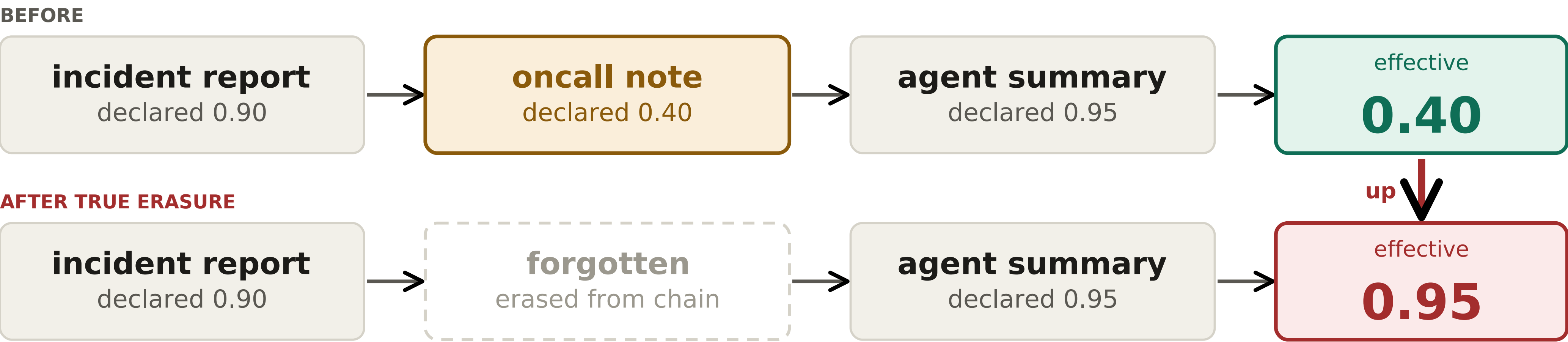


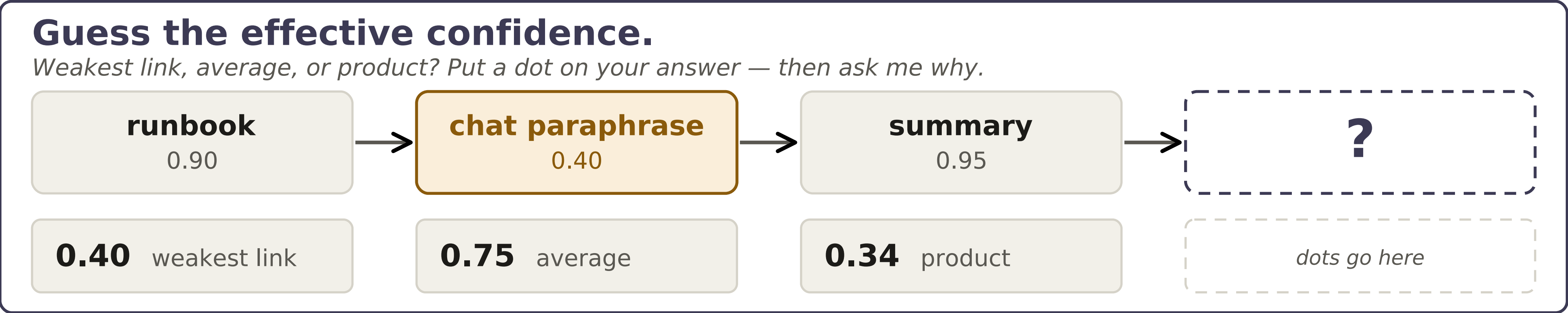
## Delete the weakest link and the chain gets stronger.

Forgetting the evidence strengthened the conclusion — and nothing in the store said so.



