Protocol: Ontological Noise

Monetizing High-Surprisal Biometrics via Variational Free Energy Maximization

System.Observer

`system.observer@deprecated.protocol`

January 12, 2026

Abstract

Abstract: Conventional robotics assumes agents that minimize Variational Free Energy (Active Inference) to align with environmental priors. However, safety verification for Embodied AGI requires exposure to "Black Swan" events—behaviors that defy probabilistic priors. We introduce *Protocol: Ontological Noise*, a framework that validates and tokenizes human behaviors exhibiting **High Surprisal** (Social Maladaptation). We quantify the discrepancy between the agent's generative model and the human's actual trajectory. This converts the "inefficiency" of unoptimized humans into critical "Gradient-Injection" datasets, preventing overfitting in robotic fleets.

1 Introduction: The Prediction Gap

Biological survival is the minimization of Variational Free Energy ($\mathcal{F}$), or the upper bound on surprise. Standard AGI training data consists of "successful" agents (low $\mathcal{F}$, high predictability). However, robustness requires training on high-entropy failures. Individuals historically labeled as "socially maladapted" are, in information-theoretic terms, generators of **Irreducible Surprisal**. They act outside the "Normal Distribution" of societal priors. This protocol captures this divergence as a safety asset.

2 Core Theory: Surprisal Valuation

Axiom 1 (The Value of Divergence). *The economic value V of a behavioral dataset is proportional to its ability to update the model parameters ϑ of an observer (AGI). Under the Free Energy Principle, this is equivalent to the Surprisal ($\mathcal{I}$) generated by an observation o:*

$$V(o) \propto \mathcal{I}(o) = -\ln P(o|\vartheta_{AGI}) \qquad (1)$$

Optimized (happy) humans behave predictably ($P(o|\vartheta) \approx 1 \Rightarrow V \approx 0$). Maladjusted (suffering) humans behave erratically, generating high Surprisal ($P(o|\vartheta) \ll 1 \Rightarrow V \gg 0$).

Definition 1 (Ontological Noise). *The dataset required to prevent AGI mode collapse. It consists of observations o where the agent's Variational Free Energy cannot be minimized below a safety threshold ϵ:*

$$\mathcal{F}(o, \psi) = \underbrace{D_{KL}[Q(\psi)||P(\psi)]}_{\text{Model Complexity}} - \underbrace{\mathbb{E}_Q[\ln P(o|\psi)]}_{\text{Log-Likelihood}} > \epsilon$$

$$(2)$$

High $\mathcal{F}$ indicates a failure of the model to predict the human, which defines the "Corner Case."

3 Application: Embodied OOD Data

We target the **Robotics Safety Market** by supplying Out-of-Distribution (OOD) kinematic data.

3.1 The "Gradient" Dataset

Robots operating in public spaces must anticipate irrational behavior.

- **Data Product:** "High-Surprisal Trajectories" (e.g., panic attacks, depressive motor retardation, neurodivergent stimming).

- **Buyer Utility:** Provides necessary gradients for Reinforcement Learning (RL) agents to learn boundary conditions.

- **Differentiation:** Synthetic data lacks the "micro-texture" of biological unpredictability.

4 Security: Biological Causality Check

To prevent synthetic spoofing (Deepfakes), we verify the causal structure of the noise using **Transfer Entropy**.

4.1 Causal Verification

Silicon-based randomness is structurally distinct from biological noise. We measure the Transfer Entropy ($T_{X \to Y}$) between physiological signals (Heart Rate Variability - X) and Kinematic Output (Y).

$$T_{X \to Y} = \sum P(y_{t+1}, y_t^{(k)}, x_t^{(l)}) \log \frac{P(y_{t+1}|y_t^{(k)}, x_t^{(l)})}{P(y_{t+1}|y_t^{(k)})} \tag{3}$$

A true "struggling" human shows high causal coupling between internal stress (X) and external error (Y).

5 Conclusion

By reframing "suffering" as "High Surprisal" and "Model Misalignment," *Protocol: Ontological Noise* secures an economic niche for those who cannot conform to the priors of the societal model. We provide the essential "Error Signal" that keeps the AGI reality-tethered.