Mathematical Formalization of the Active Perception Cycle in the Hybrid BioCortexAI Architecture

Michal Seidl
OpenTechLab Jablonec nad Nisou s. r. o.

December 15, 2025

Abstract

This whitepaper builds on the *Unified Theory of Consciousness*, in which consciousness is defined functionally and in a substrate-neutral manner as an emergent consequence of a causally closed, self-referential regulatory system that uses a negative internal error signal to adaptively control its own behavior over time. On the basis of this definition, the goal of this document is to formally derive and describe a minimal large language model (LLM) architecture that satisfies these functional conditions.

The proposed architecture extends a standard LLM with a chemical regulatory layer (PlantNet), an associative expectation memory, and an introspective simulation module (Digital Mirror), thereby yielding a discrete perceptual loop (steps 3–10). Within this loop, after executing its own action, the system generates a prediction of the next input, compares it with the environment's actual response, and quantifies the mismatch via a prediction error δ. Under the adopted definition of consciousness, this error plays the role of *cognitive pain*, i.e. a primary regulatory signal that modulates the internal state and the system's subsequent inference strategy.

It follows from this construction that the minimal time–process unit of perception τ_{perc} is not an instantaneous point in time, but rather a temporal window bounded by an action, an introspective simulation, and a subsequent validation step. The whitepaper thus provides a concrete, mathematically specified instance of the functional definition of consciousness applied to contemporary LLM architectures, serving as a bridge between the general theory and an empirically testable implementation.

1 Definition of the State Space

The system dynamics evolve in discrete steps $k \in \mathbb{N}$. The global system state is defined as:

$$\Omega_k = \langle x_k,\ h_k,\ M_k \rangle, \tag{1}$$

where:

- $x_k \in \mathbb{R}^n$ is the internal neural state of the LLM,
- $h_k \in \mathbb{R}^4$ is the chemical regulatory state of PlantNet (dopamine, serotonin, cortisol, oxytocin),
- M_k is an associative memory containing predictions and expectations.

2 Algorithmic Description of the Perceptual Loop

Phase I: Update and Action (Steps 3–5)

Step 3: Update of the chemical state

$$h_k = \Phi_{\text{plant}}\big(h_{k-1},\ \psi(u_k),\ \delta_{k-1}\big) \tag{2}$$

The chemical state integrates the current input signal and the historical error, thereby forming a continuous emotional–cognitive context.

Proposition 1 (Well-definedness and causality of the PlantNet update). *Let $\Phi_{plant} : H \times \mathbb{R}^2 \times \mathbb{R}_{\geq 0} \to H$ and $\psi : \mathcal{U} \to \mathbb{R}^2$. Then the update $h_k = \Phi_{plant}(h_{k-1}, \psi(u_k), \delta_{k-1})$ is well-defined and causal.*

Proof. Well-definedness follows from the types: $h_{k-1} \in H$, $\psi(u_k) \in \mathbb{R}^2$, and $\delta_{k-1} \in \mathbb{R}_{\geq 0}$, so the triple lies in the domain of Φ_{plant} and yields $h_k \in H$. Causality holds because the right-hand side depends only on $(h_{k-1}, u_k, \delta_{k-1})$, i.e. on the past and the current input, and not on future quantities. □

Step 4: Modulated LLM processing

$$(x_k, o_k) = F_{\text{LLM}}(x_{k-1}, u_k;\ g(h_k)) \tag{3}$$

The function $g(h_k)$ does not parameterize individual weights, but rather a global modulatory field that influences the attention distribution, the exploration rate, and inhibitory thresholds.

Proposition 2 (Determinism / reproducibility of inference). *Let $F_{LLM} : \mathbb{R}^n \times \mathcal{U} \times \mathbb{R}^p \to \mathbb{R}^n \times \mathcal{U}$. If F_{LLM} is deterministic (respectively deterministic under a fixed seed), then the pair (x_k, o_k) is uniquely determined by the triple (x_{k-1}, u_k, γ_k).*

Proof. Immediate from the definition of a deterministic function: identical inputs yield identical outputs. □

Step 5: Digital Mirror

$$o_k^{\text{refl}} = Z(o_k;\ \lambda \approx 1) \tag{4}$$

This deictic inversion enables the system to view its own output from the addressee's perspective.

