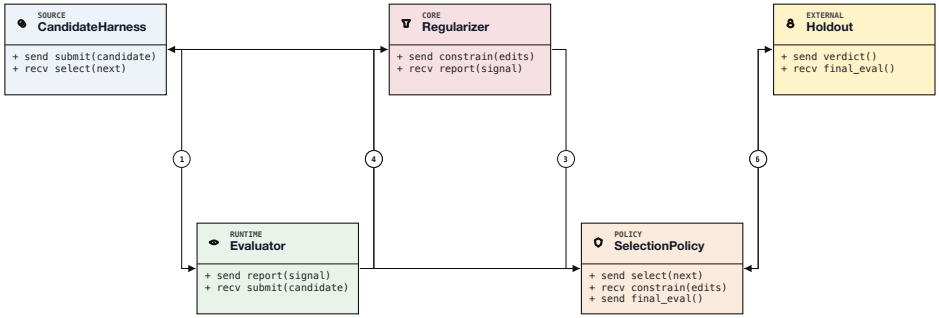
E03 Regularized harness improvement

Evolution · A² A2/A3 · funnel role collaboration

Limit overfitting during repeated edits.

STRUCTURE · PARTICIPANT COLLABORATION



PATTERN INTERFACE

context
Multiple harness edits are proposed over repeated rounds.

trigger
An evaluator accepts or rejects an edit.

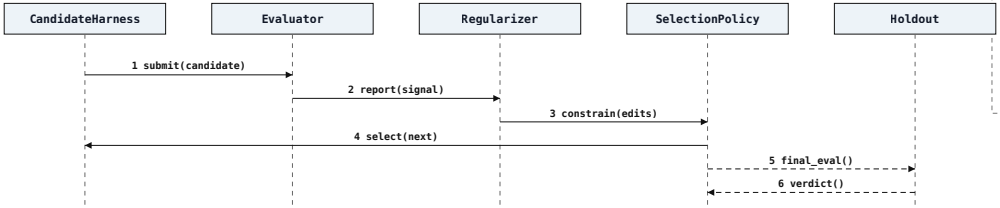
structure
Validation split, regularization signal, edit budget, and final holdout.

effect
Selection controls complexity and overfitting before final evaluation.

CONTRACT ANNOTATIONS

requires	Acceptance and final evaluation data have distinct roles.
ensures	Accepted edits include the screening evidence that justified them.
invariant	Final holdout is not used for iterative selection.
violation	Leakage or unstable gain stops acceptance or downgrades the claim.

INTERACTION



MESSAGE TRACE `1 submit(candidate) → 2 report(signal) → 3 constrain(edits) → 4 select(next) → 5 final_eval() → 6 verdict()`

Evidence: 12/41 repositories, 12/36 families; 12 trajectories (5 multi-checkpoint); exact audit: not sampled.

Variation: Holdout, critic, complexity penalty, edit budget, and pruning.

Negative control: Best-of-many tuning on the test metric is not regularized improvement.