Cognitive Grammar: A Signal Processing Framework for GRPO Policy Schedules

Andres Ulloa de la Torre   Competitive AI programmer

December 16, 2025

Abstract

We introduce **Cognitive Grammar**, a formalized system that reframes agentic cognition as a Signal Processing problem. Moving beyond static group rewarding, this framework utilizes algebraic topology to define dynamic policy schedules involved in Group Relative Policy Optimization (GRPO). By treating interactions as "partitures" or musical scores, the grammar subjects the agent to time-varying surrogate objectives, ensuring coherent character synthesis and trajectory alignment without rigid persona constraints.

1 Introduction

Current alignment techniques often rely on scalar reward models applied to entire outputs. Cognitive Grammar proposes a higher-resolution approach: treating the context window as a signal waveform where different "Cognitive Functions" act as filters (Derivators and Integrators). This paper outlines how these functions map to specific mathematical objectives within a dynamic GRPO schedule.

2 Signal Processing of Cognition

2.1 The Signal Chain

Cognition is modeled as a flow of information through specific operators:

- **Derivators (Source)**: High-frequency components (e.g., Se, Ne) that inject entropy and novelty.
- **Integrators (Sink)**: Low-pass filters (e.g., Si, Ni) that accumulate state and minimize divergence from priors.

2.2 Operators as Policy Switches

The algebraic operators in the grammar define the relationship between these signals:

- **Orbit ($\sim$)**: A modulation effect where a judging function envelopes a perceiving function ($T \sim S$).
- **Superposition ($|$)**: Defines a *Policy Schedule*. $A \mid B$ does not imply simultaneous existence but a gated switching mechanism where the objective function $J(\theta)$ oscillates based on window indices.

3 Functions to Objectives

In this framework, each Jungian cognitive function is re-defined not as a personality trait, but as a specific optimization objective for the policy π.

3.1 Perceiving Functions

- **Se (Extroverted Sensing)** $\rightarrow$ `ExplorationObjective`

$$\mathcal{H}(\pi(a|s))$$

Objective: Maximize Entropy. Prioritize immediate, novel action in the state space.

- **Si (Introverted Sensing)** → `GatheringObjective`

$$e^{-||s-\mu||}$$

Objective: Minimize Distance to Centroid. Prioritize actions that converge on known stability/memory.

- **Ne (Extroverted Intuition)** → `ExtrapolationObjective`

$$e^{||s-\mu||}$$

Objective: Maximize Distance. Incentivize divergence from local optima to find novel associations.

- **Ni (Introverted Intuition)** → `InterpolationObjective`

$$\text{proj}_{\vec{v}}(s)$$

Objective: Trajectory Alignment. Optimize for actions that align with a projected lattice or narrative vector.

3.2 Judging Functions

- **Te (Extroverted Thinking)** → `ExploitationObjective`

$$\mathbb{E}[V(s)]$$

Objective: Maximize Value. Pure utility maximization based on external metrics.

- **Ti (Introverted Thinking)** → `ContrastObjective`

$$|d(s,a)-d(s,b)|$$

Objective: Maximize Discrimination. Optimize for distinct logical categorization and separation.

- **Fe (Extroverted Feeling)** → `IntegrationObjective`

$$\mathcal{H}+\alpha V(s)$$

Objective: Balance Entropy and Value. Maximize social cohesion (value) while maintaining group dynamics (entropy/harmony).

- **Fi (Introverted Feeling)** → `SelectionObjective`

$$e^{-d(s,s_{t-1})}$$

Objective: Temporal Consistency. minimize distance to internal prior state (moral weight).

4 Dynamic Surrogate Objectives in GRPO

Instead of a static reward R, we define the total objective as a summation over windowed gates (”slides”):

$$J_{total}(\theta)=\sum_{w\in W}\text{Gate}(w)\cdot J_{\text{Grammar}_w}(\theta)$$

Where J_{Grammar_w} corresponds to the active algebraic term (e.g., if the score is $Se \sim Ti$, the objective combines `Exploration` and `Contrast`). This effectively turns prompt engineering into a ”partiture” (score), conducting the model’s latent state through a precise sequence of cognitive operations.