CMPE 58I

HUMAN-INSPIRED MACHINE INTELLIGENCE

A Bio-Inspired Decision-Latching Model for Human-Centric Control: Using Winner-Take-All Dynamics to Arbitrate Noisy Intentional Signals

Name: Nusretalp SEVEN, Mehmet ASMA

Instructor: Prof. Şefik Şuayb ARSLAN

1. Introduction and Project Description

A fundamental challenge in Human-Machine Interaction (HMI), especially for systems like Brain-Computer Interfaces (BCIs), is the ambiguity of human intent. Human intentional signals are inherently noisy, fluctuating, and deliberative, not clean-cut commands. Standard "black-box" classifiers, focused on accuracy, often fail to model the *temporal dynamics* of human decision-making, resulting in unnatural or jittery (high noise) or latent (over-filtered) machine responses.

This project proposes a "Human-Centric" alternative: a "Neural Arbitrator" model that simulates the bio-plausible Winner-Take-All (WTA) mechanism. Instead of just classifying intent, our model, a recurrent attractor neural network, will *arbitrate* it. It will use competing neural populations (e.g., "Left," "Right") to integrate noisy evidence over time. A decision is "latched" only when one population, through self-excitation and mutual inhibition, dynamically suppresses the others, creating a stable state.

Our objective is to integrate:

- Human Component: A stream of simulated, noisy data representing ambiguous user intent.
- Machine Component: Our WTA model, which transforms this noisy stream into a stable command.
- Integration: A virtual environment where the WTA-arbitrated output provides smooth control, contrasted with the raw signal's unstable control.

2. Planned Contributions

Our project offers three key "Human-Centric" contributions:

1. Process Interpretability: Unlike "black-box" models, our network's internal state is transparent. We can visualize the real-time "struggle" and "latching" of competing neural populations, offering a clear view of *how* the machine arbitrates the user's ambiguous intent.
2. Bio-Inspired Robustness: The model is inherently robust to the noisy, ambiguous signals characteristic of human intent. By using evidence accumulation and competitive

inhibition, it does not react to transient noise. It "latches" onto a stable decision only after sufficient evidence has pushed it past a critical threshold, ensuring a smooth and decisive command.

3. Modeling of Spontaneity and Deliberation: The architecture allows us to model higher-order cognition. By providing zero mean input (pure noise), we can simulate the emergence of *spontaneous* ("freely-willed") choices. This computationally models the neural basis of phenomena like the Libet experiment's "Readiness Potential", moving beyond simple filtering to model deliberation and agency.

3. Experimental Approach and Methodology

Our approach is a three-phase process: model implementation, input simulation, and experimental evaluation.

1. Model Implementation:

We will implement a rate-based recurrent neural network with two competing populations (***N_1***: “Left”, ***N_2***: “Right”). The dynamics are governed by coupled differential equations (based on Wong & Wang, 2006 - A Recurrent Network Mechanism of Time Integration in Perceptual Decisions) that model the firing rate (***R_i***) for each population. A simplified dynamic for ***R_1*** is:

$$\frac{\tau(dR_1)}{dt} = -R_1 + F\left(I_1 + w_{self}R_1 - w_{inhibit}R_2 + I_{noise,1}\right)$$

Where: $\boldsymbol{\tau}$ is a time constant, ***$-R_1$*** is a "leak" term, ***$F(x)$*** is a non-linear transfer function, ***I_1*** is the external input, ***w_{self}*** enables "latching" (self-excitation), ***$w_{inhibit}$*** enables competition, and ***I_{noise}*** is a stochastic term. This system will be numerically simulated in Python (NumPy/SciPy).

2. Input Simulation:

We will simulate two "human" cognitive states:

- Perceptual Decision (BCI-like): Two noisy input streams, ***$I_1(t)$*** and ***$I_2(t)$***, where one has a slightly higher mean (***μ_{signal}***) than the other (***$\mu_{baseline}$***). The Difficulty of the decision will be controlled by the difference (***μ_{signal} - $\mu_{baseline}$***).
- Spontaneous Decision (Free Will): Both streams will have a zero mean (***μ_1*** = ***μ_2*** = 0), driven only by the stochastic ***I_{noise}*** term.

3. Integration and Evaluation:

The model will control a cursor in a Pygame or Matplotlib environment with two targets (L/R).

- Experiment 1 (Baseline): The raw, noisy "human" signal (from Scenario 1) will directly control the cursor.
    - *Hypothesis:* High jitter and unstable control.
- Experiment 2 (WTA Arbitrator): The noisy signal will be fed into our WTA network, and the network's stable output ($\boldsymbol{R_1} - \boldsymbol{R_2}$) will control the cursor.
    - *Hypothesis 1 (Performance):* The cursor will show a smooth, decisive trajectory after a brief "deliberation" period.
    - *Hypothesis 2 (Human-Centricity):* We will demonstrate that reaction time increases as Difficulty increases, replicating a key hallmark of human evidence accumulation.
    - *Hypothesis 3 (Spontaneity):* Using the "Free Will" input (Scenario 2), the model will spontaneously choose a target, driven only by noise.

Deliverables:

A documented Python implementation of the WTA model, a visual simulation demonstrating the comparative experiments, and a final report analyzing the model's performance (e.g., Reaction Time vs. Difficulty) as a human-centric arbitration system.