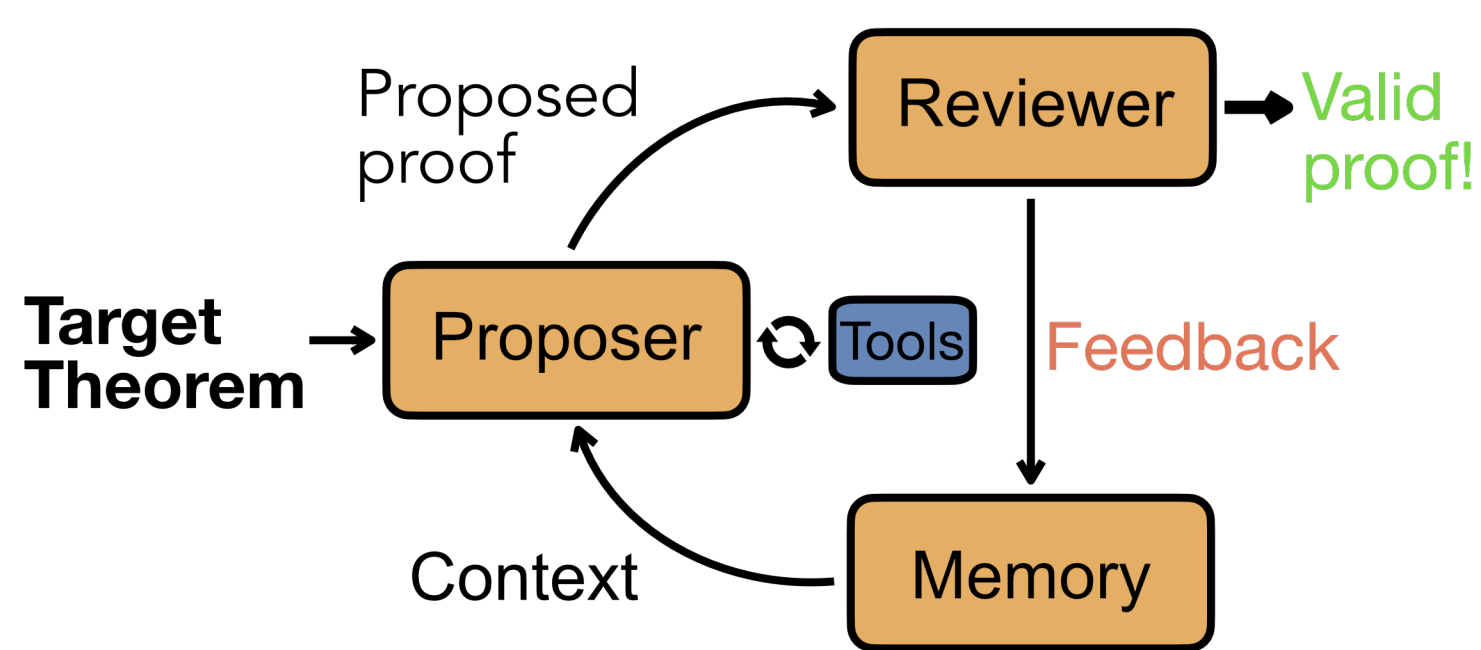
We provide **AxProverBase**: a strong baseline for future theorem proving work

