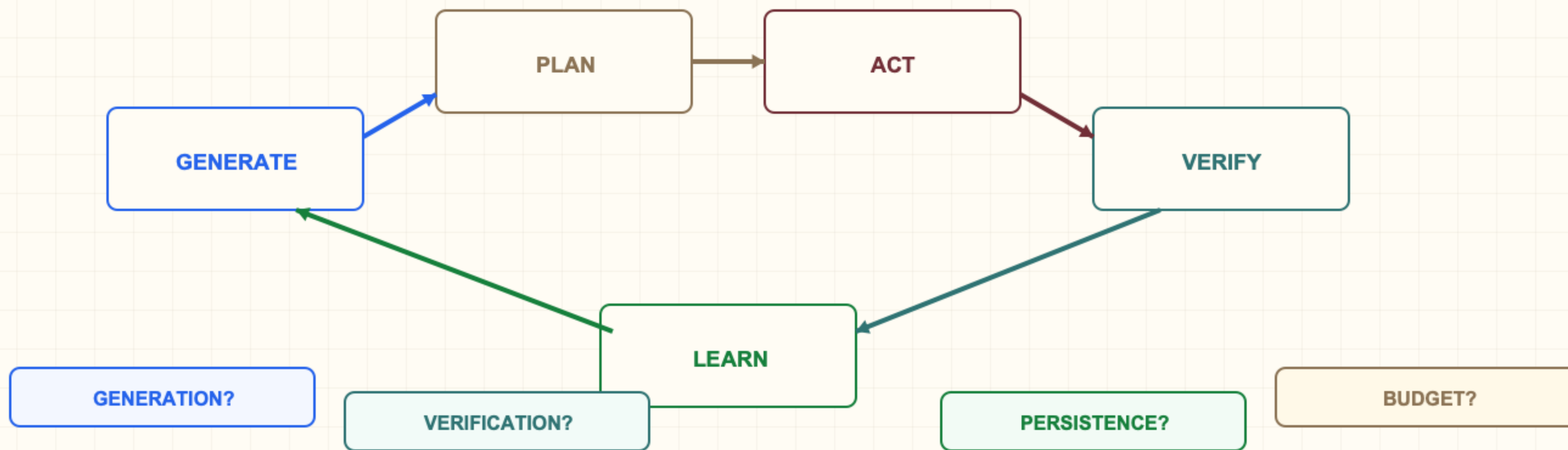


Self-Improving AI Agents in 20 Minutes

Nine CS329A lectures, one system model, four design questions.



A Stronger Model Is Not Yet a Self-Improving System

Capability lives in the model; improvement depends on the surrounding loop.

