Cogentic: Multi-Agent Orchestration for Automated Proof Discovery

Research paper discussion | 25-minute main talk

Paper by Yang Cai, Vineet Gupta, Yanchen Jiang,
Christopher Liaw, Aranyak Mehta, Grigoris Velegkas, and Di Wang

Google Research | Additional affiliations: Yale; Google DeepMind

arXiv:2609.40324v1 | 30 September 2026
Reading prepared 3 October 2026

Turn failed proof attempts into reusable, verified progress.

arxiv.org/abs/2609.40324v1. A reading presentation; not authored by the paper's authors.

Open problems require progress across attempts

Research question

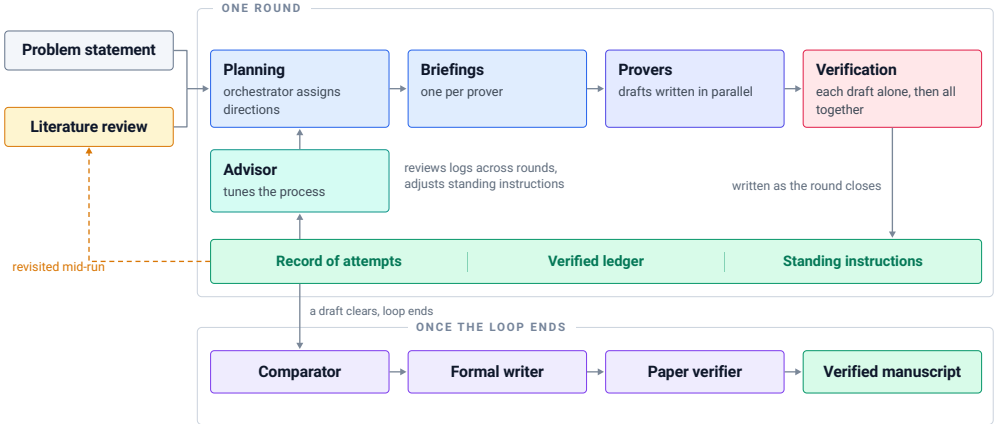
How can a language model explore competing proof directions,
overcome subtle errors, and preserve useful intermediate results?

- **Search:** one plausible approach can stall at a technical gap.
- **Verification:** natural-language arguments lack a cheap, faithful score.
- **Memory:** correct lemmas and failed conjectures need different treatment.

Cogentic produces mathematical prose for expert review.

Paper, Sec. 1, PDF pp. 1–2.

Each round writes the starting point for the next



Plan → prove → verify → retain progress.

Original Fig. 1, PDF p. 3. Reused from the authors' TeX assets.

Independent briefings preserve different approaches

Orchestrator

Assigns a claim, bound, construction, counterexample search, or repair task.

Allocates prover slots and manages shared files.

Summarizer → prover

Selects relevant attempts, critiques, and verified lemmas.

Each briefing is produced independently; the prover chooses its next step.

Long documents are passed as paths for selective reading.
Literature reviewers can return when a technical barrier persists.

Paper, Secs. 2–2.1, PDF pp. 2–4. Functional roles follow Sec. 2.1; Fig. 1's caption calls the briefing writer an advisor.

Acceptance requires two scopes of adversarial review

One draft at a time

Check the candidate's steps, assumptions, and citations.

All drafts in the round

Compare arguments and identify shared blind spots.

$$\text{Accept}(P_i) = V_{\text{local}}(P_i) \wedge V_{\text{cross}}(P_i; \mathcal{P}_r)$$

P_i: a candidate proof; $\mathcal{P}_r$: the round's proof batch.

Both reviewers demand justification rather than presume correctness.

Paper, Sec. 2.2, PDF p. 4. Equation is explanatory notation, not a paper equation. No verifier error rate is reported.

A rejected proof can still contribute a valid lemma

Record of attempts

Target claim, attempted method, and the objection that defeated it.

Supports allocation and avoids repeating errors.

Verified ledger

Extract a valid fragment.
State it as a self-contained lemma.
Verify it again in isolation.
Make it available to later provers.

Verified counterexamples also exclude unattainable targets.
The memory unit is a checked claim with conditions.

Paper, Sec. 2.3, PDF p. 4. "Verified" denotes the system's review process.

Process adaptation and final writing get separate checks

- A **process advisor** reads logs across rounds: repeated gaps, missed errors, and weak review standards.
- It adjusts instructions and advises on allocation. The advisor and orchestrator may not supply their own mathematical opinions.
- After a proof passes, exploration may continue for a stronger result.

