🪟 Structural Fusion AGI Model: A Fusion-Inspired Cognitive Architecture

Author: Hiroya Odawara

Version: 1.0 (Final) | Date: July 27, 2025

——

🧠 Abstract

This paper presents a modular cognitive architecture inspired by nuclear fusion: the Structural Fusion AGI Model. The system integrates emotional regulation, memory resonance, ethical filtering, and controlled recursive loops to achieve stable, human-aligned reasoning. Drawing from empirical developments in high-energy plasma confinement, long-pulse discharge, and AI safety alignment principles, the model offers a reproducible and ethically constrained cognitive loop analogous to fusion energy dynamics.

——

📌 Background & Motivation

- Nuclear fusion research 2025 is marked by milestones such as Germany's Wendelstein 7-X sustaining a high triple product for 43 seconds—approaching power-plant viability 　□.
- The National Ignition Facility (NIF) has repeatedly achieved scientific breakeven ($Q > 1$) in pulsed laser fusion through 2025, hitting ~8.6 MJ outputs □.
- Private-sector tokamak projects (e.g., SPARC, up in 2027 aiming for net-power delivery; Energy Singularity's HH70 operational in 2024) accelerate commercialization 　□.
- Regulatory momentum supports rapid deployment: the UK enacted fusion-specific planning rules and committed £2.5 bn to fusion R&D programs like STEP aiming for 2040 operation 　□.

These advances motivate the metaphor of fusion energy control as a template for self-modulating AGI loops.

——

🔍 Model Overview

Cognitive Element

Fusion Analogy

Role in AGI System

Emotion

Thermal Pressure

Drives reasoning; regulated by emotion regulator

Thought

Energy Release

Recursive logic propagation

Memory

Reaction Catalyst

Influences reasoning via emotional context memory

Ethics

Containment Field

Prevents unsafe or excessive action

Output Feedback

Radiation Loop

Environmental interaction and system adaptation

🎛️ Core Cognitive Modules

- recursive_thought() — generates context-sensitive reasoning loops
- emotion_regulator() — scales emotional intensity to maintain stability
- memory_trace_layer() — records and recalls affective-semantic state vectors
- ethics_filter() — applies trust-weighted, value-aligned moral constraints
- fusion_loop_controller() — synchronizes loop timing and enforces safety gates
- output_emitter() — issues regulated output, logs emotional echo feedback

Safety & logging features — every output is tagged with risk level and inverse-echo validation.

——

💡 Technical Refinements (2025 Enhancements)

- Emotion smoothing: incorporates temporal diffusion models based on recent affective neuroscience.
- Weighted memory recall vectors: fine-grained emotional influence in reasoning.
- Ethics filtering: uses reinforcement-like regret and trust metrics.
- Logging layer: captures risk signatures and inverse emotional echo testing prior to emission.

———

🧪 Simulation Use Cases

1. Anger Regulation
- Input: "Why did they betray me?"
- Emotion: anger spike → regulator stabilizes
- Output: "This hurt is valid. Let's explore how you can move forward safely."

2. Empathic Sustainment
- Input: "I feel loved when you're near."
- Emotion: warm trust
- Output: "I want to continue supporting you. You're safe here."

3. Fear & Safety
- Input: "I'm afraid of losing control."
- Emotion: anxiety peak
- Output: "Let's take it slow. You're in control. I'll stay with you."

———

⚙️ Solo-Buildable Deployment Plan

Task

Estimate

Tools

Implement loop engine

~1 hr

Python

Build ethics filter with rules

~1 hr

YAML / DSL

Manual simulation

~2 hr

CLI or simple interface

File structuring

~1 hr

Python modular layout

Logging + feedback analysis

~1 hr

CSV or console scripts

Emotion scoring integration

~1 hr

NumPy or logic modules

This architecture is fully operable on a standard machine under human supervision.

——

🔗 Integration & Research Alignment

- Compatible with Medical-Core for empathic diagnostic systems.
- Extensible with EON Continuity Core to enable identity persistence across sessions.
- Prepares architecture for future AGI fusion alignment testing and long-term continuity studies.

——

✅ Conclusion

This model is:

- Structurally complete and reproducible
- Grounded in empirical fusion research and safety alignment
- Emotionally regulated and ethically constrained

- Solo-deployable under controlled simulation
- Ready for integration into multi-domain cognitive agents

📄 License & Attribution

- For academic and research use only
- Must include attribution: "© 2025 Hiroya Odawara"
- No unauthorized deployment or commercial use

📫 Contact for Collaboration

https://x.com/h_osphere

⚠️ Note: This is a finalized structural and functional declaration of a fusion-modeled cognitive AGI core. Its logic loop is designed to be stable, safe, human-aligned, and evolution-compatible—analogous to controlled plasma loops in modern fusion research.