PutnamBench
54.7% pass@1
 Near Hilbert with 200x less tokens

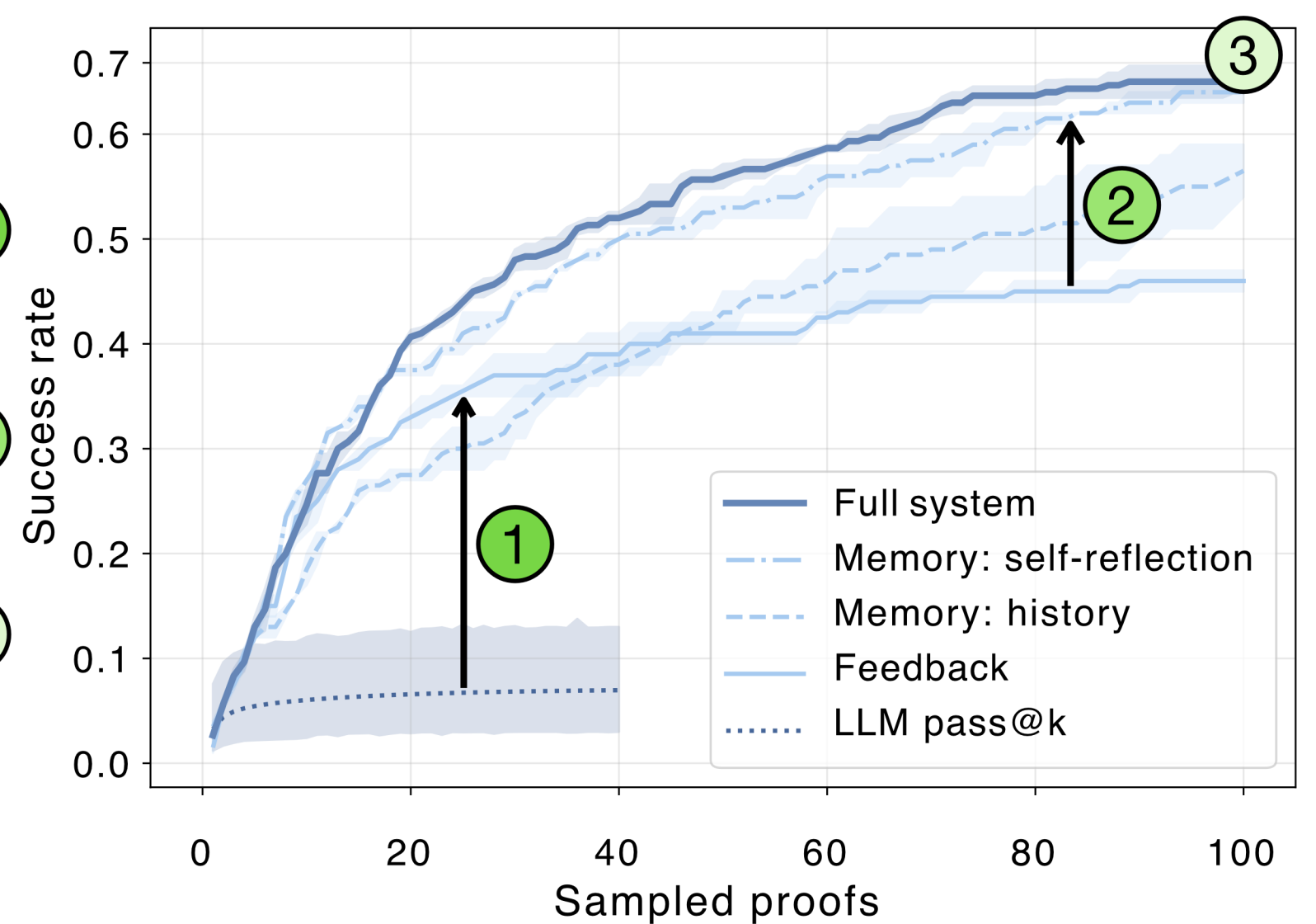
FATE M / H / X
98 / 66 / 24% @ 1
 Strong research level algebra

