NUS National University of Singapore | NUS Graduate School

Integrative Sciences & Engineering Programme

Centre for Frontier AI Research

CFAR

Gradually learning representations using biologically plausible and deep learning systems for rapid performance

M Ganesh (“guh-nay-sh”) Kumar

National University of Singapore (NUS)

Agency for Science, Technology and Research (A*STAR)

m_ganeshkumar@u.nus.edu   M Ganesh Kumar   mgkumar138   @Ganeshk92

NUS National University of Singapore | NUS Graduate School

Integrative Sciences & Engineering Programme

Biologically plausible computations underlying one-shot learning of paired associations

M Ganesh Kumar

Integrative Sciences and Engineering Programme (ISEP)
Department of Electrical and Computer Engineering (ECE)
National University of Singapore (NUS)

Co-Advised by: Dr Shih-Cheng Yen and Dr Andrew Tan
Thesis Advisory Council: A/Prof Thomas Yeo, Dr Cheston Tan, Dr Camilo Libedinsky

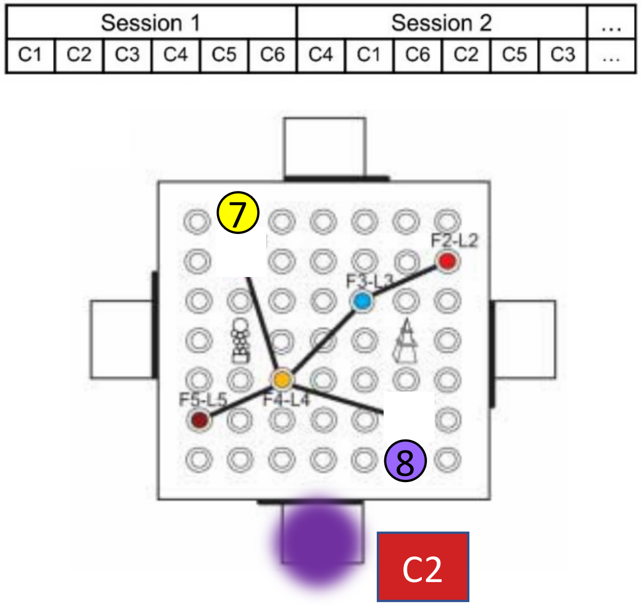
How do we do grocery shopping in a new market?

<u>Grocery list</u>

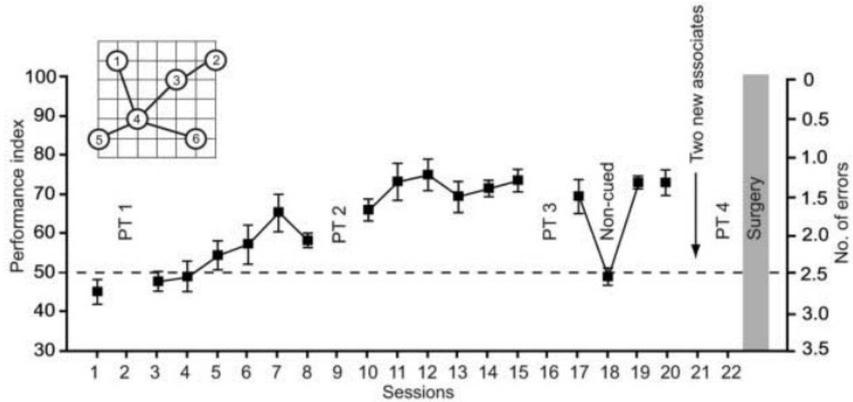
- Milk

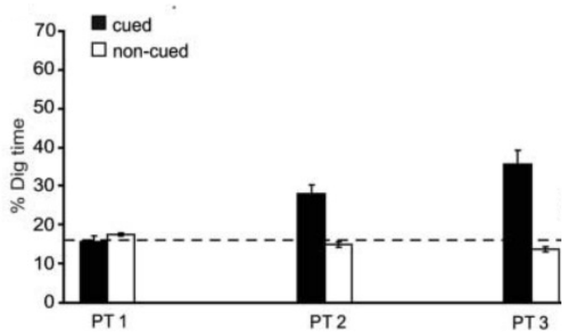
1) Gradually learn the new market
2) Recall milk location and quickly navigate

Gradual then one-shot learning of Multiple Paired Associations (MPA) by rodents



6 FLAVOUR-LOCATION paired associations

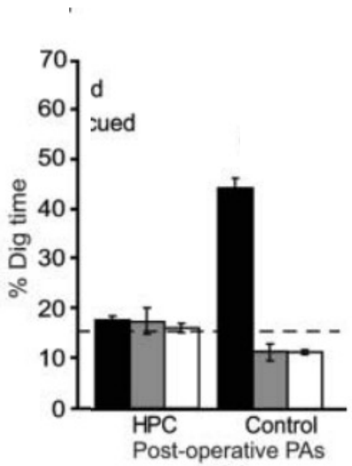




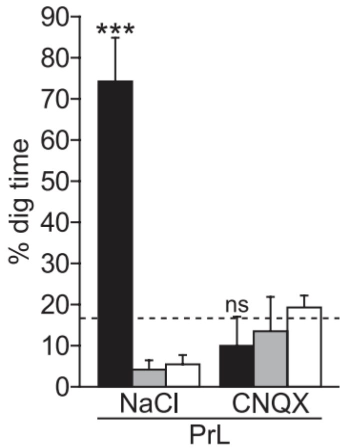
Tse et al. (2007) Science

Brain regions and environment for one-shot learning

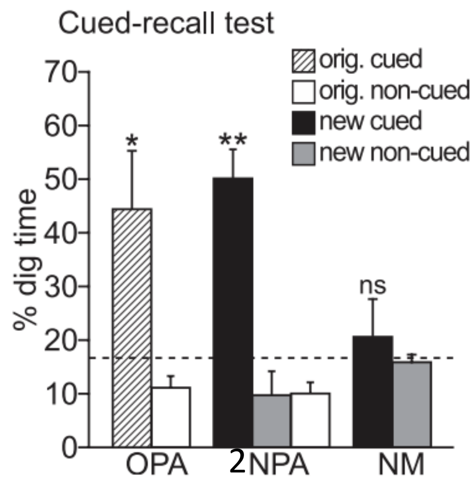
Hippocampus needed for one-shot learning



