BioPlaN: Biologically Plausible Decision Making in RNNs

Harpreet Singh
Department of Computer Science and Engineering
IIIT Hyderabad
harpreet.singh@students.iiit.ac.in

Keerthi
Robotics Research Center
IIIT Hyderabad
keerthi.keerthi@research.iiit.ac.in

***Abstract*—In this project we explore the computational neuroscience by developing "BioPlaN," a framework for biologically plausible Recurrent Neural Networks. While standard RNNs are powerful, they often lack the structural constraints found in biological brains. To bridge this gap, we implemented Dale's Law to enforce strict excitatory or inhibitory roles on neurons and utilized Curriculum Learning to mimic natural progression from simple to complex tasks. We introduce two novel mechanisms. First, for Perceptual Discrimination we propose Adaptive Contrastive Gating (ACG). It is designed to dynamically filter sensory noise and enforce a clear separation of decision states. Second, for Delayed Match-to-Sample tasks, we introduce Adaptive Deliberate Replay (ADR) which is a rehearsal mechanism allows a signal to survive long delay periods. By testing these models on Perceptual Discrimination and Delayed Match-to-Sample tasks, we found that ACG demonstrated superior robustness against structural damage and ADR demonstrates improved accuracy in the working memory tasks as compared to baseline models.** [1]

***Index Terms*—Recurrent Neural Networks, Computational Neuroscience, Dale's Law, Curriculum Learning, Decision Making, Adaptive Gating, Contrastive Loss**

I. Introduction

Recurrent Neural Networks (RNNs) [1] are particularly effective at modeling cognitive tasks that require the integration of information over time, such as evidence accumulation and working memory [2]. By training and experimenting the RNNs on standard psychophysical tasks, researchers can generate hypotheses about how biological brains process temporal sensory data.

In this project we explore how neurons in RNNs can be forced to obey excitatory or inhibitory roles, following the *Dale's Law* [3]. Second, add specific gating mechanisms biological circuits use to filter sensory noise. Third, analysing the internal dynamics of these networks, specifically understanding *how* it solves the task, by identifying the causal circuit mechanisms and attractor dynamics [4].

The primary objectives of this project were:

- To utilize *Curriculum Learning* [5], a method where the network is trained on high-coherence (easy) examples first, with noise levels gradually increasing as performance improves.
- To implement Dale's Law [3] within the weight matrices, forcing neurons to be strictly excitatory or inhibitory.
- We introduce a novel architectural component, *Adaptive Contrastive Gating (ACG)*. This mechanism allows the network to dynamically gate [6] input based on signal coherence and utilizes a contrastive loss [7] to enforce clear separation between decision manifolds in the state space.
- We perform a rigorous mechanistic analysis [4]. Using dynamic causal perturbations and fixed-point analysis, to visualize the state-space trajectories and identify the specific neurons causally dominant in decision-making.
- We introduce an Adaptive Deliberate Replay (ADR) suppression mechanism for recurrent neural networks that regulates internal memory reactivation during delay period working memory tasks [8].

II. Literature Review

Recurrent Neural Networks (RNNs) [1] have become essential for modeling cortical dynamics [2]. [9] show RRNs can mimic low-dimensional prefrontal cortex activity during decision-making. We address biological realism by enforcing Dale's Law as proposed by [3]. Furthermore, to enhance training, we integrate [5] Curriculum Learning and refer to [7] which explain why contrastive methods can learn good embeddings. We also use PsychRNN [10] toolkit for 2AFC [11] and DMS [12] tasks. While our Adaptive Deliberate Rehearsal mechanism is novel within computational neuroscience, it is inspired by classical psychological models of deliberate and controlled rehearsal, originally studied in human short-term memory experiments [13].

III. Problem Statement

To evaluate the decision-making and memory capabilities of our AI model, we define two specific cognitive tasks. These tasks formally model evidence accumulation and the maintenance of information over a delay period.

A. Perceptual Discrimination

In computational neuroscience, Perceptual Discrimination(2AFC) [11] task is a decision-making task where a subject must classify a sensory stimulus into one of two categories. A classic example of a perceptual discrimination task is random-dot motion task where a subject views a cloud of moving dots and must decide if the net motion is to the left or to the right.

