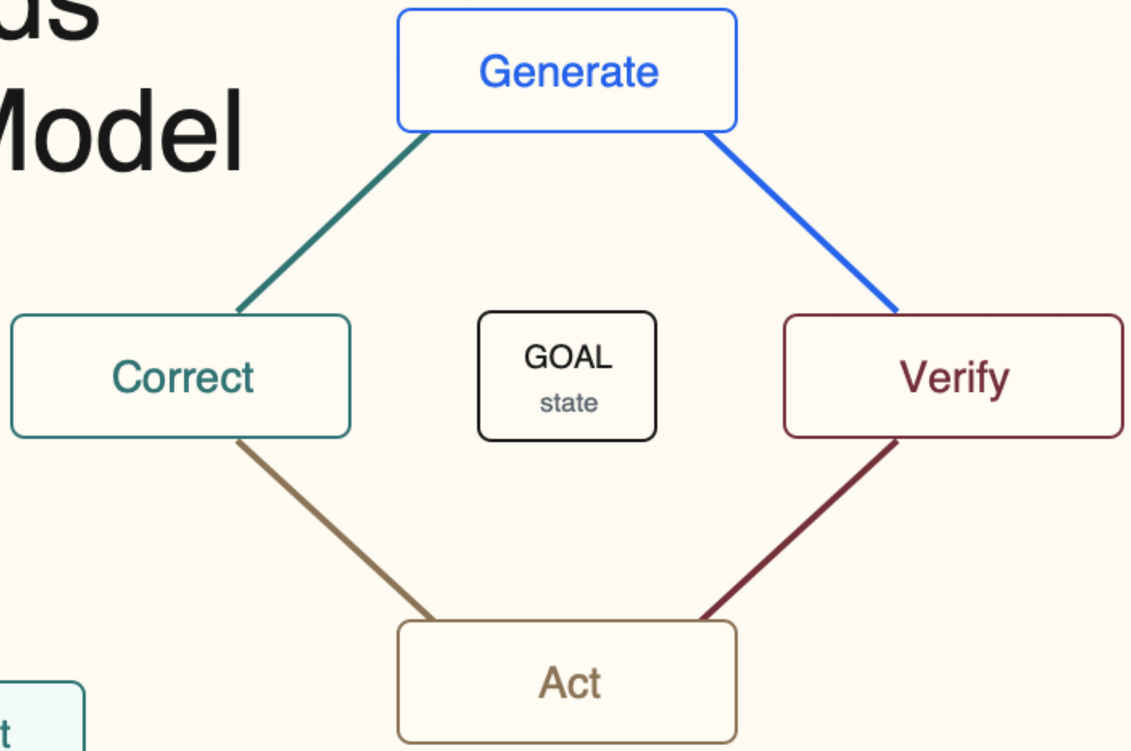
COURSE OVERVIEW

LECTURE 1 42:28-53:56

A Reliable Agent Needs More Than a Strong Model

Stanford CS329A Lecture 1 - from scaling laws to feedback-driven agent loops



Generate | Verify | Act | Correct

Verification is the narrow gate between possibility and reliability.

DECK SYNTHESIS

LECTURE 1 02:23-19:38 / 40:49-47:55

One Way to Organize AI Progress: Four Compute Frontiers

Each frontier improves a different part of the system.

01

Pretraining

What can it know?

02

Post-training

How should it behave?

03

Test-time

Which answer should it choose?

04

Orchestration

What should it do next?

CAPABILITY -> BEHAVIOR -> SEARCH -> ACTION

SCALING

LECTURE 1 02:23-10:49 / 58:12-60:32

Scaling Expanded Capability, but Emergence Is Not a Free Pass

Historical evidence supports capability gains, not an unrestricted law of emergence.

OBSERVED ASSOCIATION

lower test loss + broader benchmark capability

INTERPRETATION BOUNDARY

Larger models showed stronger few-shot and chain-of-thought behavior on selected benchmarks.