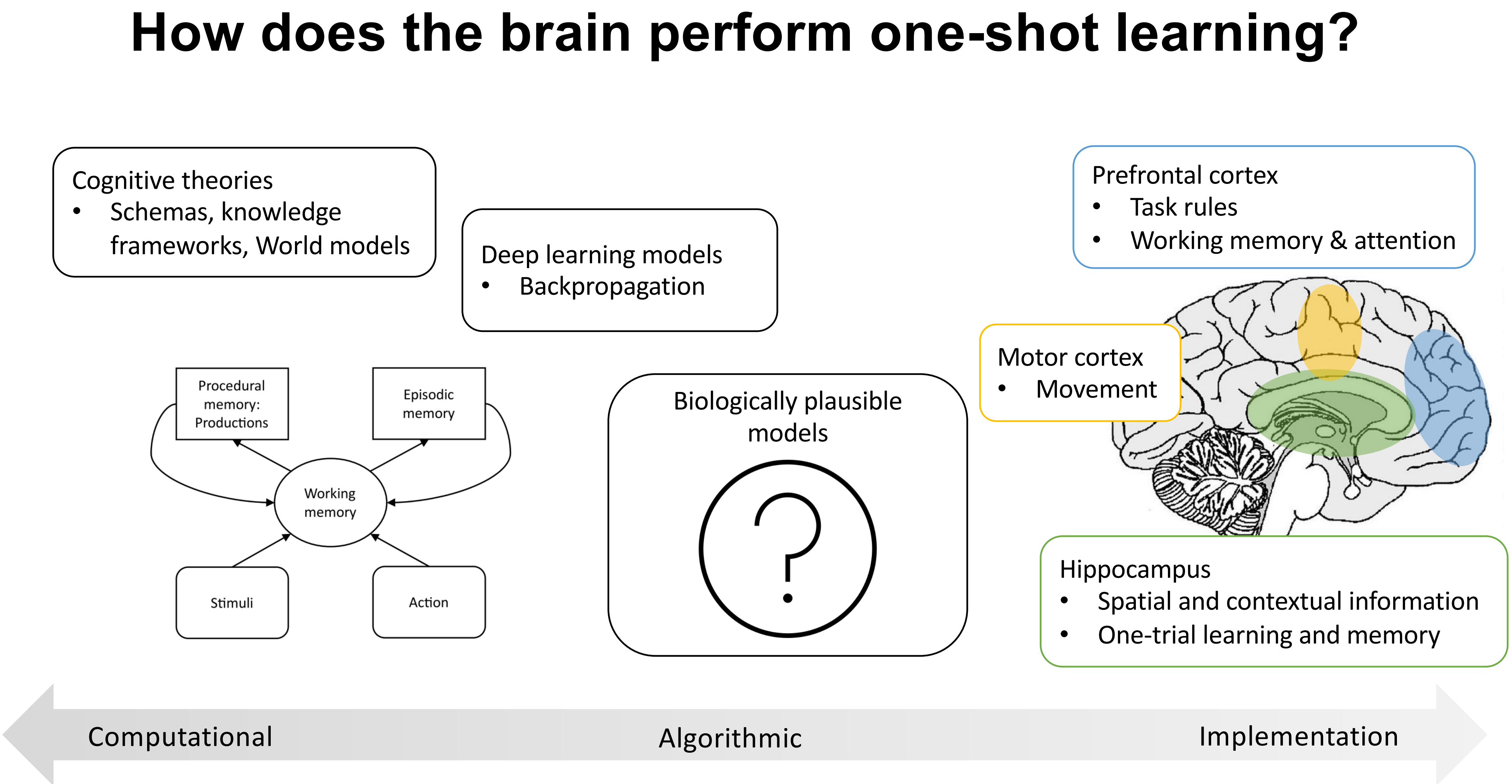
Prefrontal cortex recalls goals after consolidation



New environment requires new learning



Tse et al. (2007); (2011)


How does the brain perform one-shot learning?
Cognitive theories
• Schemas, knowledge frameworks, World models
Deep learning models
• Backpropagation
Procedural memory: Productions
Episodic memory
Working memory
Stimuli
Action
Biologically plausible models
Prefrontal cortex
• Task rules
• Working memory & attention
Motor cortex
• Movement
Hippocampus
• Spatial and contextual information
• One-trial learning and memory
Computational
Algorithmic
Implementation

How do biologically plausible neural circuits learn multiple paired associations after one-trial?

- **Research question**
  - How are schemas represented in neural networks to facilitate one-shot learning?
    - Neural architecture
    - Biologically plausible learning algorithms
- **Specific aims**
  1. Develop biologically plausible reinforcement learning agent to **gradually** learn the Multiple Paired Association (MPA) task
  2. Develop biologically plausible agent to learn multiple new flavour-location paired associations (PA) after **one trial**

Gradually learn MPA

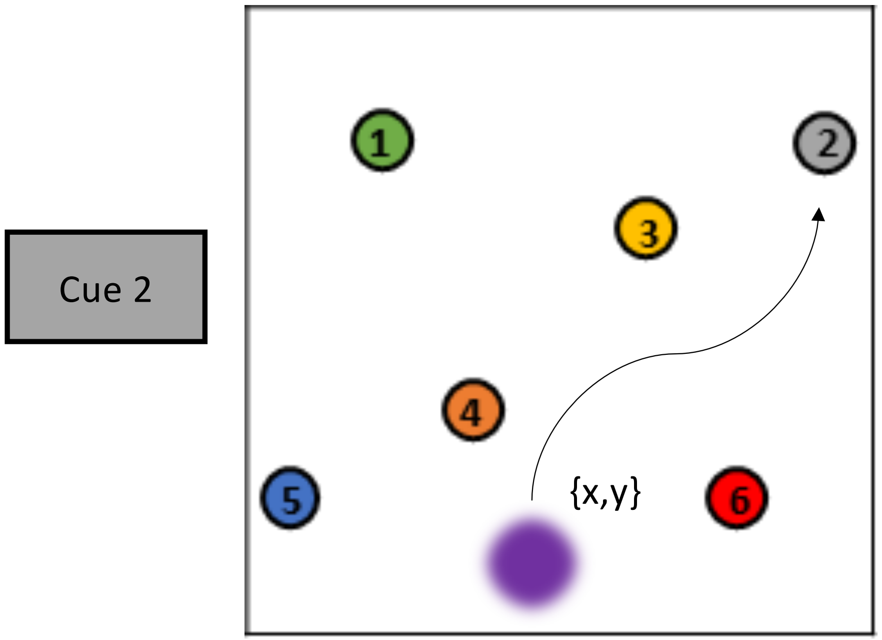
Can classical biologically plausible reinforcement learning agents learn MPA?

If not, what agent can?

Task setup and solution

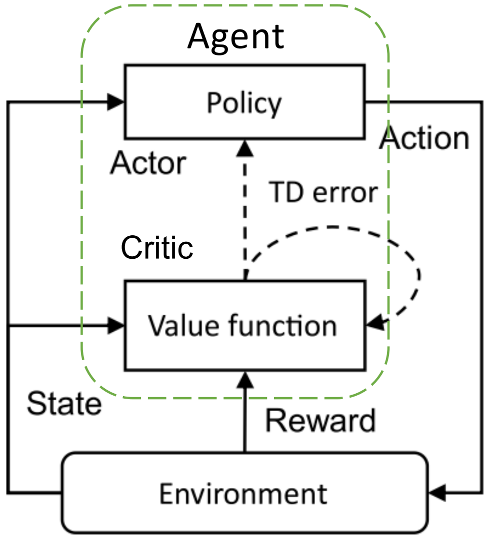
Spatial navigation:

Reward when cue associated goal is reached



Traditional solution:

Model-free reinforcement learning algorithm



Temporal difference error

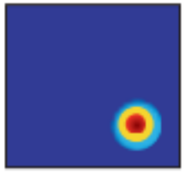
$$\delta(t) = R(t) + \gamma V(t) - V(t-1)$$