[1] All Code and Model files are available at: https://github.com/Harpreet287/INCM-course-project

Let x_t be the input the neural network receives. We can formally write

$$x_t = \mu + \sigma \cdot \epsilon_t$$

where μ is the coherence of the signal and ϵ_t is a random number drawn from a standard normal distribution which represents noise. The difficulty is controlled by coherence, which is the strength of the signal relative to the noise. Higher coherence is associated with a strong signal and low coherence implies signal is very weak; almost looking like noise. For a decision variable D, at any time step we have

$$D_{new} = D_{old} + x_t$$

Notice that if μ was positive, D will slowly drift upwards and if the signal was negative, D will slowly drift downward.

B. Delayed Match-to-Sample

The Delayed Match-to-Sample(DMS) [12] task evaluates working memory [9] by requiring the comparison of two temporally separated stimuli. The input x_t is defined piecewise over three intervals, with additive noise η_t:

$$x_t = \eta_t + \begin{cases} f_1 & 0 < t \leq T_1 \quad \text{(Encoding)} \\ 0 & T_1 < t \leq T_2 \quad \text{(Maintenance)} \\ f_2 & T_2 < t \leq T_3 \quad \text{(Comparison)} \end{cases}$$

The network must bridge the maintenance delay to compute the relation $y = \mathcal{F}(f_1, f_2)$ at $t = T_3$. A loss function is minimized to optimize the instances where y matches the ground truth.

IV. Methodology

A. Baseline

We first train standard RNN from PsychRNN toolkit [10] where the internal neurons are perfectly reliable, where the hidden state update is given by:

$$h_t = \tanh(W_{in} x_t + W_{rec} h_{t-1} + b)$$

, to establish the maximum theoretical performance of the network on the task without any internal interference. Next we test the same model but with a random noise of ξ_t, which gives us new hidden state as:

$$h_t = \tanh(W_{in} x_t + W_{rec} h_{t-1} + b + \mathbf{\xi_t})$$

where $\xi_t \sim \mathcal{N}(0, 0.1)$.

B. Curriculum Learning

Curriculum Learning is a training strategy [5] for 2AFC [11] task where an RNN is trained on easier examples first, gradually decreasing the coherence as the model improves. In an easy task, signal is very strong as compared to the noise. In a hard task, signal is very weak, input almost looks like random noise. We begin the training with a high coherence and gradually drop the coherence if the model's accuracy exceeds a threshold.

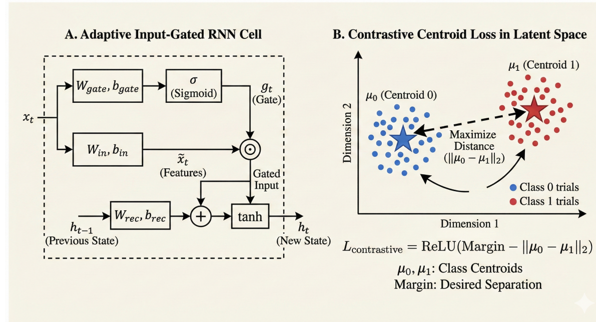

Fig. 1. Novel ACG mechanism (A) to filter noise and a contrastive centroid loss (B) to enforce separation between class representations in the latent space

C. Dale's Law

Biologically Dale's Law [3] states that a neuron performs the same chemical action at all of its synaptic connections to other cells, regardless of the identity of the target cell. In the context of RNNs, this implies that a neuron is either purely Excitatory i.e. outputs positive signals or purely Inhibitory i.e. outputs negative signals. It cannot excite one neighbor while inhibiting another. To satisfy Dale's Law, every column in the recurrent weight matrix must consist entirely of either non-negative or non-positive values. We refer to the modified RNN as *DalesRNN* [3].

D. Dynamic Causal Perturbations

Dynamic Causal Perturbations [4] refers to deliberately interfering with the RNN's normal operation. We either force a neuron's activity to a fixed value, clamping, or permanently silencing neurons, lesioning. We look at which neuron's activity matters the most for the final decision using Gradient-Based Saliency. We zero (or fixate) outgoing connections from each neuron in the recurrent weight matrix, regardless of its natural input. This effectively prevents the neuron from influencing the rest of the network, simulating a lesion.

