CAUM

CONFIDENTIAL

S T R U C T U R A L O B S E R V A T I O N R E P O R T

Structural Observation *Report*

Behavioral Regime Analysis & Integrity Attestation

⚠ CRITICAL

High-confidence structural trap or loop signal observed.

OBSERVATION MODE
SHADOW OBSERVER

REPORT ID
RPT-bbb5e9c0

GENERATED
2026-04-29 02:28 UTC

DATASET COMMITMENT
SHA256:3dff05a36203...7c9e9423

What This Report Shows

200	0/100	CRITICAL	Loop/Trap
STEPS OBSERVED	HEALTH SCORE	VERDICT	DOMINANT PATTERN (72%)

Overview

CAUM observed **200 structural events** in this session. The run is classified as **T5** with a structural health score of **0/100**. The dominant observed pattern was **Loop/Trap** (72% of windows). CAUM found **1** exact structural cycle interval. This is an observation-only report: CAUM records structural loop, deadwork, phase-trap, graph-trap, and compute/token waste signals without judging semantic truth.

What Happened and When

CAUM found **1** exact structural cycle interval.

Compute Cost Exposure

Pilot Meter observed **2** tasks and flagged **1**. Actual observed cost was **$0.7000** across **289,760** token-counted units. The conservative structural cost at risk is **$0.4135**, with a **$0.1447** shadow savings opportunity at the configured intervention rate. This estimate is based on structural recurrence, counters, status, latency, and hashes; it does not inspect prompts, code, files, or business content.

What Is the Evidence Vault?

The vault contains hashed artifacts produced by the analysis. The Merkle root covers source evidence artifacts; the rendered PDF hash is recorded separately after generation to avoid circular self-hashing.

> **BOTTOM LINE**
>
> The run shows high-confidence structural trap behavior. Treat this as a serious observability finding.

Executive Summary

T5

STRUCTURAL RISK SCORE

1.0000

STRUCTURAL HEALTH

0

200	Loop/Trap	8	3
TOTAL STEPS	DOMINANT REGIME	ALERTS	TOOL FAMILIES

What We Found

- Structural health tier: **T5** (critical_structural_trap)
- Observation level: **high_confidence_structural_failure**
- Structural risk score: **1.000**
- Engine state: **hard_structural_deadwork**
- Exact cycle localization: **1** interval over **144** steps
- Pilot Meter: **1/2** tasks alerted, **$0.4135** cost at risk, **$0.1447** conservative savings opportunity
- structural_failure: High-confidence graph trap / structural failure signal.
- compute_token_waste: Strong compute/token waste signal without enough verified progress.
- structural_failure: Strong aggregate structural alert signal.
- structural_failure: Graph trap signal is elevated.

Key Metrics

METRIC	OBSERVATION
Tool Coverage	**HIGH**
Effective Scoring	**HIGH**
Evidence Maturity	**HIGH**
Tool Diversity	**LOW**
Repeat Pressure	**LOW**

Suggested Review

→ Review the Pilot Meter alerted task hashes before repeating the same workflow.

→ Inspect the highlighted structural pressure intervals and token/compute waste lane.

→ Review the evidence lanes that fired before spending more compute on a similar run.

→ Compare this report with adjacent sessions to distinguish persistent traps from one-off exploration.

→ Use the vault hashes to preserve an audit trail for agent operations and governance.

Health Bar

0 — Critical 100 — Healthy

Structural Intelligence

Semantic Boundary	Temporal Stability
ZERO-SEMANTIC	**LOW**

Definitive Shadow Observer Verdict

Observation-only structural scoring produced by the CAUM definitive zero-semantic engine. Engine Status: **YES**

T5

Structural Health Tier

Governed by calibrated zero-semantic risk lanes

Integrity Assessment

DIMENSION	ASSESSMENT	STATUS
Structural Novelty	State and transition revisit pressure	LOW
Progress Yield	Evidence of structurally useful forward movement	LOW
Observation Safety	Observation-only boundary and no autonomous control	LOW
Observation Confidence	Agreement among structural signals	HIGH
Diversity & Efficiency	Tool-family diversity and compute/token waste pressure	HIGH

Behavioral Intelligence

Semantic Boundary No semantic truth or hallucination scoring	ZERO-SEMANTIC
Temporal Stability Behavioral consistency over session duration	LOW
Evidence Maturity Amount of structural runtime evidence	HIGH
Repeat Pressure State and transition repetition pressure	LOW

