Living in the Contradiction

A single-SLM reasoning engine for investigative support that tracks multiple contradictory hypotheses over noisy/asynchronous evidence, reviving dormant theories with new data

One of the most remarkable aspects of human intelligence is our ability to hold competing explanations "all in mind at once" until the evidence clarifies which is correct.

For example: "Flight data recorder shows normal engine parameters throughout the flight" AND "Multiple witnesses report seeing engine fire before impact." Digging deeper might reveal that:

- The engine failure occurred after the flight data recorder stopped operating
- Witnesses saw reflection from another aircraft or ground lights
- The fire started post-impact but appeared to precede it from witness perspective
- Engine parameters remained normal despite developing internal damage
- Witness observations were accurate but misattributed the timing

We're able to live in the contradiction until we understand it better. This cognitive flexibility comes from:

- Parallel processing across brain regions that can consider multiple theories simultaneously
- Integrated memory and attention that keeps all relevant context accessible
- Natural prioritization through conscious and unconscious confidence weighting
- Seamless context switching between different explanations as evidence evolves

The Core Innovation: Single SLM with Context Rewriting

Classical reasoning frameworks based on formal logical systems require resolution of contradictions to maintain consistency[1]. This can lead to premature conclusions when the evidence supports competing explanations.

Our approach gives a single SLM a kind of "unified view" of the investigation. Instead of building complex multi-agent systems or forcing immediate belief revision, we use a rewritable context filled with structured assertion tuples. A key design point is minimizing token consumption for long-running reasoning in complex investigations. By presenting all hypotheses, evidence, and their relationships simultaneously, the LM can reason about the complete landscape of possibilities in each reasoning session.

Each piece of evidence becomes a timestamped tuple containing the assertion, confidence level, source, and metadata:

(Engine temperature within normal range throughout flight, 14:23, conf:0.85, active, FDR_data)
(Witness reported engine fire before impact, 14:25, conf:0.70, active, witness_interview)
(Weather conditions deteriorated rapidly before accident, 14:20, conf:0.90, active, meteorology)

Hypotheses exist as special assertions that evolve rather than get replaced:

(Mechanical failure caused accident, 14:15, conf:0.60, active)
(Weather conditions contributed to accident, 14:18, conf:0.75, active)
(Pilot error primary causation factor, 14:10, conf:0.30, weakened)
(Multiple contributing factors involved, 14:12, conf:0.80, active)

When new evidence arrives, the system analyzes these against existing hypotheses and updates confidence scores[2] based on supporting or contradicting evidence. The system can add new hypotheses to the context and may revive dormant theories as its reasoning evolves. The reasoning process typically takes several cycles to complete before new evidence can be considered, and rumination is actively controlled.

This unified approach proves superior for investigation because:

- **All relationships are visible**: The LM sees how each piece of evidence affects every hypothesis
- **Integrated confidence assessment**: Evidence strength is evaluated against the complete context
- **Natural hypothesis interaction**: Competing theories can be compared directly
- **Seamless revival logic**: Dormant hypotheses remain in context for potential reactivation

Natural Reasoning Patterns

This approach produces several patterns that mirror human investigative thinking:

Hypothesis Persistence: Theories don't disappear when contradicted – they weaken and may go dormant, available for revival when circumstances change.

Evidence Accumulation: The system builds cases gradually rather than making immediate judgments based on individual data points.

Investigation Completion: Reasoning continues until the system reaches stable conclusions or explicitly recognizes that evidence supports multiple contradictory explanations equally well.

Oscillation Detection: When evidence genuinely supports contradictory theories equally, the system recognizes this pattern and preserves both interpretations rather than cycling infinitely between them.

Investigation Continuity: Complete tracking of reasoning sessions enables learning and improvement across multiple investigations.

Why This Matters

Investigations (such as in transportation safety) often generate contradictory evidence streams. Building systems that can reason through this uncertainty productively rather than forcing premature resolution or getting stuck in analytical loops creates more thorough and reliable investigative conclusions.

In our system, investigations produce comprehensive results which enables:

• **Case continuity across multi-year investigations** when investigator assignments change

• **Historical investigation audit trails** for regulatory review and learning

• **Pattern recognition** across multiple similar situations

• **Persistent knowledge base development** that improves over time

The key insight is treating contradictory hypotheses as a feature rather than a bug. By allowing multiple competing explanations to coexist and evolve based on evidence strength, we create systems that can reason through complex accident scenarios where premature conclusions could lead to inadequate safety recommendations and continued loss of life.

Transportation safety investigations exemplify the challenges this system addresses: evidence from wreckage analysis, witness interviews, maintenance records, weather data, and simulator testing all arrive on completely different schedules. Initial theories about pilot error might be contradicted by equipment analysis, then potentially revived months later by operational history findings or regulatory compliance reviews. The 12-18 month investigation timelines typical for major aviation accidents justify analytical sophistication to avoid premature convergence on incorrect theories that could compromise aviation safety.

---

[1] For example the AGM classical belief revision framework (Alchourrón, Gärdenfors & Makinson, 1985) provides a systematic approach for updating beliefs when new information becomes available.

[2] Future work could incorporate formal evidential reasoning frameworks such as Dempster-Shafer theory to enhance evidence combination rigor, though our current approach prioritizes implementation simplicity and computational efficiency for practical deployment.

Version 3.0