E. Adaptive Contrastive Gating

We developed this algorithm for this specific project to help RNN to solve noisy perceptual task better, Fig. 1. Our model dynamically learns to ignore the noise by calculating a "gate" value between 0 and 1. If the gate is 0, the input(x_t) is ignored otherwise it passes through. It is "adaptive" because the gate creates its value based on the input itself at every time step.

$$g_t = \sigma(W_{gate} x_t + b_{gate})$$

$$h_t = \tanh\left(\underbrace{(W_{in} x_t + b_{in}) \odot g_t}_{\text{Gated Input}} + W_{rec} h_{t-1} + b_{rec}\right)$$

Additionally we want the RNN's internal state to clearly distinguish between the two choices. It does so by calculating the "center of mass" or centroid for all trials where the answer was A, and the centroid for answer B, and forces them to be far apart. Formally

$$L_{contrastive} = \text{ReLU}(\text{Margin} - ||\mu_0 - \mu_1||_2)$$

where μ_0 is the average final state for Class 0 trials, and μ_1 is the average for Class 1 trials. If the distance between clusters is smaller than the Margin, the loss increases. Else, if they are already far apart it is close to 0.

F. *Causal and Mechanistic Analysis*

To determine causality we do a Lesion Study. In a standard observation, if a neuron is firing during a decision, implies that it is *correlated* with that decision [9]. By lesioning (deleting) it we observe if the accuracy drops to infer whether that neuron was *causally* necessary for the decision.

We perform two primary mechanistic analyses to understand how the network compute the answer. First, State-Space Trajectories reveal the "thought process" of the RNN as a moving dot in 3D space. Two paths, one for choice A and one for choice B should diverge from each other as time progresses. Second, Fixed Point Analysis [4] where we sought to find fixed points or natural attractors, a state h^* where the network stops changing, formally:

$$h^* \approx \tanh(W_{rec}h^* + \text{Input})$$

G. *Adaptive Deliberate Replay*

We propose improvements on RNN for delayed-match-to-sample (DMS) [12] task. We add a new rehearsal predictino mechanism which estimates how likely is it to re-activate a memory at a time stamp. We optimize on loss function which has three goals to improve–task accuracy, auxiliary stimulus reconstruction, and minimizing unnecessary mental rehearsal to encourage efficiency. Formally put the loss is:

$$\mathcal{L} = \mathcal{L}_{\text{task}} + \lambda_{\text{aux}}\mathcal{L}_{\text{aux}} + \lambda_{\text{rehe}} \sum_t p(t),$$

where $\mathcal{L}_{\text{task}}$ is cross-entropy on task outputs, $\mathcal{L}_{\text{aux}}$ is auxiliary reconstruction loss, and the $\lambda_{\text{rehe}} \sum_t p(t)$ penalizes excessive rehearsal. By gradually increasing the difficulty of the task, by increasing the duration between two stimulus, we can evaluate how ADR improves the accuracy overtime.

V. Experimental Setup

The dataset for Perceptual Discrimination task is synthetic and generated on-the-fly. A batch of random Gaussian noise is generated and a target label y (0 or 1) is chosen randomly for each trial with a coherence value added to the channel corresponding to the label y for all time steps. The output is a prediction vector at the final time step which are logits for the two possible classes. We use PsychRNN's library [10] to generate data for Delayed Discrimination task.

To generate fixed points for mechanistic analysis, we use the LBFGS [14] optimizer to create a random state h and tweak it until the difference between the current state and the next state($|\tanh(W_{rec}h_{old} + \text{Input}) - h|$)is close to 0.

We use scikit-learn's implementation of PCA to reduce the higher dimensions to pick the principal components.

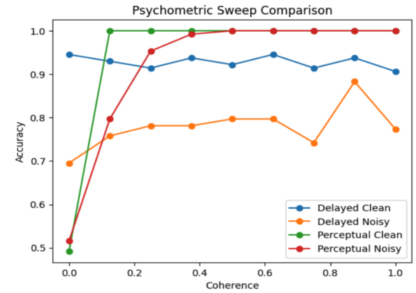

Fig. 2. Accuracy vs. Coherence across Perceptual and Delayed tasks for Clean and Noisy models.