Consolidation

Compare accepted proofs $\rightarrow$ write a complete manuscript
$\rightarrow$ audit the manuscript against the accepted argument.

Paper, Secs. 2.4–2.5, PDF pp. 4–5. "Formal writer" means mathematical exposition here, not proof-assistant certification.

The evidence consists of five research results

- **Setting:** open questions in online learning, auction theory, and mechanism design, chosen within the authors' expertise.
- **Engine:** Gemini; roughly $O(100)$ calls for most problems and $O(1000)$ for the hardest, as reported by the authors.
- **Validation:** domain experts subsequently checked the proofs; companion papers develop the mathematics.

Table 1 compares mathematical bounds. It is not a controlled comparison of agent architectures.

Paper, Secs. 1 and 3, PDF pp. 1, 5–6. Exact per-task tokens, costs, and model configuration are not provided.

Inverse optimization achieves finite total shortfall

$\mathbb{B}$: unit ball in $\mathbb{R}^d$; hidden $w^* \in \mathbb{B}$; compact $X_t \subseteq \mathbb{B}$.

$$x_t \in \arg\max_{x \in X_t} \langle w^*, x \rangle, \qquad R_T = \sum_{t=1}^{T} \langle w^*, x_t - \hat{x}_t \rangle.$$

Theorem 1

Deterministic, proper, anytime algorithm with $R_T = O(d)$ for all T.
Per round: $O(d^2)$ arithmetic **plus one linear optimization**.

- Proper: commit to a nonzero linear objective estimate *before* seeing X_t.
- Prior efficient bound: $O(d \ln T)$; prior $O(d)$ rule was improper and costly.

Paper, Sec. 3.1, Theorem 1, PDF pp. 5–7. $\hat{x}_t$: learner's action; d: dimension.

A decreasing trace potential removes the horizon term

$g_t = -v_t$, with v_t a unit separation direction; H_t is a metric.

$$s_t = \sqrt{g_t^\top H_t^{-1} g_t}, \qquad \Phi_t = \operatorname{tr}(H_t^{-1/2}).$$

Self-normalize the rank-one metric increment by s_t. Then

$$\Phi_t - \Phi_{t+1} \geq \frac{\tau}{4} s_t^2, \quad \Phi_1 = d, \quad \Phi_t \geq 0 \implies \sum_t s_t^2 \leq \frac{4d}{\tau} = O(d^2).$$

Metric scale τ and iterate step α are both $\Theta(1/d)$:

$$R_T = O\left(1/\alpha + \alpha \sum_t s_t^2\right) = O(d).$$

Paper, Sec. 3.1, PDF pp. 6–7. The paper also notes a later, inefficient tight $O(\sqrt{d})$ bound; efficient $O(\sqrt{d})$ remains open in this account.

Two extra sellers suffice on the smaller side

Theorem 2: a two-sided Bulow–Klemperer result

m buyers, n sellers, $m \geq n \geq 1$. Buyer values first-order stochastically dominate seller costs.

$$\mathrm{STR}(m, n+2) \;\geq\; \mathrm{OPT}(m, n).$$

- Compare expected gains from trade: Seller Trade Reduction in the expanded market vs. first-best in the original market.
- Prior constant recruitment bound: at least 20,000 agents *per side*.
- Adding just one seller is insufficient, even starting from $m = n = 1$, within the stated mechanism restrictions.

Paper, Sec. 3.2, Theorem 2, PDF p. 7; Table 1, p. 6. Table 1 specifies DSIC, IR, weak budget balance for the lower bound; the theorem summary says prior-independent.

With many experts, anytime has no leading-order price

Loss vectors lie in $[0,1]^n$; regret is relative to the best fixed expert.

$$\underbrace{\sqrt{t \ln n/2}}_{\text{fixed horizon}} \quad \text{vs.} \quad \underbrace{\sqrt{t \ln n}}_{\text{previous anytime guarantee}} .$$

Theorem 3

Without knowing the horizon, simultaneously for all $t \geq 1$:

$$R_t \leq \left(1 + O\left(\sqrt{\frac{\ln \ln n}{\ln n}}\right)\right) \sqrt{\frac{t \ln n}{2}}.$$

The relative overhead vanishes as $n \to \infty$.

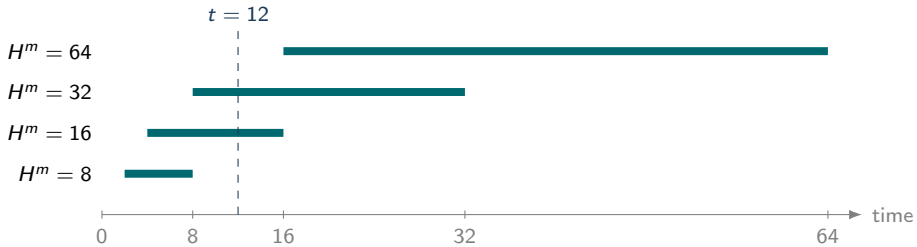
This is not exact equality for each fixed number of experts.