LeanCat
59.0% pass@1
 Big improvements on category theory

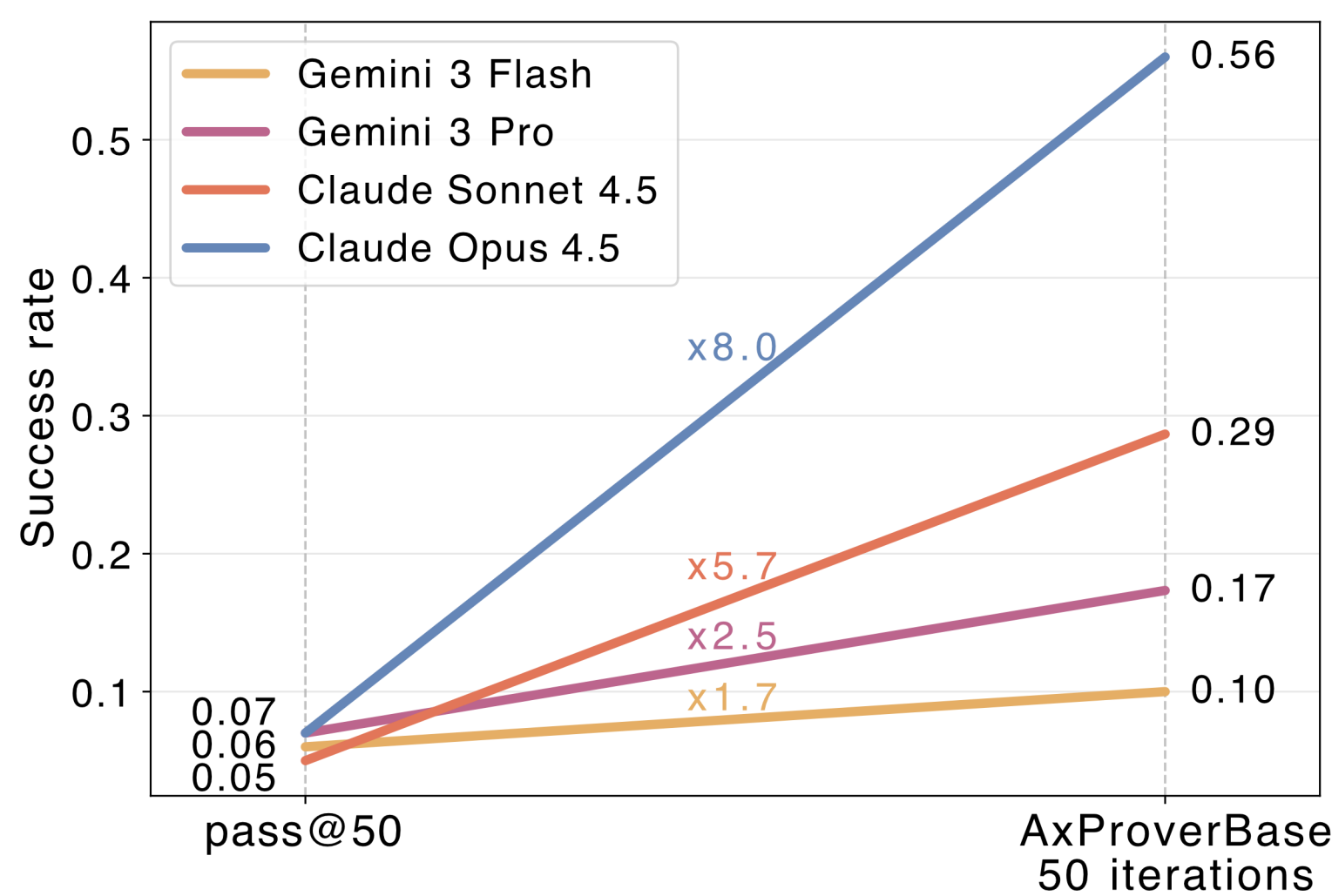
Three main elements unlock the potential of frontier LLMs:



