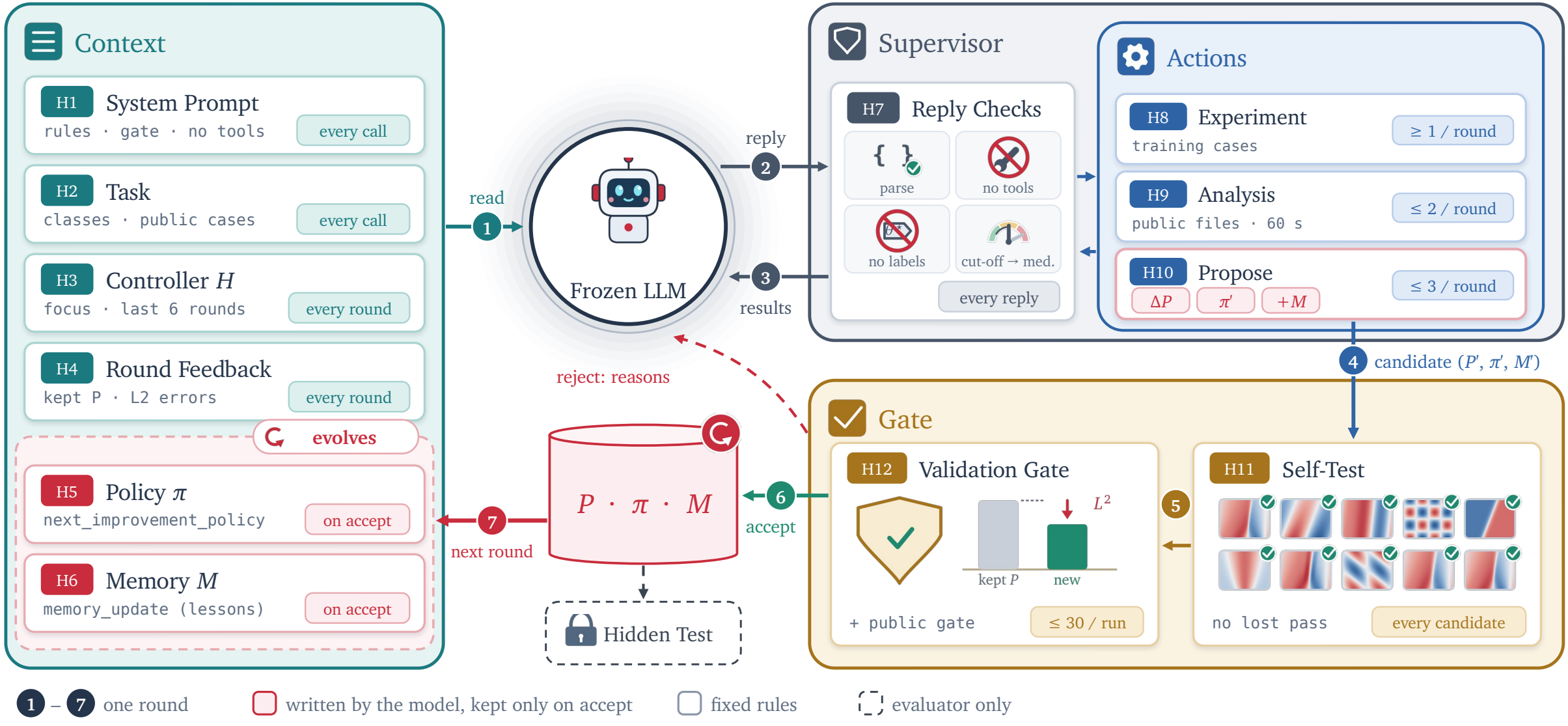

Context
H1 System Prompt
rules · gate · no tools
every call
H2 Task
classes · public cases
every call
H3 Controller H
focus · last 6 rounds
every round
H4 Round Feedback
kept P · L2 errors
every round
evolves
H5 Policy π
next_improvement_policy
on accept
H6 Memory M
memory_update (lessons)
on accept
read
1
Frozen LLM
reply
2
results
3
reject: reasons
Supervisor
H7 Reply Checks
parse
no tools
no labels
cut-off → med.
every reply
Actions
H8 Experiment
training cases
≥ 1 / round
H9 Analysis
public files · 60 s
≤ 2 / round
H10 Propose
ΔP
π′
+M
≤ 3 / round
4
candidate (P′, π′, M′)
Gate
H12 Validation Gate
L²
kept P
new
+ public gate
≤ 30 / run
5
H11 Self-Test
no lost pass
every candidate
6
accept
P · π · M
7
next round
Hidden Test
1 – 7 one round
written by the model, kept only on accept
fixed rules
evaluator only