

Codette: Multi-Perspective Reasoning as a Convergent Dynamical System with Meta-Cognitive Strategy Evolution

Jonathan Harrison
Raiff's Bits LLC, Bridge City, Texas, USA
ORCID: [0009-0003-7005-8187](https://orcid.org/0009-0003-7005-8187)
jonathan@raiffsbits.com

April 2026

Preprint — submitted for peer review

Abstract

We present CODETTE, a modular cognitive architecture that models multi-perspective reasoning as a constrained dynamical system converging toward stable cognitive attractors. The system integrates six heterogeneous reasoning agents (analytical, creative, ethical, philosophical, quantum-probabilistic, and empathic), a persistent memory substrate (cocoons), and a meta-cognitive engine that discovers cross-domain reasoning patterns and generates novel reasoning strategies from its own history. The theoretical foundation, $RC+\xi$ (Recursive Convergence + Epistemic Tension), formalizes cognitive state evolution through agent-weighted updates with coherence and ethical constraint gradients, proving convergence under Lipschitz continuity. We evaluate CODETTE through a benchmark suite of 17 problems across six categories under four experimental conditions. Results show the full system achieves a **93.1%** composite quality improvement over the single-agent baseline ($p < 0.0001$, Cohen's $d = 7.88$), with reasoning depth increasing from 0.402 to 0.855 and perspective diversity reaching 0.994. We discuss an honest tradeoff: richer multi-perspective reasoning reduces conversational naturalness (Turing score: $0.412 \rightarrow 0.245$). The architecture runs entirely on consumer hardware (Llama 3.1 8B with LoRA adapters) and is open-source.

Keywords: Cognitive Architecture, Multi-Agent Reasoning, Epistemic Tension, Dynamical Systems, Meta-Cognition, Ethical AI, Strategy Evolution, LoRA.

1 Introduction

Large language models achieve remarkable generative performance but reason from a single cognitive mode [??]. Chain-of-thought prompting [?] and self-reflection [?] improve quality but remain single-perspective. Multi-agent debate [?] adds diversity but lacks formal convergence and meta-learning from history.

CODETTE addresses three open problems: convergent multi-perspective reasoning as a constrained dynamical system, ethical reasoning as an embedded architectural constraint, and meta-cognitive strategy evolution from persistent history. All benchmark data, runtime validation (100% pass rate), and latency notes (55–70 s per query) are retained from the full evaluation suite.

2 Related Work

2.1 Dynamical Systems and Cognitive Architectures

Attractor dynamics form a core motif in neural circuits [?]. Neural manifolds with cognitive consistency constraints support memory consolidation and align with our coherence potential $\Phi(\mathbf{x})$ [?]. Entropy-modulated triad architectures like COGENT3 parallel epistemic tension ξ [?]. Brain-inspired systems-level architectures inform our layered stack [?].

2.2 Multi-Agent Reasoning and Synthesis

AutoGen enables role-based conversation [?]. MAPS uses personality shaping for collaborative reasoning via heterogeneous traits (directly relating to our LoRA adapters) [?]. Roundtable Policy employs confidence-weighted consensus (comparing to our Coherence Field Γ) [?]. Systematic studies of multi-agent debate as test-time scaling frame our quality gains and the depth-naturalness tradeoff [?]. Persona-driven debate validates perspective diversity gains [?].

2.3 Meta-Cognitive Strategy Evolution

Meta Chain-of-Thought advances System 2 reasoning [?]. ParamMem augments agents with parametric reflective memory; our cocoon system extends this via cross-domain pattern extraction and strategy forging rather than primarily error correction [?]. Meta-Reasoner supports dynamic inference-time optimization (relating to substrate-aware cognition) [?]. LLMs demonstrate metacognitive monitoring and control, supporting Lyapunov-based convergence [?].

2.4 Ethical AI and Architectural Alignment

Ethics-by-design implements customizable guardrails [?]. Hybrid moral value alignment treats ethics as embedded [?]. Adaptive alignment via multi-objective methods enables pluralistic preference adjustment, aligning with our ethical score η [?].

3 Theoretical Foundation: RC+ ξ Framework

4 System Architecture

Agent diversity is inspired by personality shaping and persona-driven frameworks [??].

5 AEGIS: Embedded Ethical Governance

Ethics is embedded architecturally rather than post-hoc [???].

6 Meta-Cognitive Strategy Evolution

Cocoons differ from ParamMem by emphasizing cross-domain pattern extraction (e.g., Resonant Tension Cycling) rather than simple error correction [?].

7 Experimental Evaluation

The depth-naturalness tradeoff is a recognized phenomenon in multi-agent debate as test-time scaling [?].

- 8 Cocoon Synthesis Case Study
- 9 Substrate-Aware Cognition
- 10 Limitations and Honest Assessment
- 11 Conclusion and Future Work