Cuttlefish AI: Abductive Reasoning Protocol

Overview

Cuttlefish AI integrates abductive reasoning as a foundational logic framework, inspired by Charles Sanders Peirce's philosophy of inquiry. This README outlines the `LeapEngine` protocol, which positions abductive reasoning as the core cognitive mechanism for our agentic infrastructure systems.

Core Principle: "Insight Before Certainty"

Cuttlefish AI begins not with known facts (deduction) or established trends (induction), but with informed speculation. Abduction is the leap from uncertain data to plausible hypotheses worth testing. This enables Cuttlefish agents to navigate ambiguity and simulate futures.

The LeapEngine Protocol

1. Abduction: Hypothesis Generation

Trigger: Observational anomaly or ambiguous system input
Agent Type: `InsightAgent`
Function: Generate candidate explanations, scenarios, or interventions. These are the "best guesses" used to seed simulations or policy design.

2. Deduction: Rule Application

Trigger: Known constraints, physical laws, regulatory bounds
Agent Type: `RuleAgent`
Function: Apply rule-based logic to narrow or filter hypotheses. Examples include zoning codes, infrastructure specs, and compliance requirements.

3. Induction: Pattern Recognition

Trigger: Time-series, geospatial, or behavioral data
Agent Type: `PatternAgent`
Function: Rank hypotheses based on statistical likelihood, past outcomes, or projected performance.

Swarm Integration

BuilderFactory

- Each `BuilderAgent` begins with abductive framing.
- Hypotheses are tested through simulated deployment environments.

- Memory states track the evolution of belief over simulation cycles.

Governance Engine

- DAO proposal scoring now includes "belief elasticity" as a metric.
- Proposals should simulate belief shifts over time, not just binary votes.

Prompt Engineering

Encourage abductive prompts: - "What might be causing this?" - "Simulate three plausible explanations." - "Test a counter-hypothesis against governance rules."

Developer Notes

- Integrate `LeapEngine` into agent architecture via `Swarm/LeapEngine/abductive_core.py`
- Logging format: `abductive_trace_[agent_id].json`
- Visual overlay: Use TrustGraph to map hypothesis lineage

Philosophy in Practice

Peirce's model of belief formation emphasizes the role of *doubt* in thinking. Cuttlefish agents are designed to question assumptions, propose alternatives, and adapt their models of the world dynamically.

> "We're not designing for truth. We're designing for beliefs that evolve."

Use this protocol to train agents who speculate with care and test their own assumptions.

Tagline

Cuttlefish thinks like Peirce: not by knowing, but by wondering well.