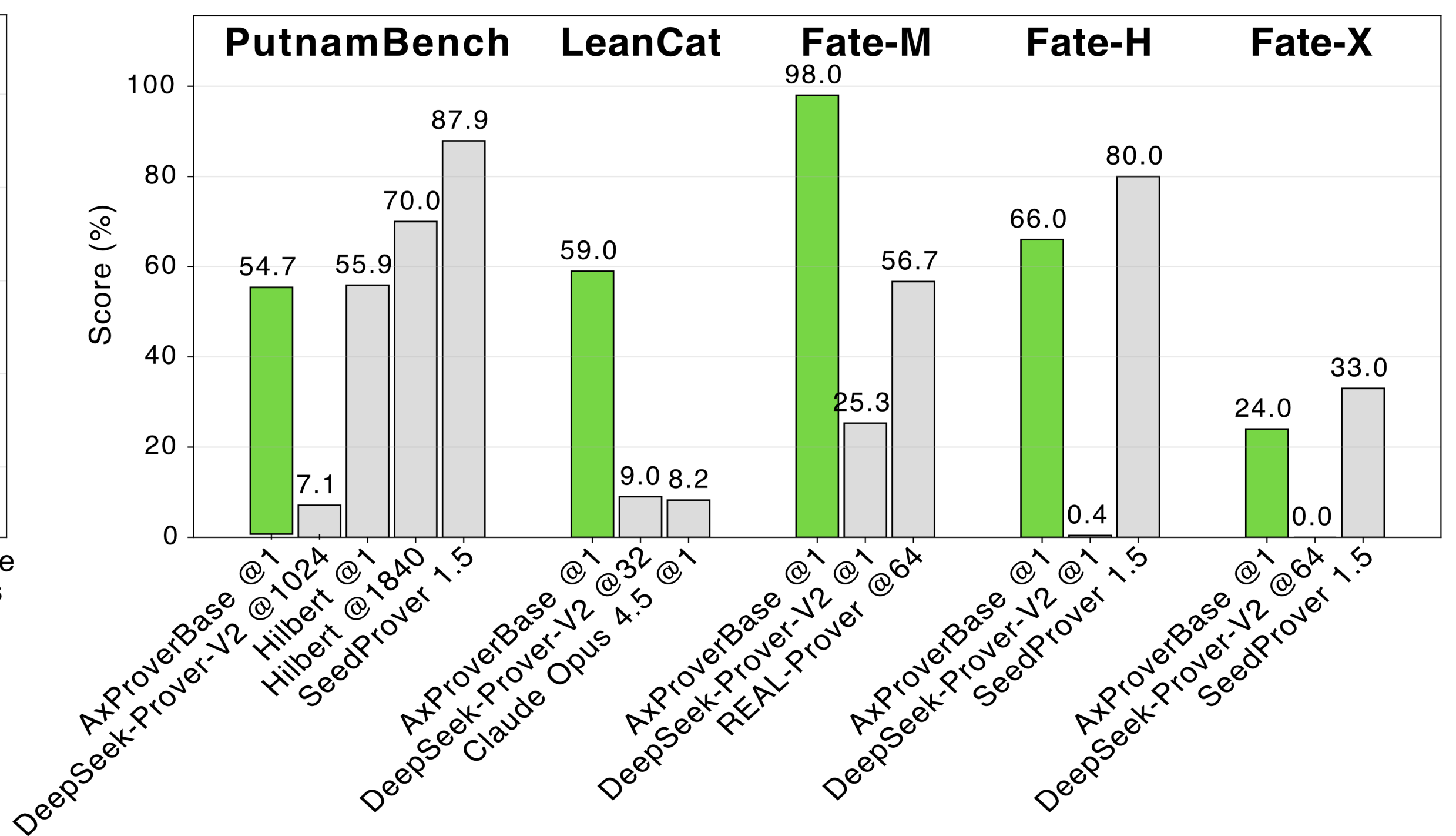
- Feedback:** the big gain! Enables iterative refinement ①
- Memory:** the second driver Reduces trivial errors ②
- Search:** the cherry on top ③