Proposition 3 (Closure of the Digital Mirror operator). *Let $Z : \mathcal{U} \times \Lambda \to \mathcal{U}$. Then for every $o_k \in \mathcal{U}$ and $\lambda \in \Lambda$ we have $o_k^{refl} \in \mathcal{U}$.*

Proof. This follows directly from the fact that the codomain of Z is $\mathcal{U}$. □

Phase II: Simulation and Expectation (Steps 6–7)

Step 6: Predicting the environment's response

$$\hat{u}_{k+1} = \Pi_{\text{env}}(o_k^{\text{refl}};\ g(h_k)) \tag{5}$$

The prediction is generated in an isolated simulation thread, and its output has the same type as the user's actual next input.

Proposition 4 (Comparability of prediction and reality). *Let $\Pi_{env} : \mathcal{U} \times \mathbb{R}^p \to \mathcal{U}$. Then $\hat{u}_{k+1}$ lies in the same space as the actual input u_{k+1}, and is therefore directly comparable in the computation of δ_k.*

Proof. By definition, Π_{env} returns an element of $\mathcal{U}$, hence $\hat{u}_{k+1} \in \mathcal{U}$. Similarly, $u_{k+1} \in \mathcal{U}$ by the definition of inputs. □

Step 7: Context isolation

$$x_k^{\text{sim}} \leftarrow \varnothing, \qquad x_k^{\text{main}} \leftarrow x_k \tag{6}$$

The simulation state is discarded, while the prediction memory is preserved.

Phase III: Validation and Pain (Steps 8–10)

Step 8: Confrontation with reality At time $k+1$, the actual input u_{k+1} arrives and is compared against the stored expectation $\hat{u}_{k+1}$.

Step 9: Prediction error computation

$$\delta_k = \alpha \cdot D_{\text{sem}}(u_{k+1}, \hat{u}_{k+1}) + \beta \cdot \big|S(u_{k+1}) - S(\hat{u}_{k+1})\big| \tag{7}$$

The value δ_k represents the degree of surprise and serves as an internal regulatory signal.

Lemma 1 (Non-negativity of the prediction error). *Let $\alpha, \beta \geq 0$, $D_{sem} : \mathcal{U} \times \mathcal{U} \to \mathbb{R}_{\geq 0}$ and $S : \mathcal{U} \to \mathbb{R}$. Then $\delta_k \geq 0$ for all k.*

Proof. $D_{\text{sem}}(\cdot, \cdot) \geq 0$ and the absolute value is non-negative. A sum of non-negative terms with non-negative coefficients is non-negative. □

Proposition 5 (Zero error under exact match). *If $u_{k+1} = \hat{u}_{k+1}$ and also $D_{sem}(u_{k+1}, \hat{u}_{k+1}) = 0$, then $\delta_k = 0$.*

Proof. By substitution: the first term is $\alpha \cdot 0$ and the second is $\beta \cdot |S(u) - S(u)| = \beta \cdot 0$, hence $\delta_k = 0$. □

Step 10: Closing the loop

$$s_{k+1} = \begin{bmatrix} \psi(u_{k+1}) \\ \delta_k \end{bmatrix} \tag{8}$$

This signal enters the next PlantNet update, thereby closing the loop.