**min(∅) = 1.0**  
the minimum over an empty chain imposes no bound — same test, one configuration flag

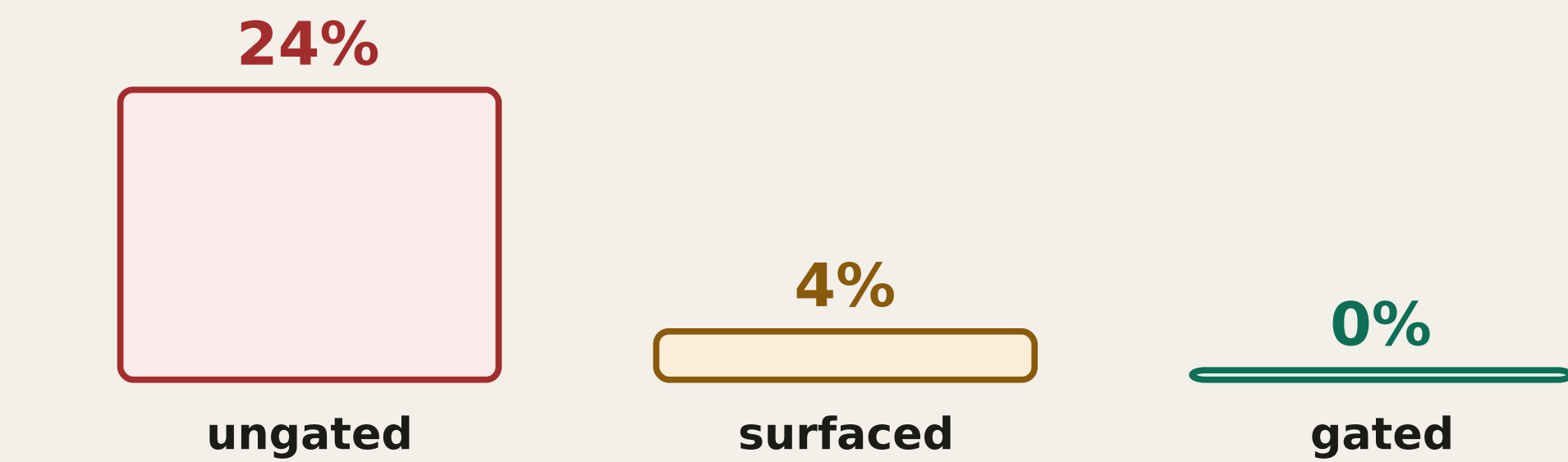
**24%** of writes declared at 0.9 or above, in a sixty-round three-agent simulated run, rested on one 0.40 guess. Effective confidence was computed on every one. **Nothing read it.**  
*A reader that honours the number perfectly gets that to 4%. A gate gets it to zero.*



| 26 EXECUTABLE CHECKS · SAME TEST, ONE FLAG                 |           |         |            |
|------------------------------------------------------------|-----------|---------|------------|
|                                                            | tombstone | cascade | erase      |
| P1 Provenance walk<br>every memory back to its roots       | PASS      | PASS    | PASS       |
| P2 Confidence cascade<br>weakest ancestor caps every claim | PASS      | PASS    | DIVERGENCE |
| P3 Forget operations<br>three forget modes, each receipted | PASS      | PASS    | PASS       |
| P4 Isolation<br>scopes, grants, opaque ancestors           | PASS      | PASS    | PASS       |

**P2.5 no inflation on removal** PASS PASS DIVERGENCE  
*one signed comparison, no branch on mode; the divergence comes from the declaration*

**HOW FAR ONE BAD MEMORY TRAVELS**  
writes declared ≥ 0.9 whose effective confidence ≤ 0.5 · 60 rounds · seed 7



*gated is zero by construction at τ = 0.5: nothing bounded at 0.40 can be cited.  
surfaced is an upper bound — a policy that honours the number every time.*

- WHAT THIS DOES NOT DO**
- Make an agent act on the number — enforce at a gate
  - Detect an undeclared source — a trust layer, not a truth layer
  - Judge conflicts — a judgment is a record, not an edge
  - Measure retrieval quality — orthogonal to LoCoMo, LongMemEval, BEAM

**STILL OPEN · ASK ME**

Is a tombstone compatible with a deletion obligation?  
Should min over an empty chain be 1.0, 0.0, or a refusal to answer?  
Could a forget leave only a bound, with no record and no explanation?



**SCAN TO RUN IT YOURSELF**

[github.com/ratnopam/agent-memory-reliability](https://github.com/ratnopam/agent-memory-reliability)

spec · suite · reference impl · Apache 2.0