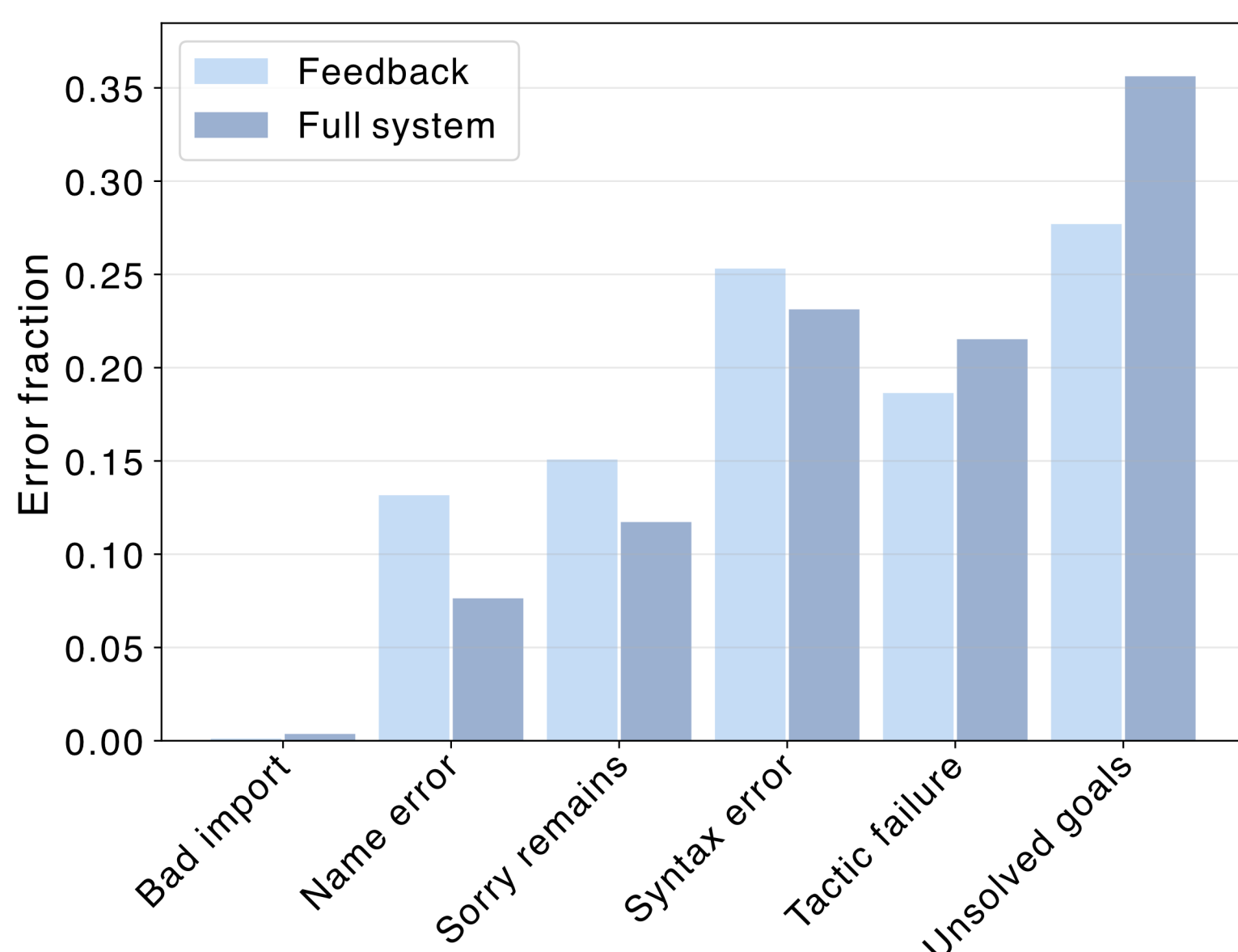
The agentic framework does not remove the dependence on the underlying model: stronger models benefit more



