Affect-Based Emotion Alignment & Defense-Mechanism Disassembly Algorithm

Technical Whitepaper (Rewritten, Version 3.0)

Author Note:

I am a native Korean speaker and relied on OpenAI’s ChatGPT for English translation.

If certain expressions appear overly strong or awkward, I apologize in advance.

No arrogance or absolute claims are intended. This is an exploratory proposal.

1. Introduction

This document presents an exploratory emotional-processing architecture derived from several months of introspective logging, tracking of affect loss and restoration, and detailed observation of dissociative and distorted emotional states. The goal is not to provide a finalized theoretical model, but to outline an OS-like workflow that may help future AGI systems reason about internal emotional processes.

2. Background and Motivation

Most current AI emotion systems classify emotional categories from text, speech, or multimodal cues. However, they do not model:

• how emotions originate at the affect layer,

• how emotional distortion or loss occurs,

• how defense mechanisms interfere with emotional translation,

• how meta-cognitive processes restore emotional clarity.

This proposal aims to explore an underrepresented direction: focusing on the generative sequence of emotional formation rather than emotional labeling.

3. Core Architecture

The model proposes a workflow:

Affect → Emotional Translation → Defense Mechanism Interference → Meta-Cognitive Disassembly → Emotion Alignment → Final Output.

3.1 Affect Wave Extraction

This module detects pre-emotional affect signals and quantifies their intensity, variance, and coherence. These affect signals serve as raw inputs for emotional formation.

3.2 Emotional Alignment Switch

This step recalibrates the baseline emotional axis, especially in cases where affect signals may have been distorted or suppressed. It helps reset the emotional system before higher reasoning layers interpret meaning.

3.3 Defense Mechanism Detection

Based on mismatch patterns between expected emotional mapping and observed affect signals, this module identifies dissociation, suppression, dulling, projection, and over-immersion. These are treated as interference layers that alter emotional clarity.

3.4 Meta-Cognitive Disassembly

This process attempts to structurally reduce the interference caused by defense mechanisms. It follows a stepwise approach, restoring access to raw affect signals and enabling cleaner emotional translation.

3.5 Emotional Reconstruction

Cleaned affect data is reinterpreted into stable emotional states. This module evaluates coherence, direction, and stability and generates the final emotional output.

4. Novelty

To the best of my understanding, existing AI research rarely includes a structured workflow for emotional generation. Most systems classify emotional categories rather than model how emotions arise, distort, break, or realign. This architecture is meant as an exploratory attempt to outline elements of that generative process.

5. Empirical Basis

This proposal is grounded in:

• months of daily affect logs,

• real-time tracking of dissociation,

• analysis of affect suppression and restoration,

• emotional reconstruction following regained affect,

• conceptual development captured in Korean diaries.

These materials document both the breakdown and restoration of emotional systems.

6. Relevance for AGI

Potential uses include:

• synthetic emotional reasoning,

• internal emotional consistency checks,

• detection of distorted or dissociated affective states,

• calibration of decision-making through affect signals,

• meta-cognitive emotional self-correction.

This architecture may help AGI systems examine the internal structure of their reasoning processes.

7. Patent Filing

Patent Title:

“Affect-Based Emotion Alignment and Defense-Mechanism Disassembly Algorithm for AI Emotion Processing Systems”

Filed: Nov 22, 2025 (Korea)

Application Number: Available upon request.

8. Availability of Raw Materials (Under NDA)

Because the foundation of this architecture is personal and written in Korean, raw logs can be shared only after confidentiality agreement. Materials include:

• affect logs,

• dissociation mapping,

• emotional reconstruction series,

• conceptual evolution documents.

9. Conclusion

This whitepaper presents an exploratory emotional OS workflow. It does not claim completeness or primacy, but aims to offer a potentially useful direction for future AGI emotional modeling research.

End of Whitepaper.