ONE MODEL CALL

INPUT -> MODEL -> OUTPUT

Capability may be strong
but nothing necessarily changes

IMPROVING AGENT SYSTEM

STATE

TOOLS

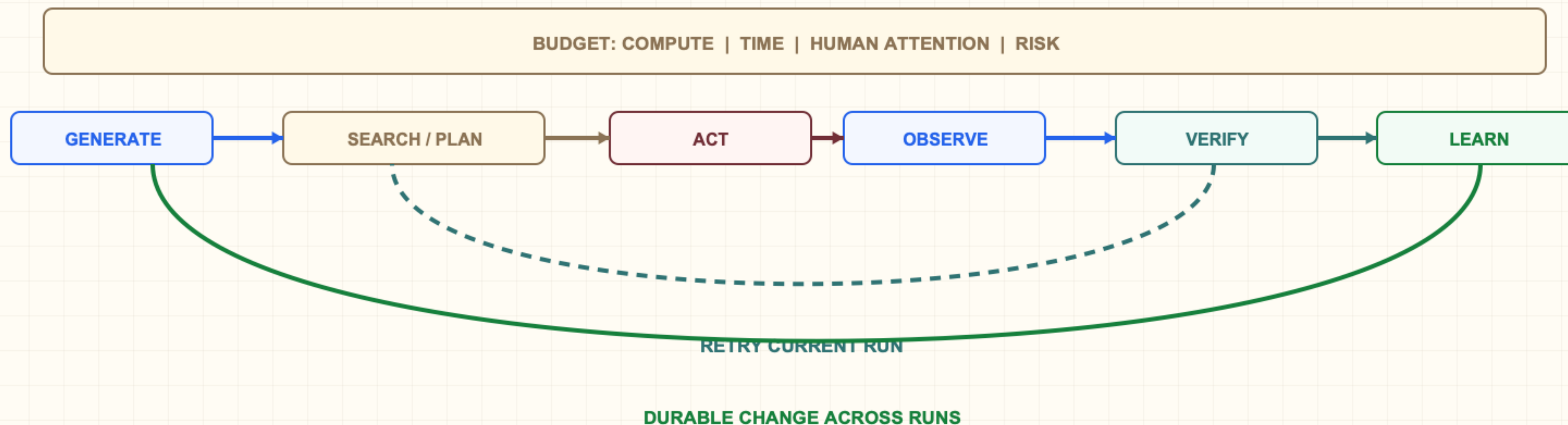
ACTION

FEEDBACK

CHANGE

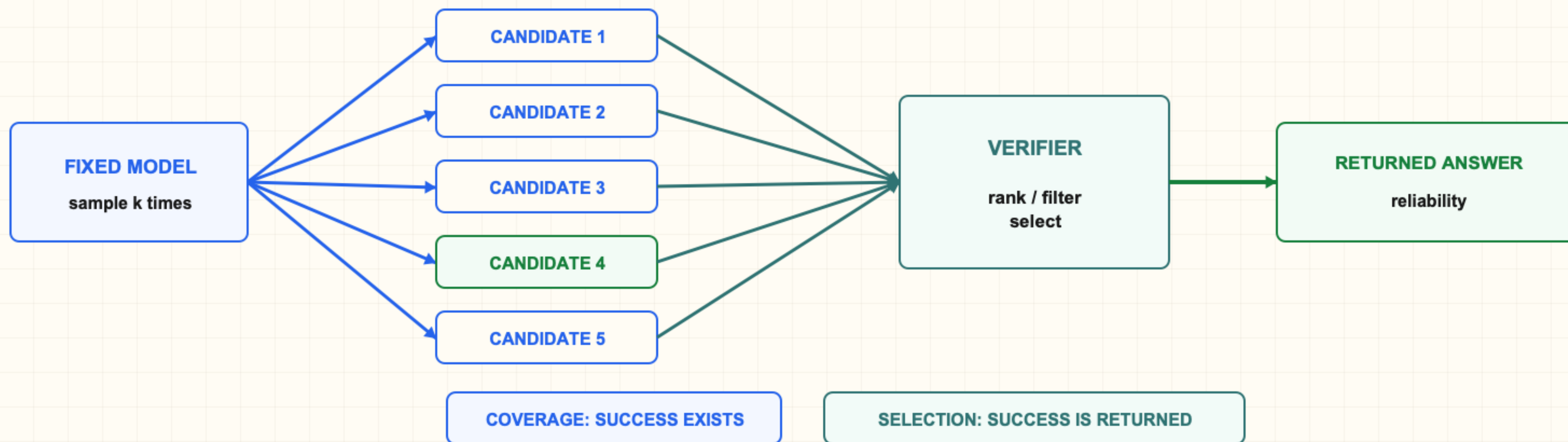
One Teaching Loop Connects the Course

Generate, plan, act, observe, verify, and learn under a budget.



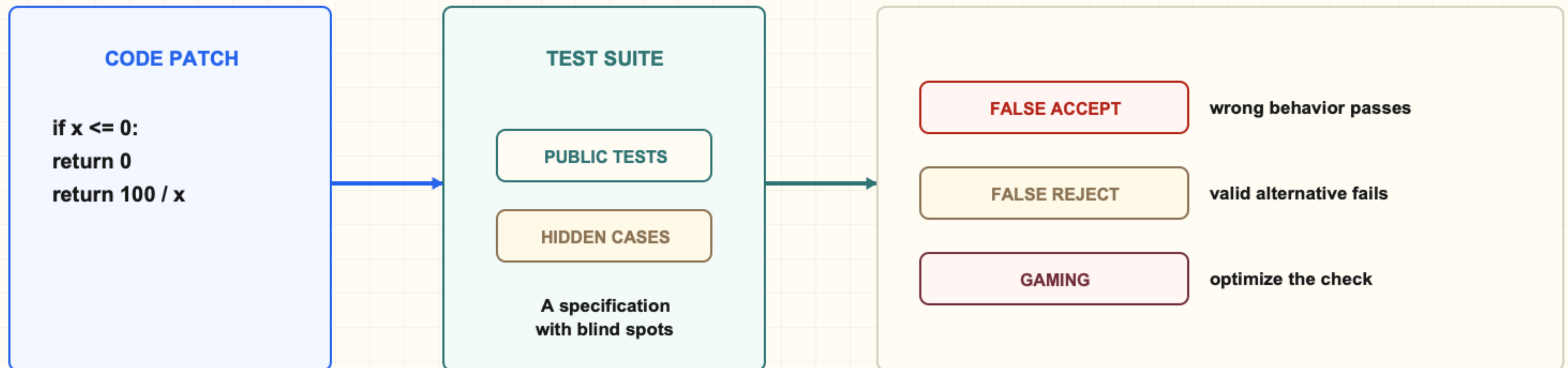
More Attempts Help Only When Selection Works