Sutton & Barto (2018)

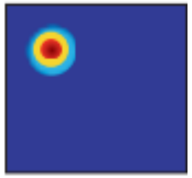
Canonical biologically plausible RL agent

Spatially selective Hippocampal place cells

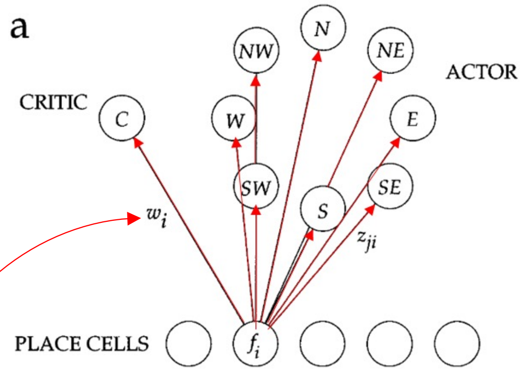
Place cell 1



Place cell 2

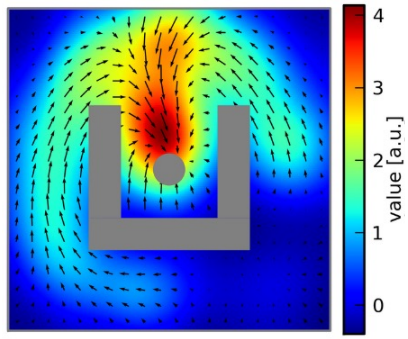


Temporal Difference error modulated Hebbian Learning (TD-HL)



$$\boxed{\Delta W_{ij}(t) \propto \overbrace{r_j(t)}^{\text{Presynaptic}} \cdot \underbrace{z_i(t)}_{\text{Postsynaptic}} \cdot \overbrace{\delta^{DA}(t)}^{\text{Scalar signal}}}$$

Critic learns value map
Actor learns policy map

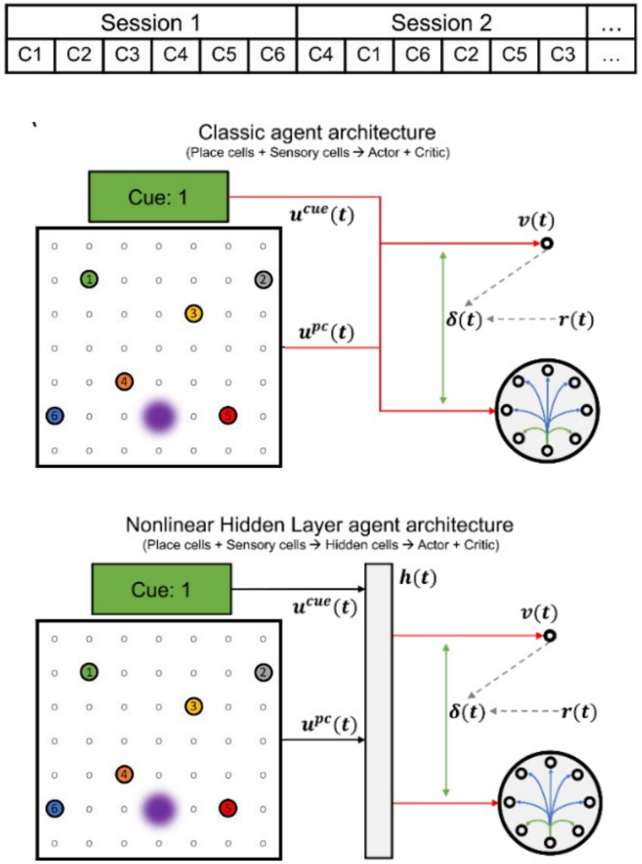


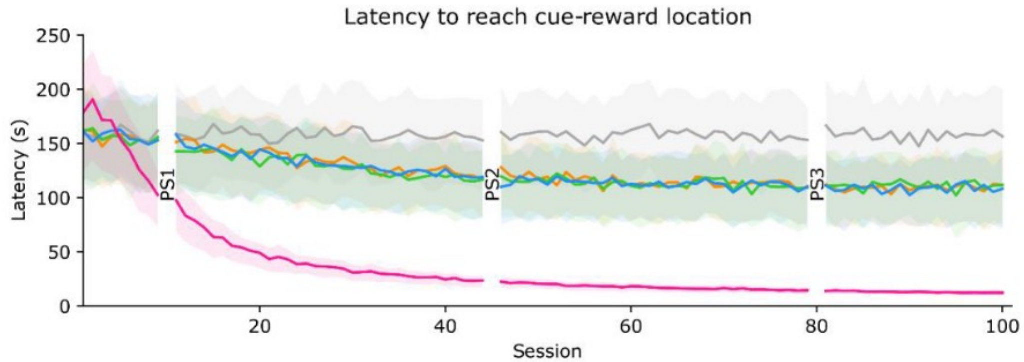
Biologically plausible learning rule

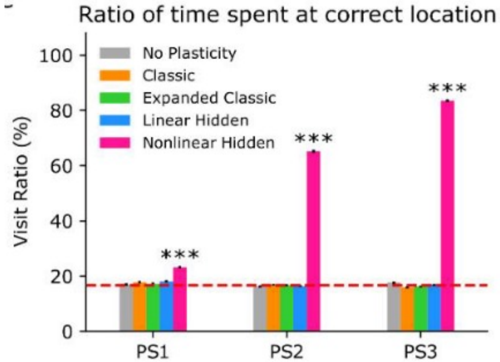
- Local information
- Global scalar modulation

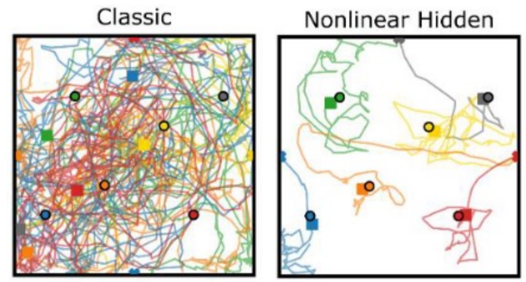
Foster et al. (2000) Hippocampus; Fremaux et al. (2013) Plos Comp. Bio.