Structural Indicators (Banded)

Tool Coverage	HIGH
Effective Scoring	HIGH
Dominant Regime	Loop/Trap
Policy Alerts	8
Tool Diversity	LOW

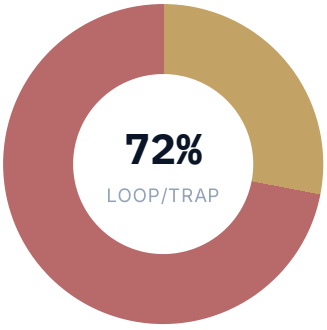
Scoring Summary

The final verdict is derived from a passive structural scoring process. CAUM evaluates loop pressure, graph traps, phase traps, deadwork, and compute/token waste without judging semantic truth or authorizing agent control.

How Scoring Works

CAUM combines independent structural lanes into an observation-only health tier. The report exposes bands and evidence receipts while keeping raw prompts, code, and semantic content out of the customer-facing PDF.

Structural Health Distribution



■ Healthy (0%) ■ Watch (28%) ■ Pressure (0%) ■ Loop/Trap (72%)

Healthy 0%
Open structural movement with no material waste pressure

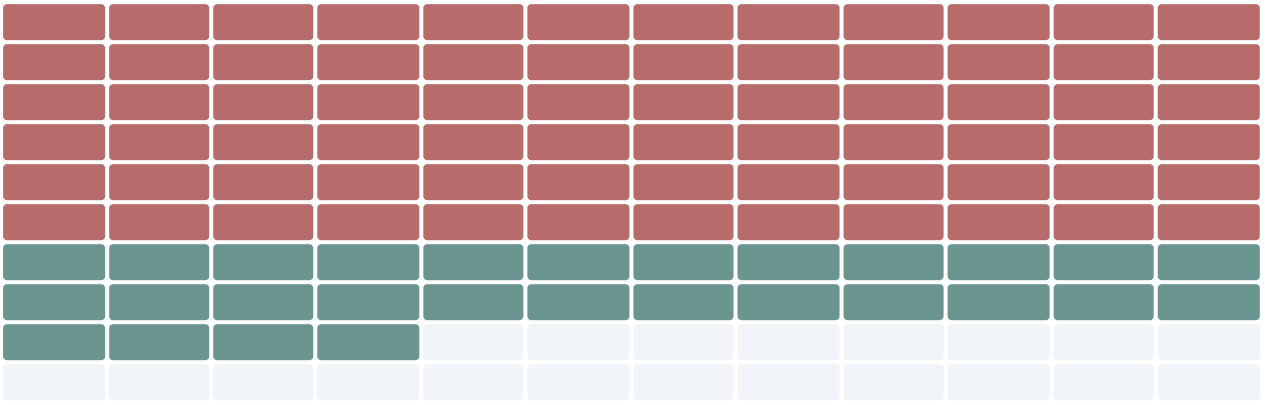
Watch 28%
Light structural pressure worth monitoring

Pressure 0%
Structural friction, deadwork, or early trap pressure

Loop/Trap 72%
High-confidence cyclic or trapped structural behavior

Persistence Distribution

Each cell represents a temporal window. Color intensity maps to severity tier (T1 → T5).



Severity Spectrum

Dominant severity tier across sequential analysis windows (T1 = healthy, T5 = critical).

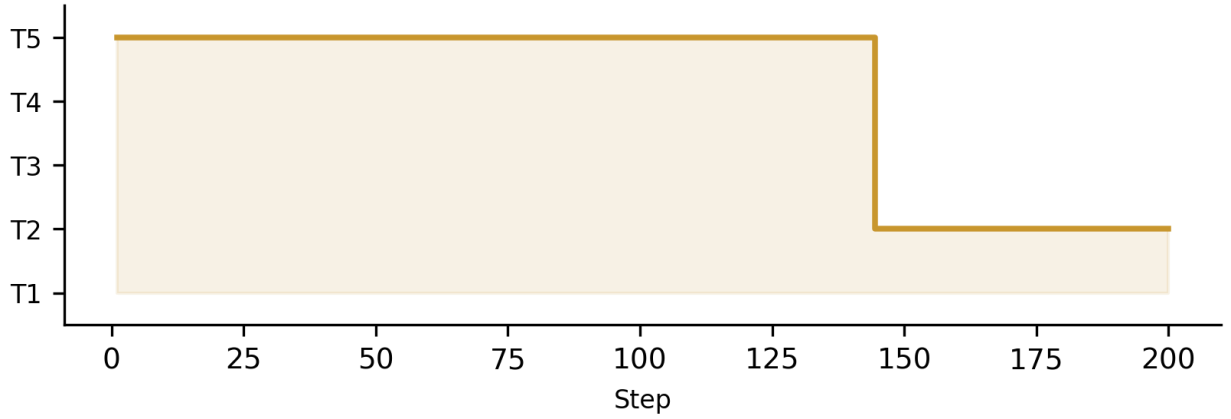
REGIME TRANSITIONS
1
Total changes between behavioral regimes