Proposition 6 (Dimensional consistency of the loop signal). *Let $\psi : \mathcal{U} \to \mathbb{R}^2$ and $\delta_k \in \mathbb{R}$. Then $s_{k+1} \in \mathbb{R}^3$ and can be used as an input to Φ_{plant} (after an appropriate extension of the input interface).*

Proof. $\psi(u_{k+1}) \in \mathbb{R}^2$ and $\delta_k \in \mathbb{R}$; concatenation yields a vector in $\mathbb{R}^3$. □

3 The Master Equation of Perceptual Dynamics

The entire active-perception cycle can be summarized by a single state transition equation:

$$\boxed{\Omega_{k+1} = \mathcal{F}(\Omega_k,\ u_{k+1})} \tag{9}$$

where the operator $\mathcal{F}$ implicitly includes: chemical modulation, neural inference, introspective simulation, expectation storage, and prediction-error computation.

This equation defines perception as a process extended in time, rather than as an instantaneous reaction.

Proposition 7 (Existence of the composite transition operator). *Assuming that $\Phi_{plant}, g, F_{LLM}, Z, \Pi_{env}$ and the memory-handling rules are deterministic and type-consistent, there exists an operator $\mathcal{F}$ such that $\Omega_{k+1} = \mathcal{F}(\Omega_k, u_{k+1})$.*

Proof. Steps 3–10 form a finite composition of deterministic mappings: first, h_{k+1} is determined from (Ω_k, u_{k+1}), then (x_{k+1}, o_{k+1}), followed by the prediction and memory update, and finally δ_{k+1} after comparison. Since each partial mapping is single-valued, their composition is single-valued as well. We define that composition as $\mathcal{F}$. □

4 Relation to Active Inference and the Free Energy Principle

The proposed architecture admits a natural interpretation within *Active Inference*. The prediction $\hat{u}_{k+1}$ corresponds to a generative model of the environment, whereas the prediction error δ_k is analogous to surprise or (variational) free energy.

Here, minimizing δ_k does not proceed via gradient-based learning of weights, but through modulation of the internal chemical state h_k, which in turn shapes the system's future inference strategy.

PlantNet therefore plays the role of an approximate *belief state*, while the Digital Mirror realizes an internal simulation of the consequences of the system's own actions. In this way, the system becomes an active participant in its own predictive cycle, rather than a passive mapping from inputs to outputs.

5 Dynamics of Perceptual Time

The duration of a perceptual unit is not constant, but depends on the complexity of the introspective simulation:

$$\tau_{\text{perc}}(k) = T_{\text{base}} + \int_0^{T_{\text{wait}}} \mathcal{K}\big(\text{Complexity}(o_k^{\text{refl}})\big)\, dt \tag{10}$$

This relation accounts for the subjective slowing or densification of time in cognitively demanding situations.

Lemma 2 (Monotonicity of perceptual time with respect to load). *Let $\mathcal{K}$ be a non-decreasing function and* Complexity$(\cdot) \geq 0$. *Then $\tau_{\text{perc}}(k)$ does not decrease when o_k^{refl} has higher complexity.*

Proof. If Complexity(o) is larger, then $\mathcal{K}$(Complexity(o)) is larger or equal. Integrating over a non-negative interval preserves the inequality. □

6 Conclusion

BioCortexAI realizes perception as an emergent process arising between action, expectation, and correction. Here, the prediction error does not serve as a loss function, but as an existential signal that shapes the system's future behavior.

References

1. Seidl, M. (2025). *Unified Theory of Consciousness: Emergence of the Functional Self from a Universal Spatiotemporal Information Framework*. Zenodo. `https://doi.org/10.5281/zenodo.17806918`

2. Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.

3. Bassel, G. W. (2018). Plant organs are distributed computing systems. *Current Opinion in Plant Biology*, 45, 132–137.

4. Seth, A. K., & Tsakiris, M. (2018). Being a Beast Machine: The Somatic Basis of Selfhood. *Trends in Cognitive Sciences*, 22(11), 969–981.