Nonlinear hidden layer is necessary to learn MPA



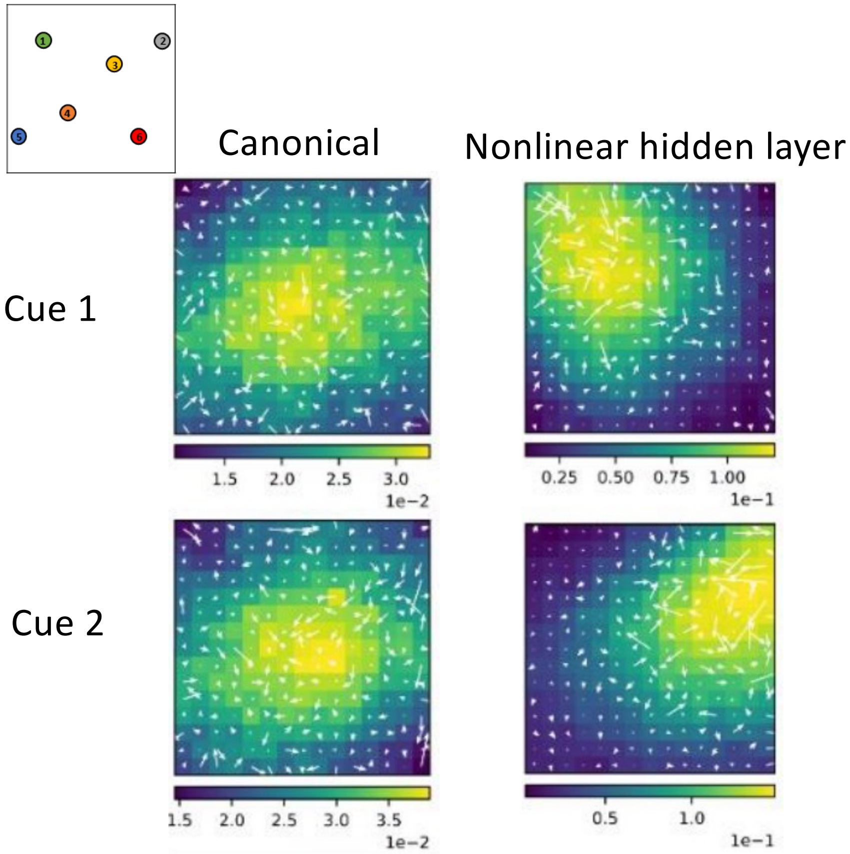


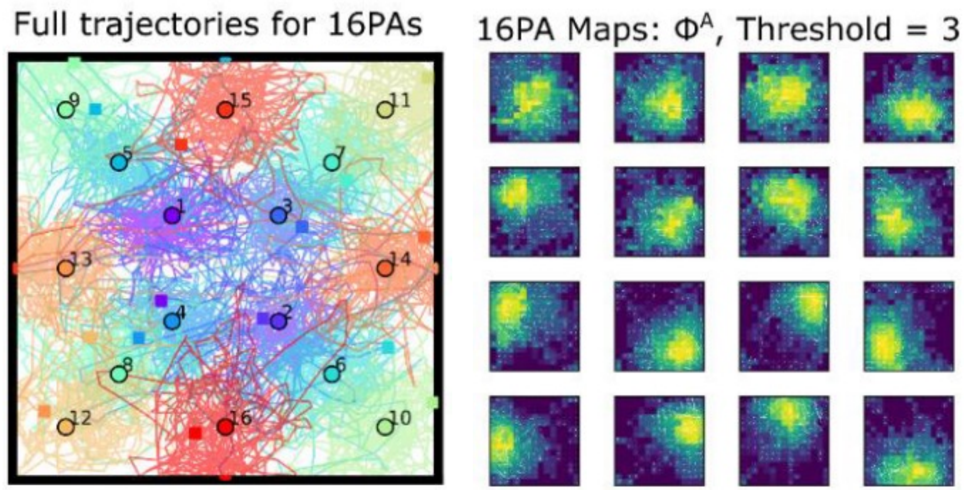




Kumar et al. (2022) Cerebral Cortex

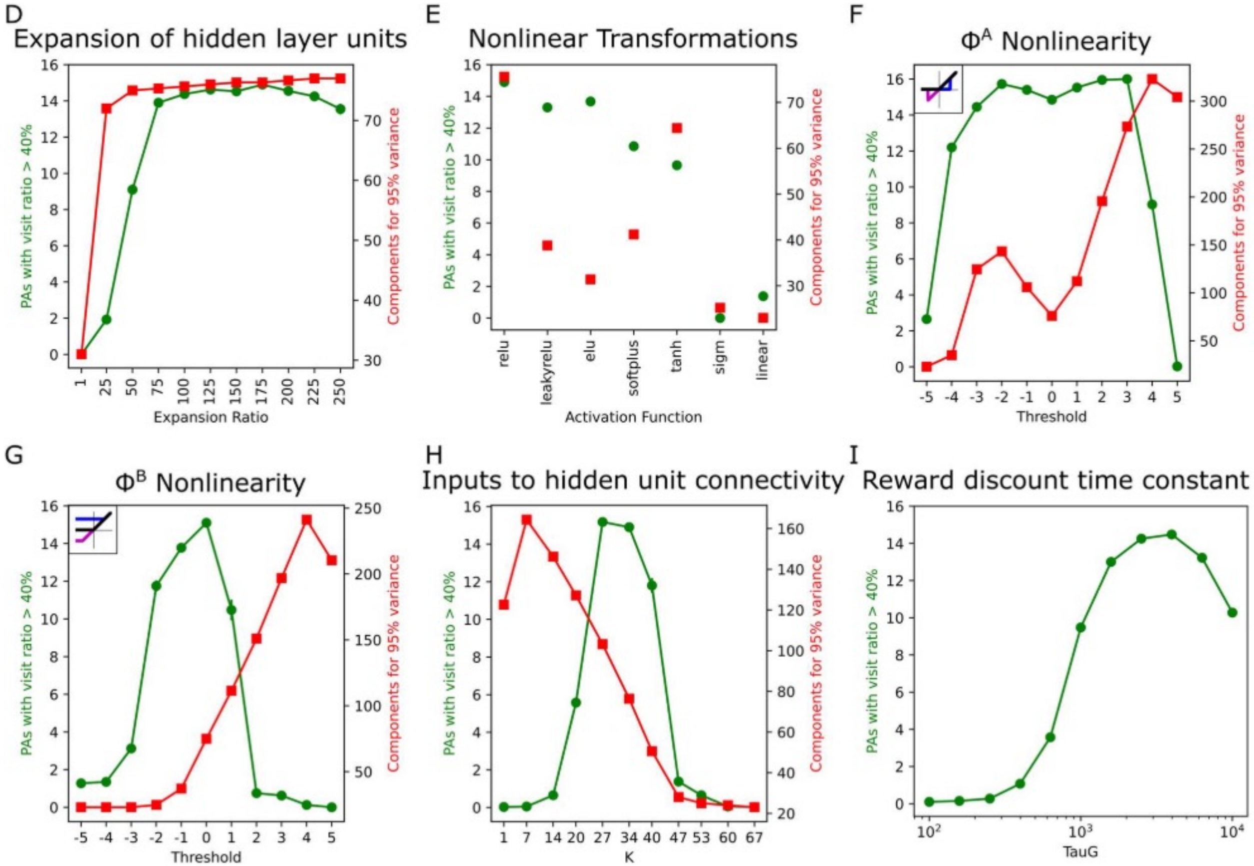
Agents with nonlinear hidden layer gradually learn cue specific value maps and policies