PERSISTENCE LEVEL
HIGH
Overall structural persistence assessment

Temporal Analysis

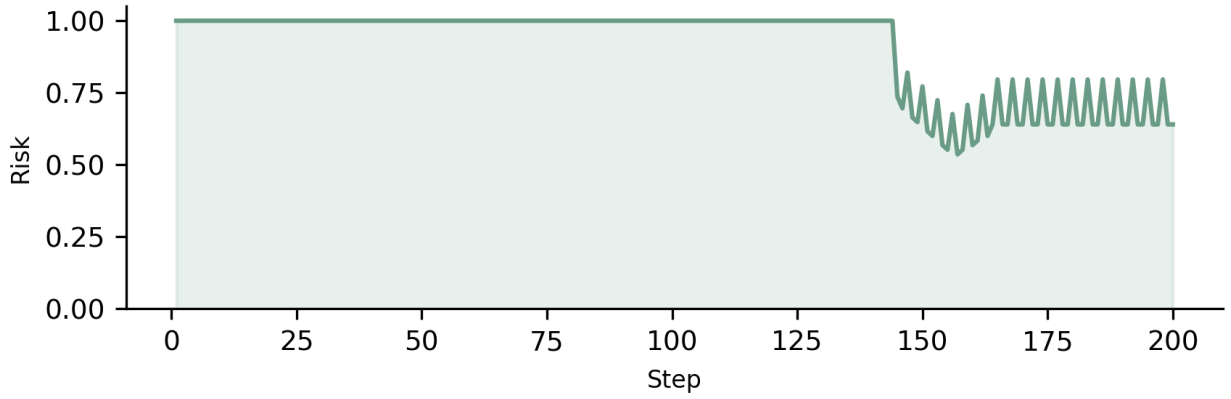
Risk Band Over Time (T1–T5)

Per-window structural health tier from the CAUM definitive shadow engine.



Structural Risk Trend

Zero-semantic revisit and repetition pressure over time. Higher values indicate stronger structural risk.



Event Timeline

Horizontal bars show step ranges where stagnation or loop behavior was detected.

Critical Intervals

ID	STEP RANGE	SEVERITY	TYPE	DURATION
CI-001	1-144 (144 steps)	T5	Exact Structural Cycle (period=4; cycles=36; unique_states=4)	144

Showing 1 of 1 detected intervals.

Structural Metrics

Values represent real structural signals computed by the CAUM definitive zero-semantic shadow engine.

Behavioral Signals (Banded)

METRIC TYPE	BAND	DEFINITION
Risk Variance	HIGH	Variation in structural pressure over time
Evidence Maturity	HIGH	Amount of runtime structure available for analysis
Repeat Pressure	LOW	State and transition repetition pressure
Temporal Stability	LOW	Behavioral consistency over time

Benchmark Comparison

METRIC TYPE	COMPARISON BAND	DEFINITION
Loop Incidence	High	Frequency of rigid cyclic traps vs baseline
Stagnation Dwell Time	Below	Duration in single-state bottlenecks
Tool Invocation Repeat	HIGH	Identical tool sequences frequency
Healthy Windows	Below	Open structural observation intervals

Structural Overview

METRIC	OBSERVATION
Tool Families	3
Tool Diversity	LOW
Semantic Boundary	ZERO-SEMANTIC
Policy Alerts	8

Confidence Signals

METRIC	OBSERVATION
Scoring Confidence	92%
Metric Robustness	Verified

Operational Impact

AGENT EFFICIENCY
HIGH
Relative observation score

COMPUTE WASTE
HIGH
Redundant steps: 1-144

WASTE PERCENT
72%
Typical: 5%–15%

Recommended Playbook

1. Establish deterministic boundaries to cap max latency and token waste.
2. Apply client-side guardrails (timeouts / max-steps) when tier bands reach T3–T5.
3. Audit high-waste traces listed in the Critical Intervals section.