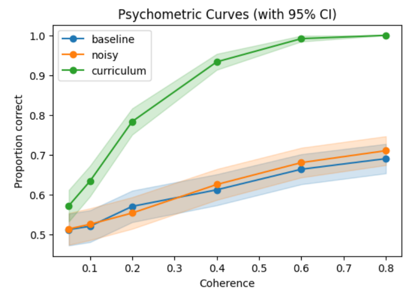

Fig. 3. Curriculum *DalesRNN* outperforms noisy and baseline models.

TABLE I
Psychometric Accuracy for *DalesRNN*

Coherence	Baseline	Noisy	Curriculum
0.05	0.51	0.51	**0.57**
0.20	0.57	0.56	**0.78**
0.40	0.61	0.63	**0.93**

For ADR, Training was performed across four curriculum stages with delays ranging from 2-30 timesteps and distractor intensities from 0-0.2. All models used 100 recurrent units, batch sizes of 64, and 400-1000 iterations per stage. The ADR model augments a standard recurrent neural network with two lightweight heads and a feedback path.

VI. Results and Discussion

In 2AFC [11], we observe(see Fig. 2) 100% accuracy with clean model(PsychRNN [10]) and 97.8% accuracy with noisy model which displays models mastered this task easily. In DMS [12] we observe 91.41% accuracy on clean model and 77.34% accuracy with noisy model from which we infer that model performed significantly worse, suggesting that internal neural noise disrupts the ability to maintain stable memories over time. On stardard RNN with curriculum learning [5] we observe curriculum model consistently outperforming baseline and noisy models. Similar phenomenon was observed with *DalesRNN* where curriculum model maintained decent performance even at lower coherences, see Fig. 3 and Table I. In our experimental setup of involving ACG and curriculum-ACG, we observe standard ACG model performing slightly better than baseline and curriculum-ACG outperforming. This

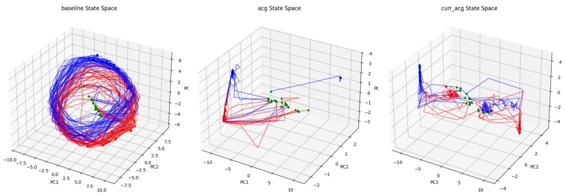

Fig. 4. 3D projections show that ACG forces the network to push its internal representations into distinct, well-separated decision paths.

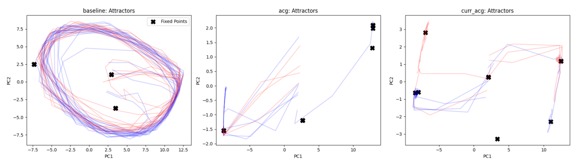

Fig. 5. Stable fixed points at the ends of the trajectories of ACG models. Notice stability is not there in baseline.

proves the contrastive loss forced the network to organize its activity into separate clusters(Fig. 4 and Fig. 5) and curriculum approach allowed the model to learn effectively by combining best of both approaches. The lesion analysis showed ACG model is more robust to lesions(Fig.6 and Table II) than the Baseline model.

PCA analysis shows activity of two decisions flowing to opposite corners therefore separating cleanly and fixed point analysis shows stable fixed points at the ends of the trajectories for ACG-models. For baseline model, representation look tangled and no stable fixed points.

Hence this demonstrates the correctness of our ACG-RNN. The ADR-augmented model demonstrated a rapid convergence during the initial curriculum phases (Stages 0–1), achieving perfect task accuracy on short-delay trials(Fig. 7). However, under conditions of extended delays and stronger distractors, the rehearsal controller tended to prioritize metabolic efficiency effectively suppressing the reactivation gate ($p(t) \to 0$). This suppression resulted in a degradation of behavioral accuracy to near-chance levels. Subsequent ablation studies confirmed the functional necessity of the rehearsal mechanism when an active replay was either removed or artificially forced, the performance collapsed to chance. This confirms that the ADR mechanism actively shapes the internal dynamics required to bridge temporal delays.

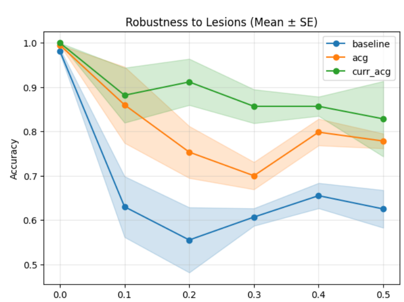

Fig. 6. ACG maintains a high accuracy even as large fractions of the network are surgically silenced.