CAVEAT

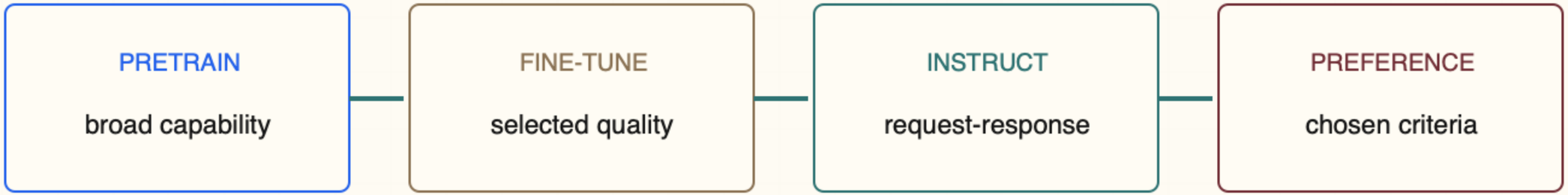
That observation does not prove a universal law, nor that the behavior was absent from training data.

BEHAVIOR SHAPING

LECTURE 1 11:20-19:25

Post-Training Turns Capability into Assistant Behavior

Capability and behavioral alignment are related, but they are not the same layer.



CAVEAT

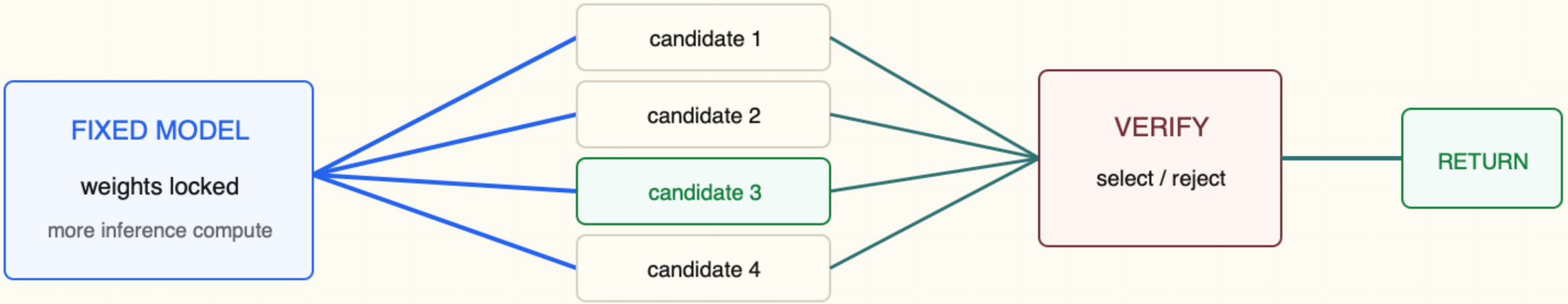
Preference optimization approximates evaluator preferences; it is not automatically a truth detector.

INFERENCE SEARCH

LECTURE 1 19:38-31:18

Test-Time Scaling Searches a Fixed Model

Spend more inference compute without changing the model weights.



Repeated sampling is one strategy; search, longer reasoning, and tools are distinct.

METRICS

LECTURE 1 20:50-27:16 / 36:33-37:04

Coverage Is Not Deployed Reliability

Finding a correct candidate does not guarantee returning it.

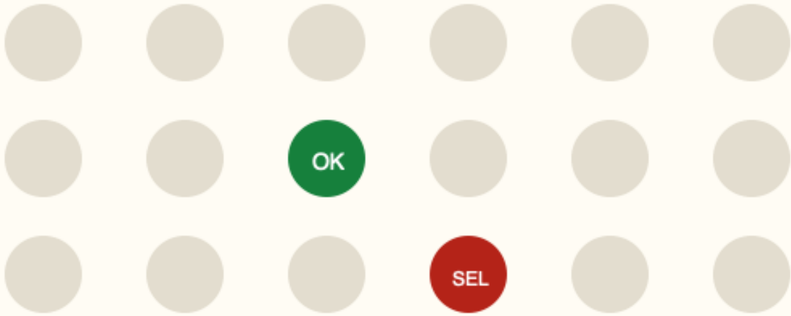
COVERAGE

Did any sampled candidate contain the correct answer?