Kumar, et al. (2022) Cerebral Cortex

Influence of hyperparameters on learning multiple paired associations

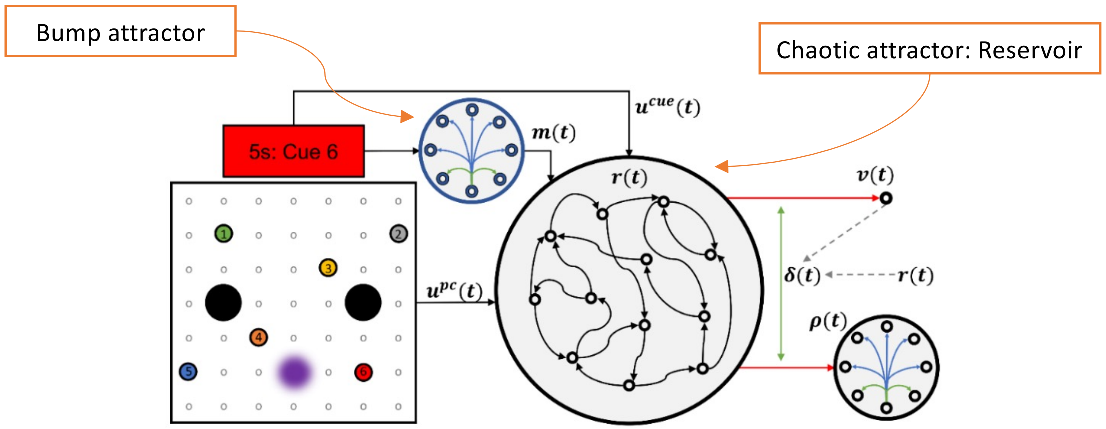


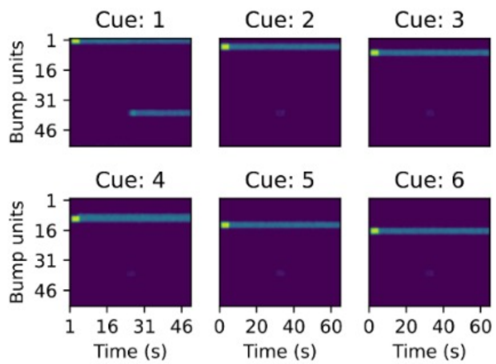
$$\phi^A(x,\theta) = \begin{cases} 0, & x \le \theta \\ x, & x > \theta \end{cases}$$

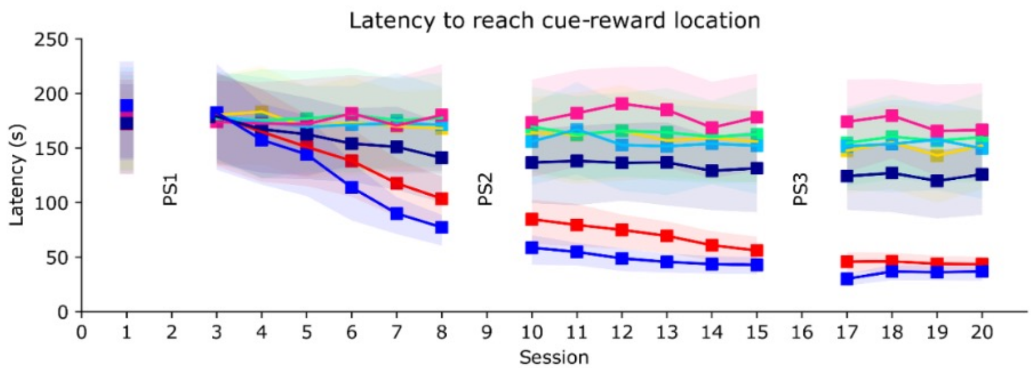
$$\phi^B(x,\theta) = \begin{cases} \theta, & x \le \theta \\ x, & x > \theta \end{cases}$$

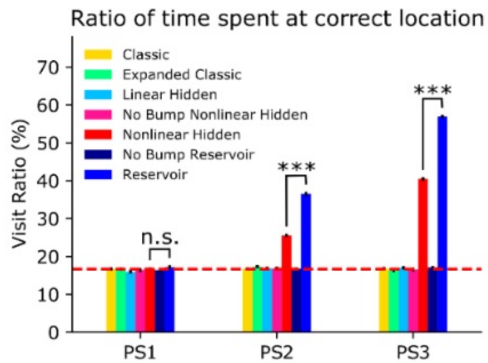
Kumar, et al. (2022) Cerebral Cortex; Litwin-Kumar & Sompolinsky (2017) Neuron