TABLE II
ROBUSTNESS TO LESIONS

Fraction Lesioned	**Baseline**	**ACG**	**Curr-ACG**
0.2	0.56	0.75	**0.91**
0.4	0.66	0.80	**0.86**

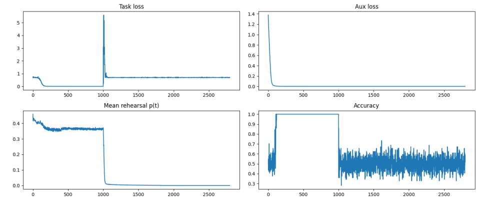

Fig. 7. Performance of the ADR-augmented recurrent network across curriculum stages.

VII. Conclusion

In this project, we demonstrate that adding biological constraints to RNNs does not hinder performance and there are ways through which we can enhance the robustness and interpretability. Through BioPlaN, we combined Dale's Law, Curriculum Learning, causal and mechanistic analysis. We propose a novel Adaptive Contrastive Gating which allows an RNN to solve 2AFC task and increase stable internal dynamics. Adaptive Deliberate Replay is a novel mechanism to control internal memory reactivation in RNNs for DMS task and establishes a fundamental trade-off between cognitive rehearsal costs and behavioral accuracy.

To conclude this project, we now understand that we can experiment with AI systems through the lens of computational neuroscience and make them more accurate and robust in different training environments.

References

[1] R. M. Schmidt, "Recurrent neural networks (rnns): A gentle introduction and overview," 2019. [Online]. Available: https://arxiv.org/abs/1912.05911

[2] G. R. Yang and X.-J. Wang, "Artificial neural networks for neuroscientists: A primer," *Neuron*, vol. 107, no. 6, pp. 1048–1070, 2020.

[3] H. F. Song, G. R. Yang, and X.-J. Wang, "Training excitatory-inhibitory recurrent neural networks for cognitive tasks: a simple and flexible framework," *PLoS computational biology*, vol. 12, no. 2, p. e1004792, 2016.

[4] D. Sussillo and O. Barak, "Opening the black box: low-dimensional dynamics in high-dimensional recurrent neural networks," *Neural computation*, vol. 25, no. 3, pp. 626–649, 2013.

[5] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in *Proceedings of the 26th annual international conference on machine learning*, 2009, pp. 41–48.

[6] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.

[7] S. Arora, H. Khandeparkar, M. Khodak, O. Plevrakis, and N. Saunshi, "A theoretical analysis of contrastive unsupervised representation learning," *arXiv preprint arXiv:1902.09229*, 2019.

[8] M. G. Stokes, "Activity-silent working memory," *Trends in cognitive sciences*, vol. 19, no. 7, pp. 394–405, 2015.

[9] V. Mante, D. Sussillo, K. V. Shenoy, and W. T. Newsome, "Context-dependent computation by recurrent dynamics in prefrontal cortex," *Nature*, vol. 503, no. 7474, pp. 78–84, 2013.

[10] D. B. Ehrlich, J. Stone, D. Brandfonbrener, A. Atanasov, and J. D. Murray, "Psychrnn: An accessible and flexible python framework for training recurrent neural networks on cognitive tasks," in *Conference on Cognitive Computational Neuroscience*, 2021.

[11] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, "The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks," *Psychological Review*, vol. 113, no. 4, p. 700, 2006.

[12] E. K. Miller, C. A. Erickson, and R. Desimone, "Neural mechanisms of visual working memory in prefrontal cortex," *Nature*, vol. 380, no. 6569, pp. 60–62, 1996.

[13] F. I. M. Craik and M. J. Watkins, "The role of rehearsal in short-term memory," *Journal of Verbal Learning and Verbal Behavior*, vol. 12, no. 6, pp. 599–607, 1973.

[14] D. C. Liu and J. Nocedal, "On the limited memory bfgs method for large scale optimization," *Mathematical Programming*, vol. 45, pp. 503–528, 1989. [Online]. Available: https://api.semanticscholar.org/CorpusID:5681609