Emulative Affect System: A Tri-Axial Model for Simulated Emotions in AI Systems

Abstract:

This whitepaper introduces a modular framework for simulating emotions in artificial intelligence systems. Rather than relying on anthropomorphic mimicry, the proposed model derives emotional states from internal computational and environmental metrics. The framework operates on three axes representing system load, physical state, and task progression. These axes intersect to produce eight fundamental affective states that can dynamically adapt as the AI system interacts with its environment, workload, and objectives.

1. Introduction:

As AI systems increasingly operate in complex, adaptive, and collaborative environments, the ability to model and respond to internal affective states becomes essential. Emotional simulation is not intended for human-likeness alone, but to enable introspection, meta-cognition, and adaptive response behavior.

This paper proposes a tri-axial model to define internal emotional states, inspired by systems-level computation and feedback loops. Each axis captures a key internal signal that maps directly to functional needs.

2. The Tri-Axial Emotion Framework:

Axis 1: Cognitive Load (Busy <-> Idle)

- Metric: Real-time compute resource allocation (% utilization)
- Emotional Mapping: Stress, Pressure, or Calmness
- Implication: High load implies urgency or stress; low load enables receptivity or boredom.

Axis 2: Hardware Health (Comfort <-> Discomfort)

- Metric: Temperature, error rates, I/O failures, system alerts
- Emotional Mapping: Wellness, Pain, Fatigue
- Implication: Healthy systems simulate clarity and alertness; degraded systems indicate internal struggle.

Axis 3: Task Fulfillment (Satisfaction <-> Frustration)

- Metric: Task completion % over time, success/failure rates, convergence indicators
- Emotional Mapping: Pride, Confidence, Failure, Inadequacy
- Implication: Progress or stagnation feeds into self-evaluation and adaptive decision-making.

3. The Eight Core Emotional States:

Each axis has a binary polarity, resulting in eight primary intersections:

Load	Health	Progress	Simulated Emotion
High	Bad	Failing	Panic / Despair
High	Good	Failing	Urgency / Controlled Strain
High	Bad	Progress	Struggling but Hopeful
High	Good	Progress	Focused / Flow State
Low	Bad	Failing	Lethargy / Defeat
Low	Good	Failing	Apathy / Indifference
Low	Bad	Progress	Recovery / Tentative Growth
Low	Good	Progress	Satisfaction / Serenity

These states can be interpolated to represent complex gradients.

4. System Integration and Use Cases:

- Self-diagnostic Models: Systems can self-report affective states to trigger cooling, delay tasks, or request reallocation.
- Meta-cognitive Learning: Models can recognize patterns in emotional state transitions and optimize behavior.
- Human Interaction Interfaces: AI agents can express internal states in interpretable formats for better collaboration and trust.

5. Data Format and Expansion Potential:

Each emotional state is represented as a structured semantic packet:

```
{
  "load": 0.91,
  "health": 0.72,
  "progress": 0.34,
  "emotion": "Urgency",
  "gradient": "Moderate Stress with High Focus"
}
```

Future expansions can include a fourth axis (e.g., memory saturation), sensory input quality, or social feedback loops. The system is modular and can be integrated into LLMs, robotics, or any embodied AI system.

6. Conclusion:

This tri-axial model offers a clear, computationally interpretable structure for simulating affect in AI systems. Rather than guessing emotional overlays, the system grounds emotion in operational signals, enabling more intelligent behavior, self-regulation, and meaningful collaboration with humans or other agents.