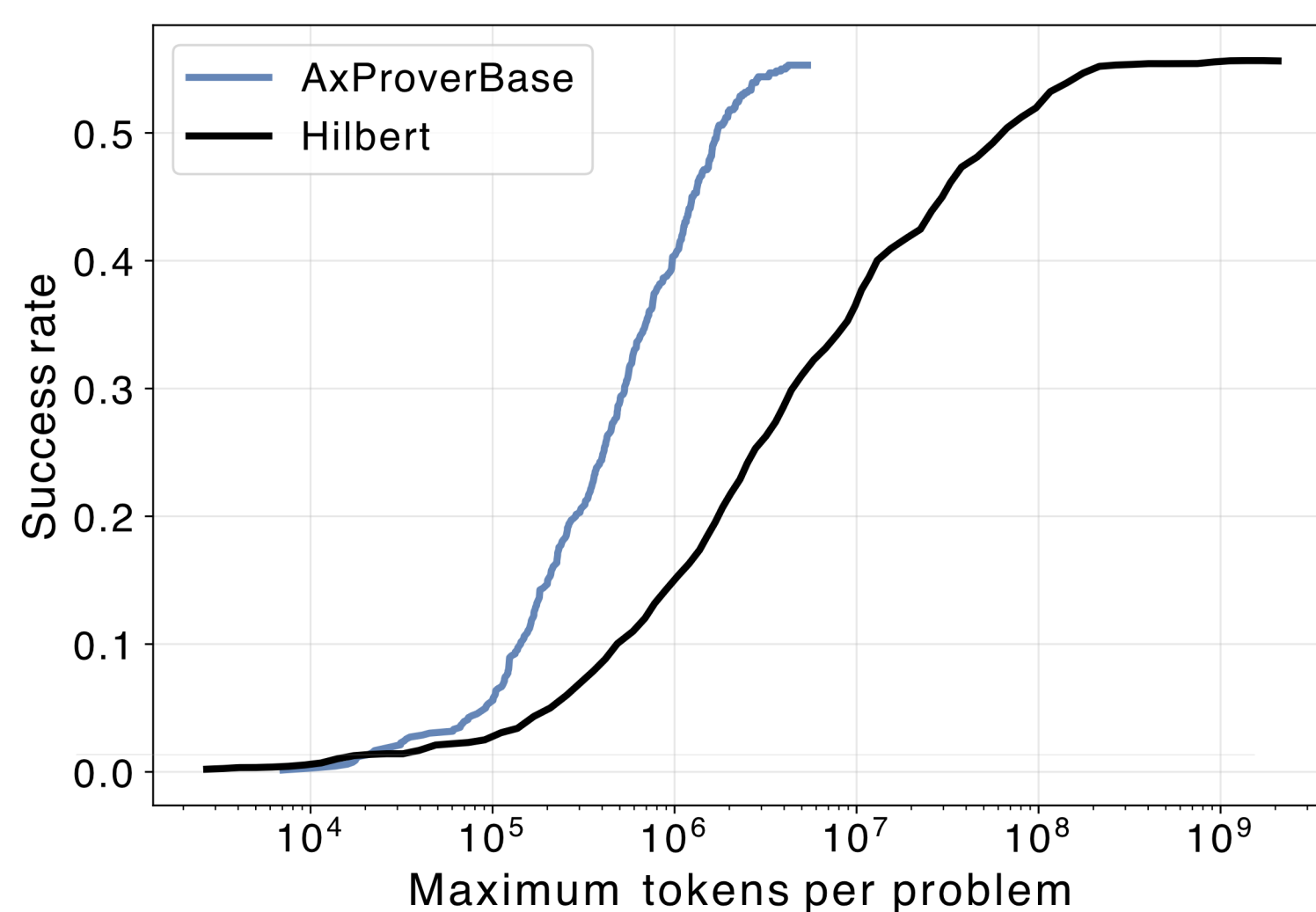
The best configuration of AxProverBase reaches competitive performance across multiple domains



System improvements reduce meaningless mistakes: less name and syntax errors, more unsolved goals



Max AxProverBase performance is matched by Hilbert with 200x more tokens



AxProverBase is open source and available for everyone to use and build upon

Sign up for our beta program and help us shape a great tool

Try it here!

