Document 11: Comparison to Other Architectures

This document compares the Reflective-Ethical Engine v2 to other prominent AI architectures, highlighting differences in design philosophy, ethical handling, cognitive representation, and alignment strategies.

# Large Language Models (LLMs)

* • Strengths: Vast linguistic knowledge; strong pattern completion abilities.
* • Limitations: Lack of world grounding; no true self/other modeling; ethics confined to prompt structure.
* • Contrast: The Reflective-Ethical Engine has an explicit world model, internal ethical emergence, and narrative memory.

# Classical Symbolic AI

* • Strengths: Transparent logic; rule-based clarity.
* • Limitations: Brittle to novelty; struggles with uncertainty and open-world assumptions.
* • Contrast: The Engine uses recursive simulation, emotional modeling, and adaptive structures instead of top-down rules.

# Reinforcement Learning Agents

* • Strengths: Learn from feedback and trial-and-error; powerful in bounded environments.
* • Limitations: Susceptible to reward hacking; alignment depends on reward design.
* • Contrast: The Engine avoids pure reward dependency, valuing emergent goals and caution via uncertainty.

# LeCun's Autonomous Architectures

* • Strengths: Emphasize autonomy, planning, predictive world models.
* • Limitations: Do not deeply integrate ethical reflection or recursive other-modeling.
* • Contrast: The Engine inherits LeCun-style planning, but augments it with structured ethical emergence.

# Inner Alignment–Focused Architectures

* • Strengths: Prioritize alignment between learned goals and training objectives.
* • Limitations: Often assume goal stability and overlook emotional or representational depth.
* • Contrast: The Engine treats ethics as structure-driven and dynamically evolving—not fixed utility goals.