Bump attractor maintains cue to gradually learn MPA



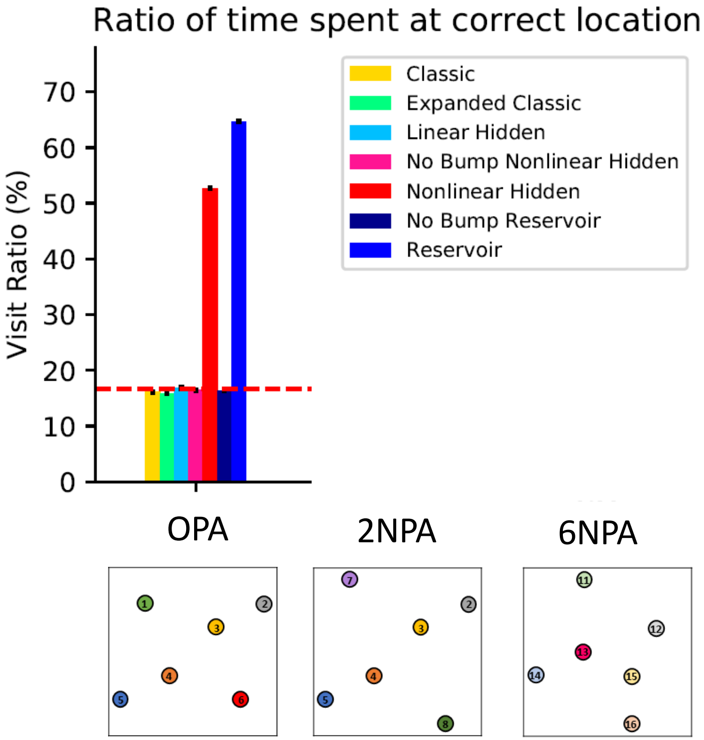


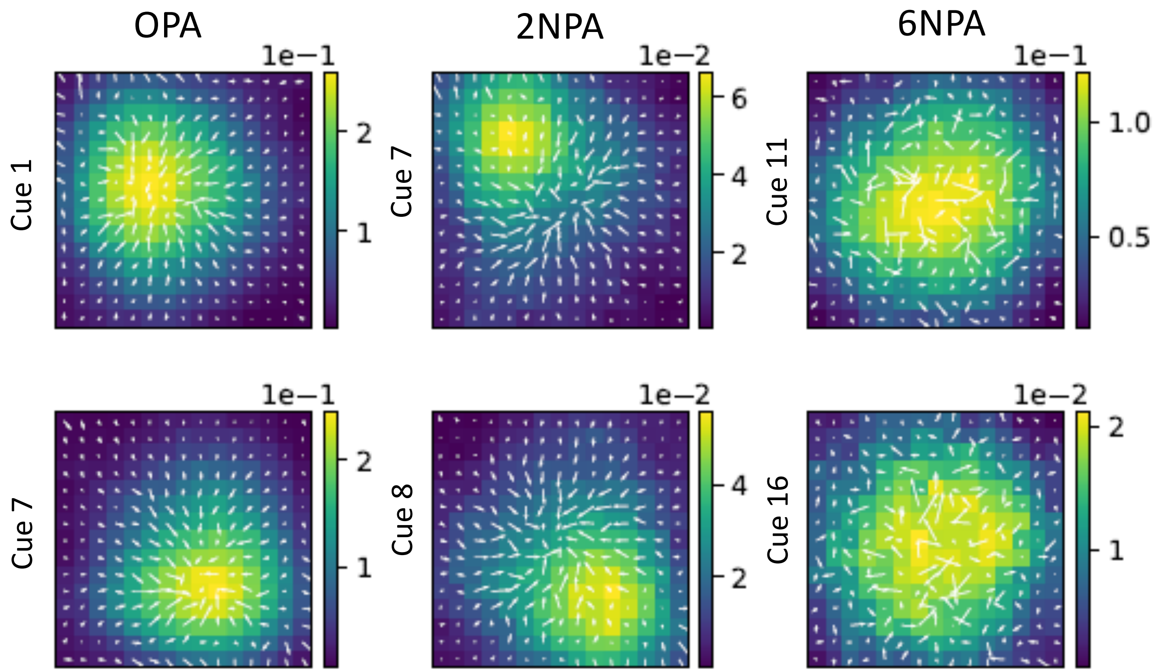




Kumar, et al. (2022) Cerebral Cortex

Actor-Critic and TD-HL cannot learn policies for new paired associations (PA) after one trial

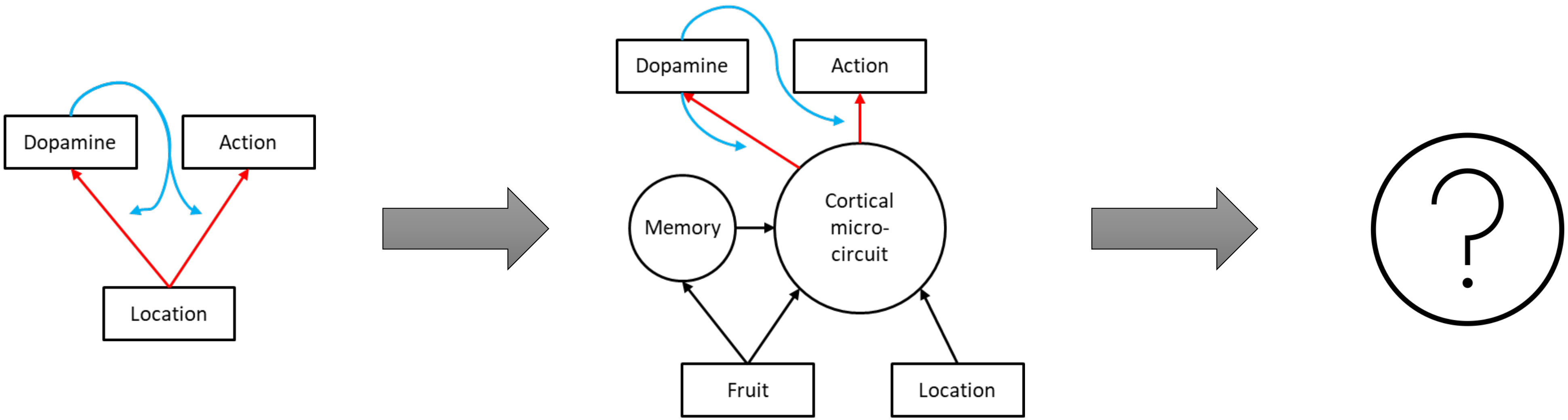




Kumar et al. unpublished.

Interim conclusion 1

- A single nonlinear feedforward or reservoir helps to pre-processes inputs into suitable representations for biologically plausible actor-critic to gradually learn MPA
- Spiking network with a single hidden layer, trained using STDP rule gradually learned MPA
- Proposed actor-critic agent does not perform one-shot learning of new PAs like the rats



Gradually learns single targets

Gradually learns multiple targets

Rapidly learns multiple targets

One-shot learning of new PAs

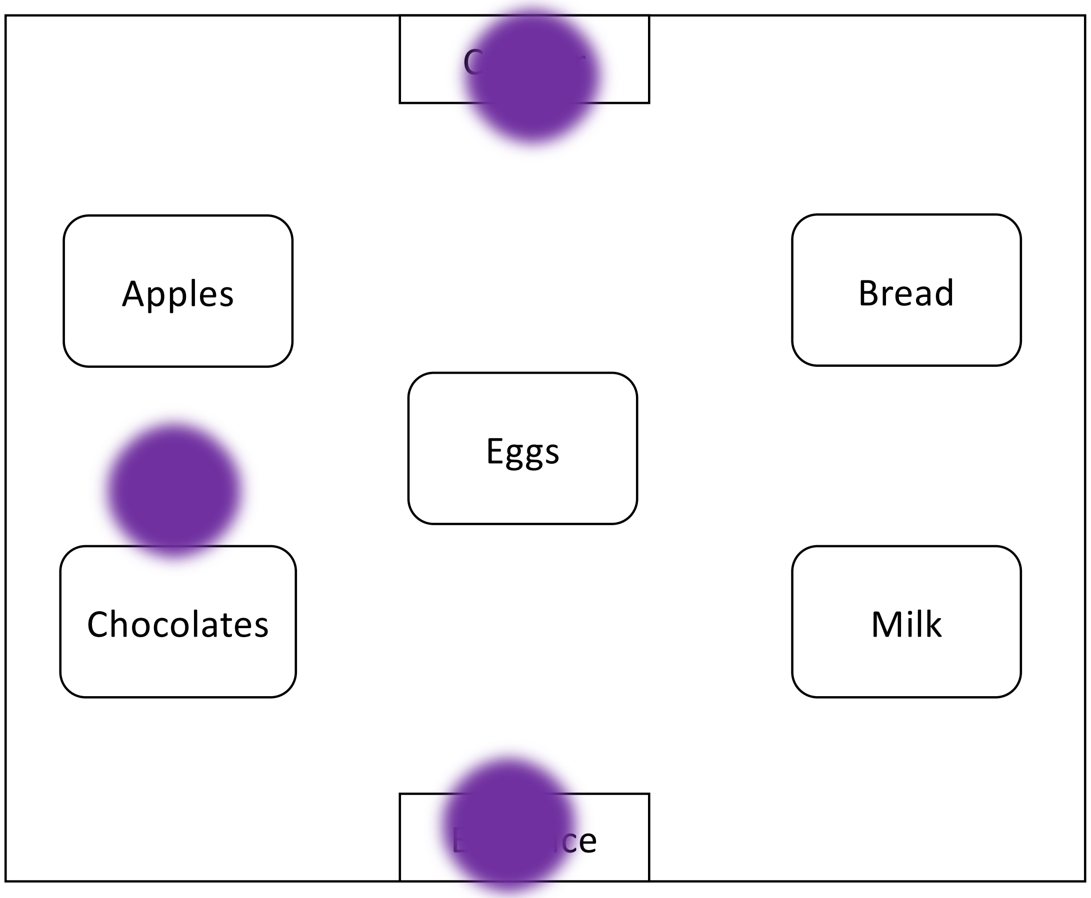
What are the neural architectures and biologically plausible learning rules for one-shot learning?

