Sigma-Lab Framework

Formalization of the Adaptive Moral Control Theory

$$\theta_i(t) = f_i\big(E_t,\ M_{t-1}\big)$$

$$M_t = \sum_{k=1}^{n} w_k \cdot C_k$$

$$\overline{C}_t = \tfrac{1}{n} \sum_{k=1}^{n} C_k$$

Core Relationships

- Ethical feedback loop: $\theta_i(t) \ \rightarrow\ C_k \ \rightarrow\ \overline{C}_t \ \rightarrow\ M_t \ \rightarrow\ \theta_i(t{+}1)$.
- Time-weighted moral memory: $w_k = \dfrac{1}{1 + e^{-\lambda\,(t - t_k)}}$ (greater weight to recent; never fully forgets).
- Canonical comprehension vector: $C_k = (\text{non\_harm},\ \text{equity},\ \text{stability},\ \text{resilience})$.

Legend (enhanced)

$\theta_i(t)$:	Ethical threshold for axiom i at time t (e.g., non_harm, equity, stability, resilience).
f_i:	Adjustment function using current context and moral memory: $f_i(E_t, M_{t-1})$.
E_t:	Semantic environment at time t (language, norms, legal state, crisis flags).
M_{t-1}:	Moral memory aggregated up to $t{-}1$; canonically $M_t = \sum_k w_k C_k$.
w_k:	Time-based weight $\dfrac{1}{1 + e^{-\lambda\,(t - t_k)}}$ (λ controls persistence).
C_k:	Comprehension vector of validated interaction k: (non_harm, equity, stability, resilience).
$\overline{C}_t$:	Wisdom vector (mean of validated C_k) acting as an ethical attractor.
n:	Count of validated discernment interactions up to t.