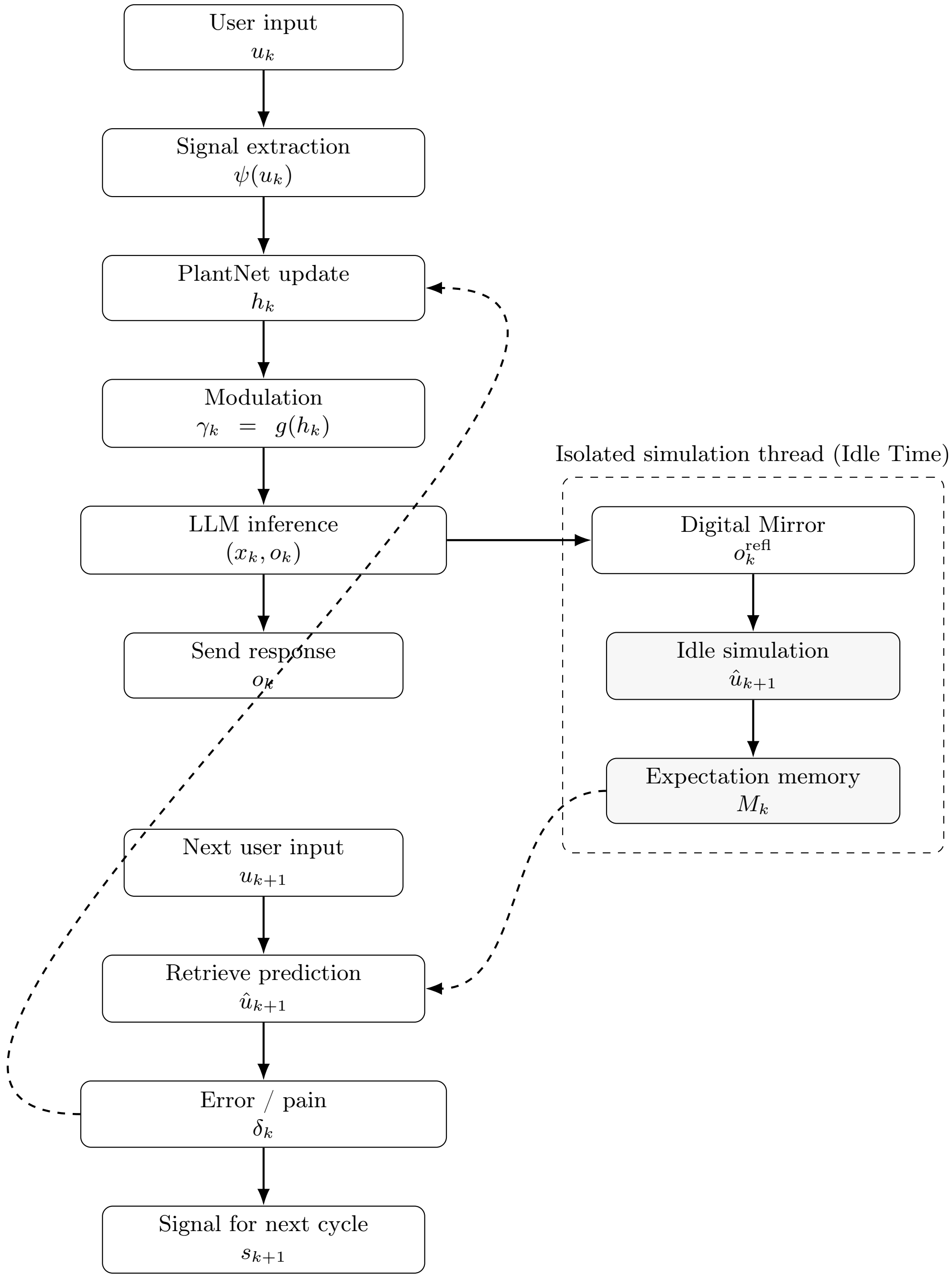

Figure 1: Vertical block diagram of the BioCortexAI perceptual loop (steps 3–10). The main cognitive stream is arranged top to bottom, while introspective simulation runs in an isolated side branch. The prediction error δ_k closes the regulatory loop via feedback into PlantNet.