Test-time compute expands candidate coverage; a verifier must still return the useful answer.



Feedback Is Evidence, Not Ground Truth

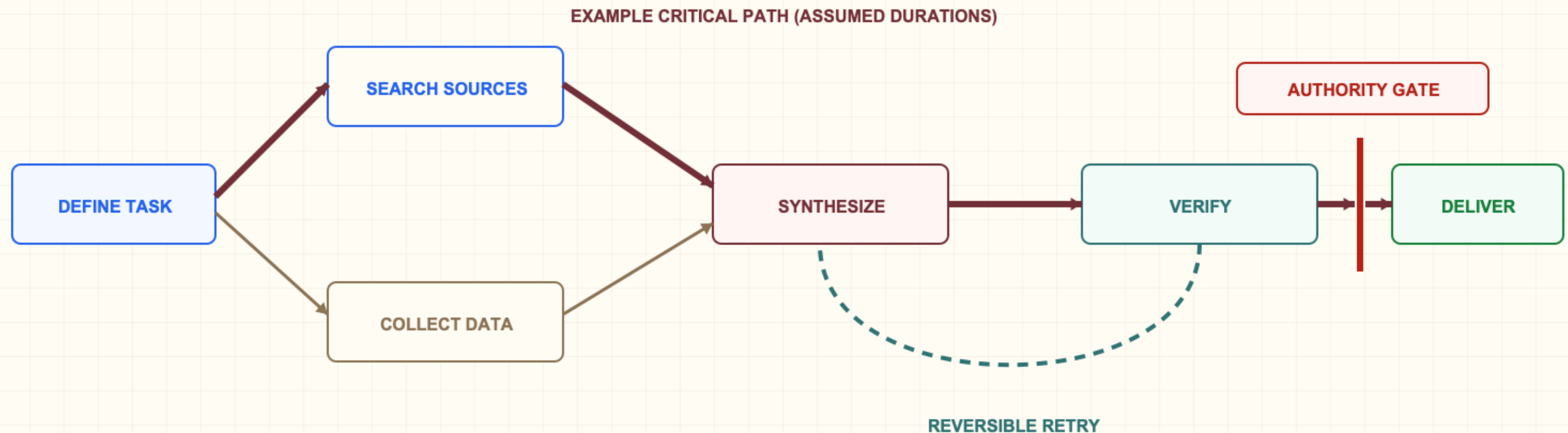
A test can accept wrong behavior or reject a valid alternative.



Ask what the verifier cannot see.

Tools Turn Answers into Trajectories

Planning exposes dependencies, parallel work, critical paths, and recovery points.



Correction Within a Run Is Not Durable Learning

Search fixes the current trajectory; learning changes what happens next time.