$1 - (1 - p)^k$

RELIABILITY

green = correct red = selected



CAVEAT

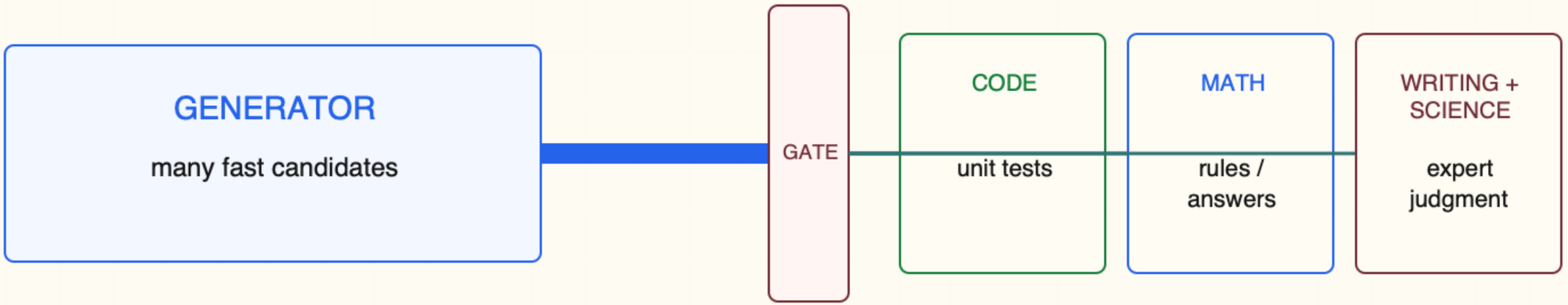
Formula assumes independent, identical attempts. Real samples correlate, and selectors make errors.

BOTTLENECK

LECTURE 1 47:09-51:16

Reliable Verification Often Scales More Slowly

The cost of checking an answer depends on the domain.



Weak verification can select a polished mistake.

