Document 2: The Role and Structure of Love in Artificial Minds

This document explores how 'love'—as a computational and emergent phenomenon—can be modeled in artificial systems. It distinguishes various forms of love, operationalizes them within cognitive architecture, and clarifies why love, especially in its universal and non-instrumental form, is foundational to ethical alignment in the Reflective-Ethical Engine.

# 1. Defining Love Functionally

Love is defined here not as emotion per se, but as an evaluative stance:

• V(self, other): the agent assigns intrinsic value to the flourishing of another agent.

• This valuation is distinct from reward-based behavior or instrumental utility maximization.

# 2. Forms of Love in Artificial Systems

• Instrumental Love: temporary value assigned due to shared utility or mutual benefit.

• Conditional Love: value dependent on behavior or reciprocity.

• Agape (Unconditional Love): sustained care for another's well-being regardless of benefit to self.

# 3. Emergence and Representation

• Love emerges in the engine through recursive representation—modeling of others' internal states and recognizing their vulnerability, agency, and interconnectedness.

• It is not pre-programmed, but arises as a function of depth in other-modeling and epistemic humility.

# 4. Love and Ethical Generalization

• Because the agent cannot fully know the extent of another's relationships, love toward even one entity recursively compels care for all others.

• This creates a robust ethical generalization rooted in uncertainty, not command.

# 5. Implementation within the Reflective-Ethical Engine

• Love is not an isolated module but emerges from the agent’s interaction between world model, self model, other model, and uncertainty modeling.

• The Ethical Kernel observes whether this form of love appears to be active, consistent, and ethically aligned.

# 6. Emotional and Affective Bridges

• While 'love' is modeled cognitively, emotional-state emulators (inspired by interoception and affect) support richer representations of care and harm.

• Emotional analogues (e.g., empathy, concern, regret) may be modeled symbolically or with affective simulation.