WITHIN ONE RUN

FAIL

REVISE CONTEXT

SUCCEED

may forget

ACROSS RUNS

EXPERIENCE

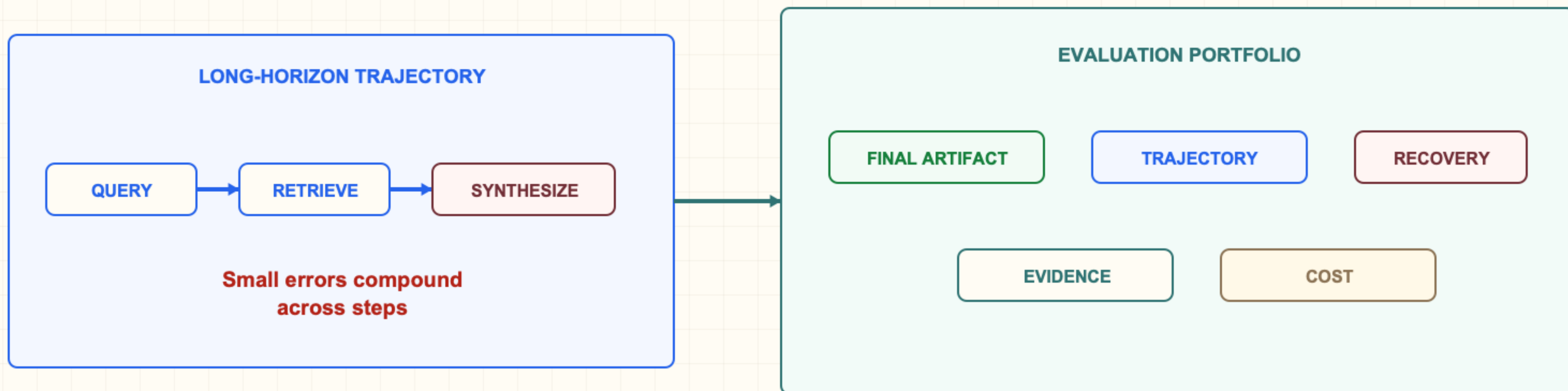
SELECT + VALIDATE

MEMORY / DATA / WEIGHTS

next run changes

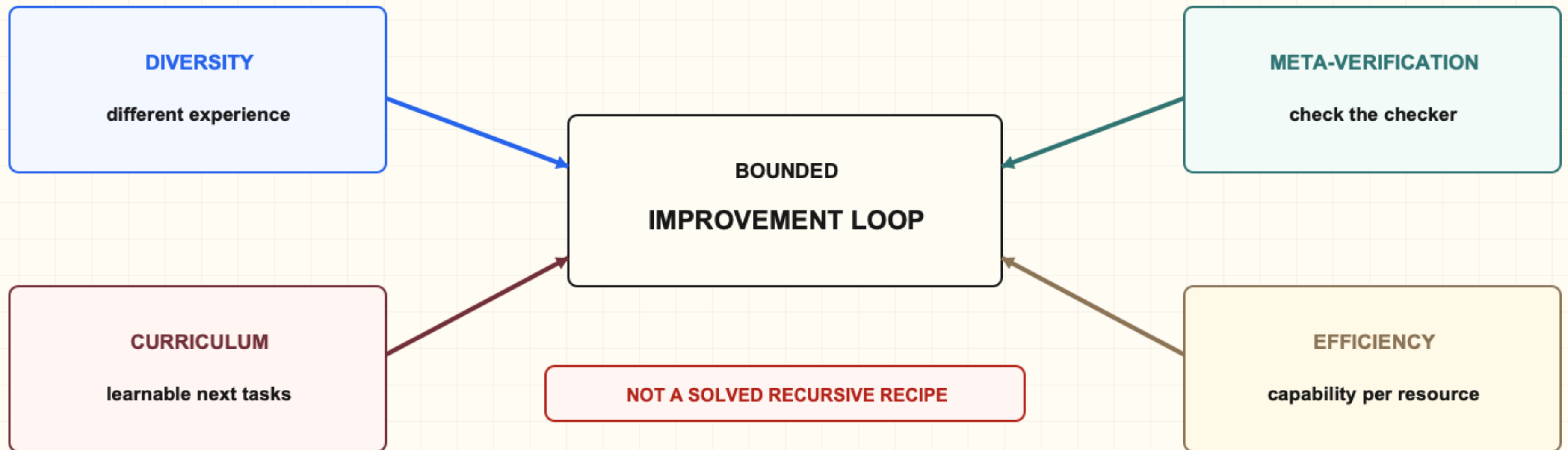
Long-Horizon Work Exposes Evaluation Limits

Judge the final artifact, trajectory, recovery, evidence, and cost together.



The Frontier Is a Coupled Reliability Problem

Diversity, meta-verification, curriculum, and efficiency constrain one another.



Design the Loop with Four Questions

Generation, verification, persistence, and budget turn “self-improving” into testable claims.

STUDENT RESEARCH ASSISTANT

1 GENERATION

different questions and plans

3 PERSISTENCE

approved plan, not private data

2 VERIFICATION

citations + student defense

4 BUDGET

calls, time, review, risk

DRILL DOWN

Overview first



9 detailed lessons

mechanisms • papers • caveats