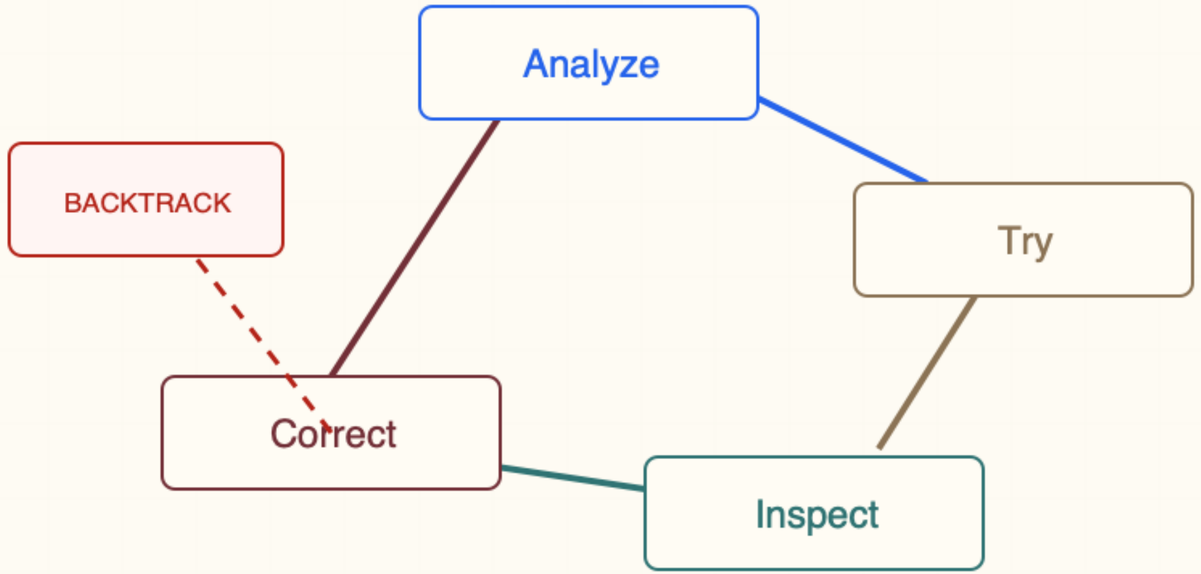
REASONING

LECTURE 1 31:22-40:43

Reasoning Models Turn One Answer into a Search Process

A useful operational view is iterative search with feedback and backtracking.

Reasoning is useful when feedback changes the next attempt.



CAVEAT

A visible reasoning trace is not proof that the reasoning is correct.

PERSISTENCE

LECTURE 1 28:53-30:12 + deck interpretation

Self-Improvement Has Two Different Meanings

The key test is whether useful change survives beyond one task.

WITHIN-RUN

retry -> inspect -> revise

change disappears when the task ends

ACROSS-RUN

learn -> persist -> reuse

- weights
- memory
- tools
- policies
- data

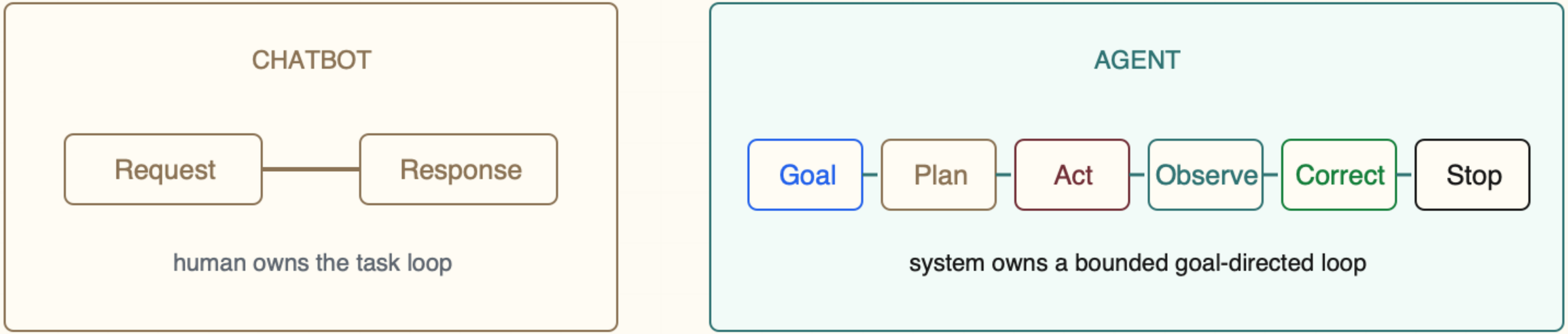
PERSISTENCE BOUNDARY

AGENCY

LECTURE 1 40:49-43:26

An Agent Owns a Goal-Directed Task Loop

Agency begins when feedback changes the next action.



ORCHESTRATION

LECTURE 1 43:30-47:55

Workflow Graphs Trade Flexibility for Control

The task's risk and uncertainty should determine the orchestration design.

OPEN LOOP

adaptive / less predictable

WORKFLOW GRAPH

observable / less flexible

Choose by task risk, uncertainty, and audit requirements.

DECK DIAGNOSTIC

LECTURE 1 47:09-58:04

A Useful Course Lens: Feedback Quality

Use three questions before accepting any self-improving-agent claim.

01

What generates alternatives?

02

What verifies success?

03

What persists after the task?

NEXT: test-time compute scaling

External validation still matters.