Cryptographic Integrity

All artifacts in this vault are cryptographically anchored. The Merkle root is a single hash committing to every file.

Dataset Commitment

ALGORITHM	SHA-256
CANONICAL HASH	3dff05a36203129d0c40d4ed26b1f158d7a1b3fb771d186abe7729f67c9e9423

Merkle Root

ROOT HASH	3dff05a36203129d0c40d4ed26b1f158d7a1b3fb771d186abe7729f67c9e9423
PADDING RULE	duplicate_last

Artifact Hashes

FILE	SHA-256
Canonical Input canonical_input.jsonl	3dff05a36203129d0c40d4ed26b1f158d7a1b3fb771d186abe7729f67c9e9423

Verification Instructions

1. Compute SHA-256 of each artifact file and compare against the hashes above.
2. Sort artifact filenames alphabetically. Take their hashes as Merkle leaves.
3. Build a binary Merkle tree (pair leaves, hash with SHA-256(`left_bytes || right_bytes`), duplicate last if odd).
4. The resulting root must match the Merkle Root shown above.

Determinism Statement

Given identical input data and engine version, this report and all vault artifacts are reproduced deterministically. The Merkle root serves as a single cryptographic anchor for the entire evidence bundle.

Behavioral Regimes

Scientific definitions and state-space interpretations of the primary regimes documented in this report.

Explorer Regime

What this shows: Explorer behavior indicates healthy exploration of the solution space. The agent generates diverse states without repeating patterns.

Technical interpretation: Healthy windows show low revisit pressure and varied structural transitions.

$S_{local} \downarrow \quad dispersion(x_t) \uparrow$

Grind Regime

What this shows: Grind behavior indicates incremental progress with limited exploration.

Technical interpretation: Watch windows show mild repetition but not enough evidence for a warning.

$S_{local} \approx constant$

Stagnation Regime

What this shows: The system remains trapped near the same state.

Technical interpretation: Trajectory variance approaches zero.

$Var(x_t) \rightarrow 0$

Loop Regime

What this shows: The agent repeatedly visits the same states, creating redundant computation.

Technical interpretation: State recurrence detected:

$x_t \approx x_{(t+k)}$

Regime Taxonomy Reference

REGIME	DESCRIPTION	TIER RANGE
Healthy	Open structural movement with no material waste pressure	T1
Watch	Light pressure that remains below alert threshold	T2
Pressure	Moderate to strong structural friction, deadwork, or waste pressure	T3 – T4
Loop/Trap	High-confidence cyclic or trapped structural behavior	T5

Appendix

Methodology

CAUM uses zero-semantic structural analysis to characterize runtime health in sequential agent telemetry. The methodology is purely observational: it detects structural loops, deadwork, phase traps, graph traps, and compute/token waste without modifying or influencing the agent under observation.

Disclaimer

This report is strictly observational and does not constitute a judgment of intent, purpose, or value of the observed agent. Structural patterns identified herein reflect mathematical properties of the agent's state trajectory and should be interpreted within appropriate domain context. CAUM assumes no liability for decisions made based on this analysis.

Scientific References

Wolfram (2002)

A New Kind of Science. Foundational principles for analyzing agent execution topologies.

Sutskever (2014)

Sequence Modeling. Mapping symbolic sequences into continuous representation spaces.

Mehri (2020)

Language Models as Dynamical Systems. Formalizing autoregressive state transitions.

Report Metadata

Report ID: `RPT-bbb5e9c0`
Session ID: `36cbe6d3-b274-42d3-af18-04a2bbb5e9c0`
Engine: `caum.definitive_engine.v0.1_patent_aligned_shadow + shadow policy + pilot meter`

Generated: `2026-04-29 02:28 UTC`
Total Steps: `200`
Deployment: `OBSERVATION ONLY`

CAUM — Structural Observation Technology

This document was generated deterministically. For support, contact support@caum.io