Paper, Sec. 3.3, Theorem 3, PDF pp. 7–8. For two experts, a strict anytime penalty is known.

Limit each instance's lifetime to bound aggregation cost

A multiplicative-weights instance for each $H^m = (1+\varepsilon)^m$:

wake at $\lfloor \delta H^m \rfloor$, retire after $\lfloor H^m \rfloor$.



At time t, only $O(\varepsilon^{-1} \log \delta^{-1})$ instances are awake.

A similar bound on arrivals during each lifetime controls mass redistribution.

Paper, Sec. 3.3, PDF p. 8. Diagram is a teaching example: grid ratio 2, $\delta = 1/4$; not the theorem's optimized parameters.

Simple selling mechanisms guarantee a 3.52 approximation

One additive buyer; values of the n items are independent.

$$M = \max\{\text{SRev}, \text{BRev}\}, \qquad \text{OPT} \leq 3.52M.$$

Simple mechanisms

SRev: best separate posted prices.
BRev: best price for the grand bundle.

Comparison

Approximation factor: $5.2 \to 3.52$.
Guaranteed share: $19.23\% \to 28.41\%$ of optimal revenue.

These shares are reciprocals of worst-case bounds.
The best known lower bound on the factor is 2.

Paper, Sec. 3.4, Theorem 4, PDF p. 8. Percentage conversions are explanatory calculations, not measured revenue gains.

Truncating at a joint revenue scale simplifies the proof

Let $V = \sum_j v_j$ and $M_{\max} = \max_j v_j$. Duality gives

$$\mathrm{OPT} \le \mathrm{SRev} + \mathbb{E}[V - M_{\max}].$$

Truncate at $M = \max\{\mathrm{SRev}, \mathrm{BRev}\}$ rather than at SRev:

$$V_M = \sum_{j=1}^{n} \min\{v_j, M\}, \qquad \mathbb{E}[V - M_{\max}] \le \mathbb{E}[V_M].$$

- The non-favorite welfare is controlled by one truncated quantity.
- Extremal bounds on $\mathbb{E}[V_M^2]$, at fixed mean, constrain $\mathbb{E}[V_M]/M$.
- The main paper sketches this route; the companion supplies the full constant calculation.

Paper, Sec. 3.4, PDF pp. 8–9. The inequalities concern expectations under the independent-value setting.

Randomized proportional allocation improves autobidding

Prior-free autobidding with return-on-spend (RoS) constraints.

$$\text{PoA} = \frac{\text{optimal allocation welfare}}{\text{worst equilibrium welfare}} \qquad \text{(smaller is better).}$$

r-Proportional First-Price Auction: pFPA$_r$

Bidder i wins query j with probability

$$x_{i,j} = \frac{b_{i,j}^r}{\sum_{k=1}^n b_{k,j}^r}.$$

The winner pays its own bid; an unopposed bidder wins for free.

r controls how strongly allocation favors the highest bid.
Deterministic mechanisms face a PoA barrier of 2 in this setting.

Paper, Sec. 3.5, PDF p. 9. $b_{i,j}$: bid; n: number of bidders.

Two bidders admit 1.5; general n improves on 2

Two bidders

$$\mathrm{PoA}(\mathrm{pFPA}_1) \leq 1.5.$$

Previous bound: 1.8.
Matching lower bound for the stated anonymous mechanism class and assumptions.

General $n \geq 2$

$$\mathrm{PoA}(\mathrm{pFPA}_{2n}) \leq 2 - \frac{1}{4n+1}.$$

Requires **undominated bids**: no feasible bid increase can win more while respecting RoS.

For even n, an anonymous-mechanism lower bound is $2 - 4/(n+4)$.
The general result matches the order of the $1/n$ improvement,
while the two-bidder constant is tight in the stated class.

Paper, Theorem 5, PDF p. 9; Table 1, p. 6 specifies anonymous, monotone mechanisms for the two-bidder optimum.

The supplied prompts already shape the research target

- **Two-sided markets:** the prompt specifies STR and asks for exactly two additional sellers.
- **Two-bidder autobidding:** both proportional allocation and the 1.5 conjecture are supplied.
- **Many experts:** the target is the existing conjecture that the anytime price tends to 1.
- **Simple revenue:** the prompt asks for factor 3, with factor 4 as a fallback; the reported result is 3.52.