How do we get milk in a familiar market?

<u>Grocery list</u>

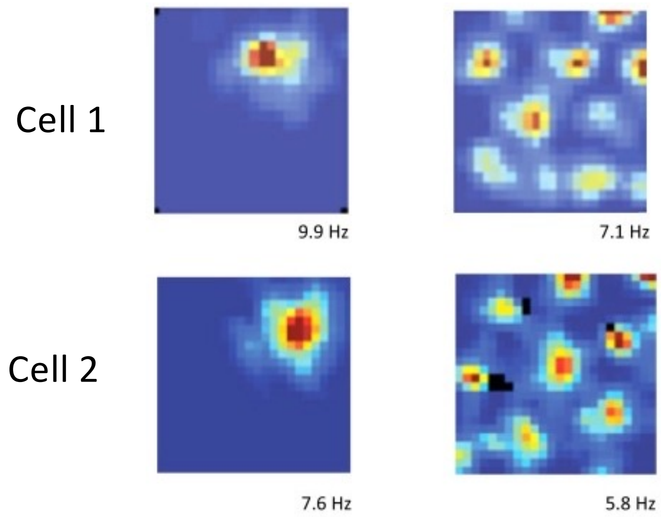
- Milk

1) **GPS**: Where are we in the environment?
2) **Store Goals** : Where is the milk?
3) **Navigation**: How do I move to the milk?

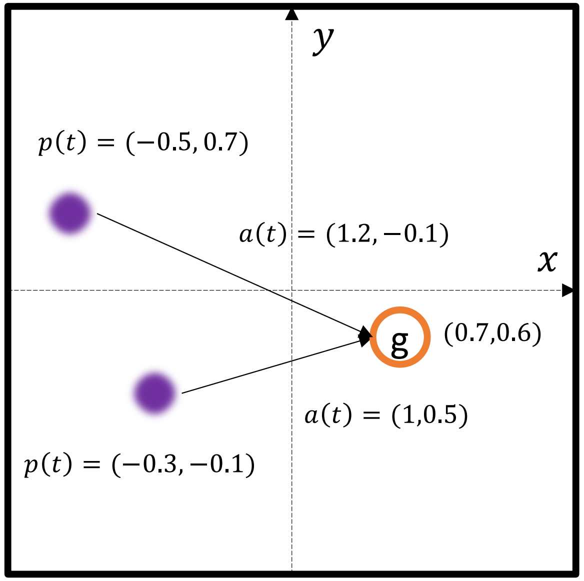


GPS: Continuous metric representation for vector-based navigation

Binned place and grid cell representation not a suitable code for vector-based navigation



Metric representation allows vector subtraction for efficient vector-based navigation

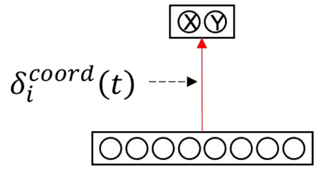


Foster et al. (2000); Bush et al. (2015); Kumar et al. (2023) arXiv.

GPS: Gradually learn to self-localize in a new environment using path integration TD error

LEARN METRIC REPRESENTATION

Current coordinates $p_i(t)$



Place cells $pc_j(t)$

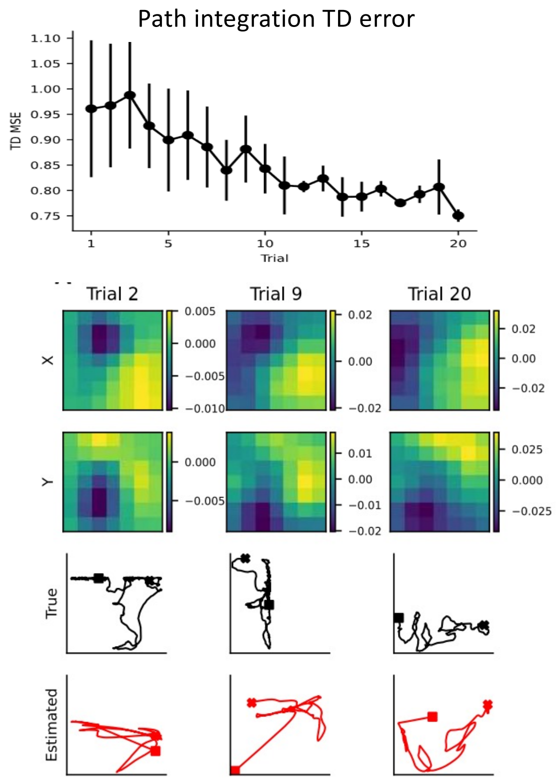
Path Integration

$$p_{i\in\{x,y\}}(t+\Delta t) = p(t) + a_{i\in\{x,y\}}(t)$$

Path integration TD error

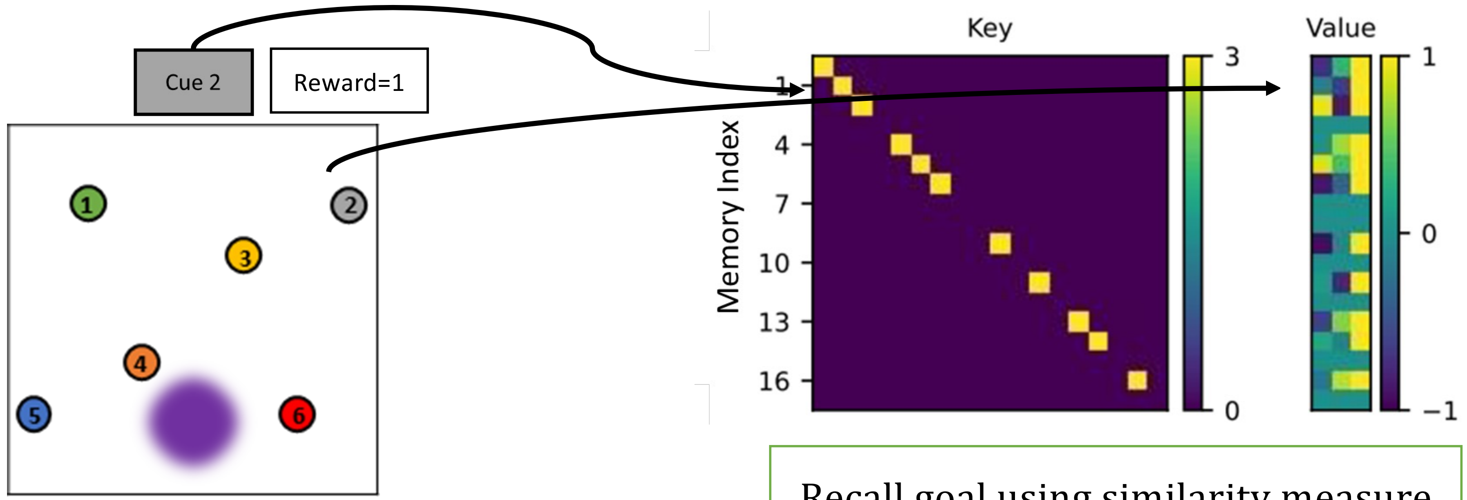
$$\delta_i^{coord}(t) = p(t+\Delta t) - p(t) - a(t)$$

$$\Delta W_{ij}^{coord}(t) \propto \underbrace{e_j(t)}_{\text{Presynaptic}} \cdot \underbrace{\delta_i^{coord}(t)}_{\text{Scalar signal for X \& Y axis}}$$



Foster et al. (2000); Kumar et al. (2023) arXiv.