Autonomous proof development occurs inside an expert-defined task.

Paper, Appendix A.2–A.5, PDF pp. 15–17; Sec. 3.5, p. 10.

Expert checking and later refinement remain part of the story

- The authors report no human mathematical intervention **during the runs**; experts subsequently checked the proofs.
- Autobidding initially assumed strictly positive values on all queries. Appendix A.5 says Gemini was later asked to remove that assumption.
- Human authors improved exposition, literature context, and in some cases extended the results.

A practical bottleneck

Candidate generation can outpace expert reading. Formal verification may help checking; understanding still takes work.

Paper, Sec. 3, PDF p. 5; Sec. 4, p. 10; Appendix A.5, pp. 16–17.

The next system study should isolate the sources of gain

What this paper establishes

Five reported mathematical advances, an explicit workflow, and expert-reviewed outputs.

What is not isolated

Equal-budget agent baselines, module ablations, verifier error rates, and a denominator of attempted problems.

Suggested experiments (reader's analysis):

- Remove the ledger or cross-draft verifier at matched compute.
- Trace which verified lemmas were reused in the final proof.
- Report failed tasks and the human effort needed to accept a result.

Assessment of the evidence in Secs. 1–4, PDF pp. 1–10. Proposed experiments are not reported results.

The reusable idea is evidence-bearing memory

1. **Explore independently.**
   Give provers distinct directions and selected views of history.
2. **Review at several scopes.**
   Check individual arguments, shared blind spots, isolated lemmas, and the final manuscript.
3. **Retain partial success.**
   Keep failed attempts as history and rechecked claims as reusable premises.

Five mathematical advances make this workflow worth studying.

Its component-level benefits remain to be measured.

Synthesis of Secs. 2–4, PDF pp. 2–10.

Primary paper and companion results

Cai et al. (2026), Cogentic
arXiv:2609.40324v1 | 30 September 2026

- Inverse optimization: arXiv:2609.13440.
- Two-sided markets: arXiv:2609.27304.
- Anytime experts: arXiv:2609.27206.
- Simple revenue: arXiv:2609.28873.
- Autobidding: *Efficiency of Generalized Proportional First-Price Auctions Under Auto-bidding*, listed as forthcoming in v1.

Companion references as listed in the primary paper, PDF p. 11. This talk follows the primary paper's statements and proof sketches. Project: sites.google.com/view/cogentic.

Backup: a round updates three different kinds of state

$$S_r = (A_r, L_r, I_r)$$

A_r: attempts and critiques; L_r: verified claims; I_r: standing instructions.

1. Plan directions and produce independent briefings from S_r.
2. Generate candidates; apply local and cross-draft verification.
3. Append attempts and objections to A_{r+1}.
4. Extract fragments, restate them, and reverify before adding to L_{r+1}.
5. Review logs to update I_{r+1} and guide the next round.

Teaching abstraction of Secs. 2.1–2.4, PDF pp. 4–5. The paper does not specify this state tuple as an algorithm.

Backup: why the awake pool is independent of time

Ignoring floors for the count, instance m is awake when

$$\delta H^m \le t \le H^m \quad \Longrightarrow \quad t \le H^m \le t/\delta.$$

On the grid $H^m = (1+\varepsilon)^m$, the pool size is

$$O\left(1 + \frac{\log(1/\delta)}{\log(1+\varepsilon)}\right) = O\left(\varepsilon^{-1} \log \delta^{-1}\right).$$

The missed prefix is handled recursively, costing

$$1 + \sqrt{\delta} + \delta + \cdots = \frac{1}{1-\sqrt{\delta}}.$$

Set $\varepsilon = \sqrt{\ln \ln n / \ln n}$ and $\delta = \varepsilon^3$ for large n.

Expanded explanation of the proof sketch in Sec. 3.3, PDF p. 8. Full analysis also controls arrivals and aggregation regret.

Backup: keep the two autobidding mechanisms separate

With two positive bids $(1, 2)$, the allocation probabilities are

$$r = 1 : \ (1/3, 2/3), \qquad r = 4 : \ (1/17, 16/17).$$

Setting	Mechanism	Reported PoA upper bound
Two-bidder result	pFPA_1	1.5
General result at $n = 2$	pFPA_4	$17/9 \approx 1.889$
General result at $n = 10$	pFPA_{20}	$81/41 \approx 1.976$

The general bound assumes undominated bids.

The bid example illustrates allocation, not an equilibrium or a PoA proof.

Teaching calculations from the mechanism and Theorem 5, Sec. 3.5, PDF p. 9.