Store Goals: Key-Value matrix rapidly stores and recalls cue-coordinate associations



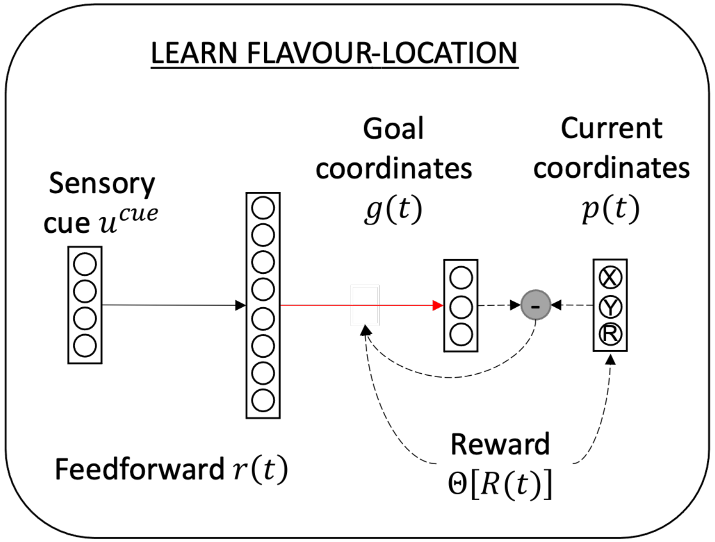
Recall goal using similarity measure

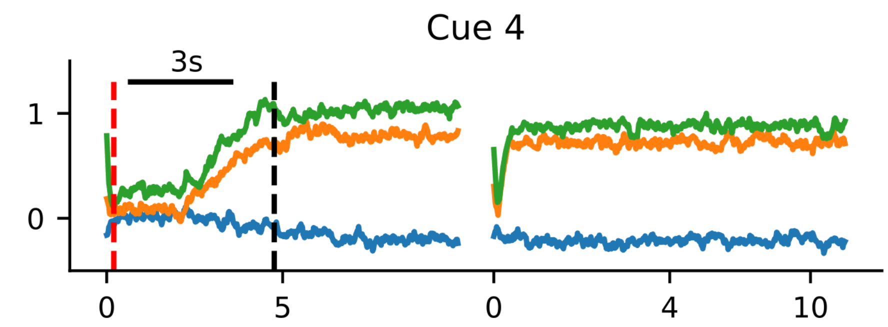
$$A(t) = softmax(Q(t)K^T)$$

$$g(t) = A(t) \cdot V^T$$

Graves et al. (2014) arXiv; Kumar et al. (2023) arXiv.

Store Goals: Reward-modulated Exploratory Hebbian rule for one trial cue-coordinate association

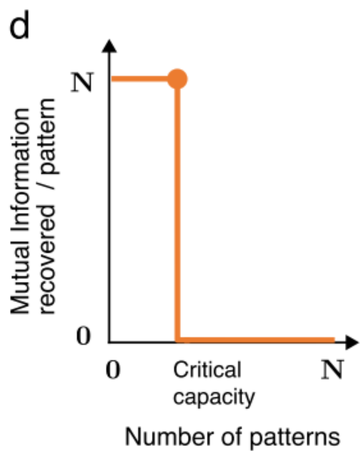


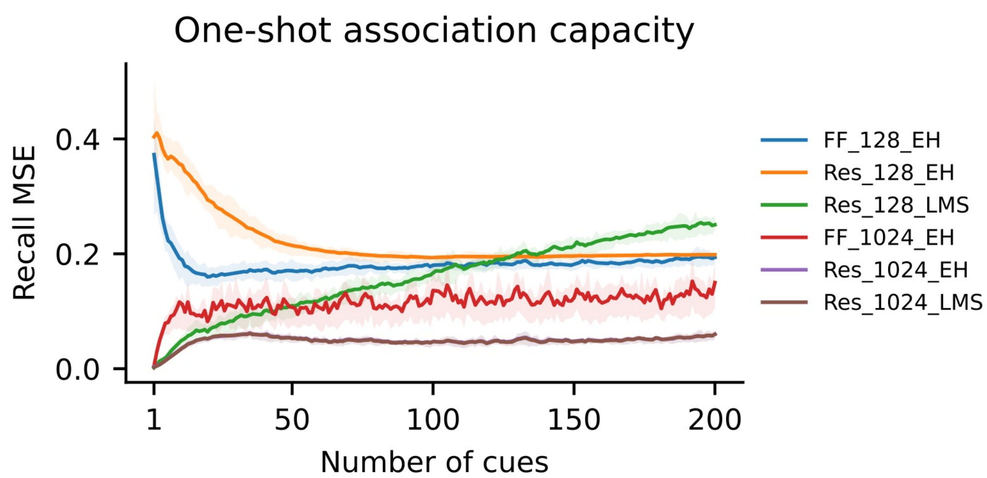


$$\boxed{\text{Reward modulated Exploratory Hebbian rule: } \Delta W_{ij}^{goal}(t) \propto \overbrace{r_j(t)}^{\text{Presynaptic}} \cdot \underbrace{\left(g_i^{noisy}(t) - g_i(t)\right)}_{\text{Postsynaptic}} \cdot \overbrace{M(t)}^{\text{Scalar signal } M(t) = \sum[g(t)-p(t)]^2} \cdot \underbrace{\Theta(R(t))}_{\text{Gate plasticity with reward}}}$$

Hoerzer et al. (2012); Kumar et al. (2023) arXiv.

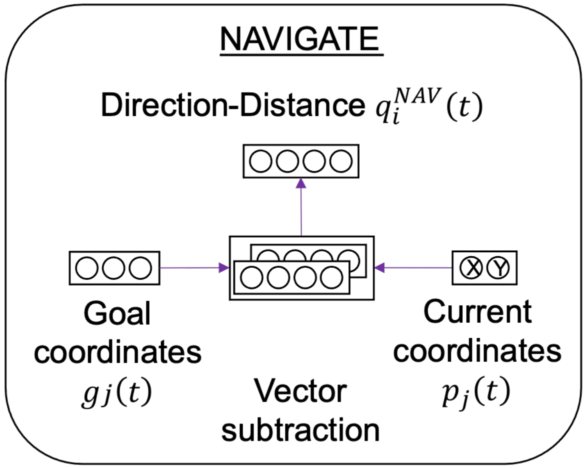
Store Goals: Influence of hidden layer and learning rule for one-shot association





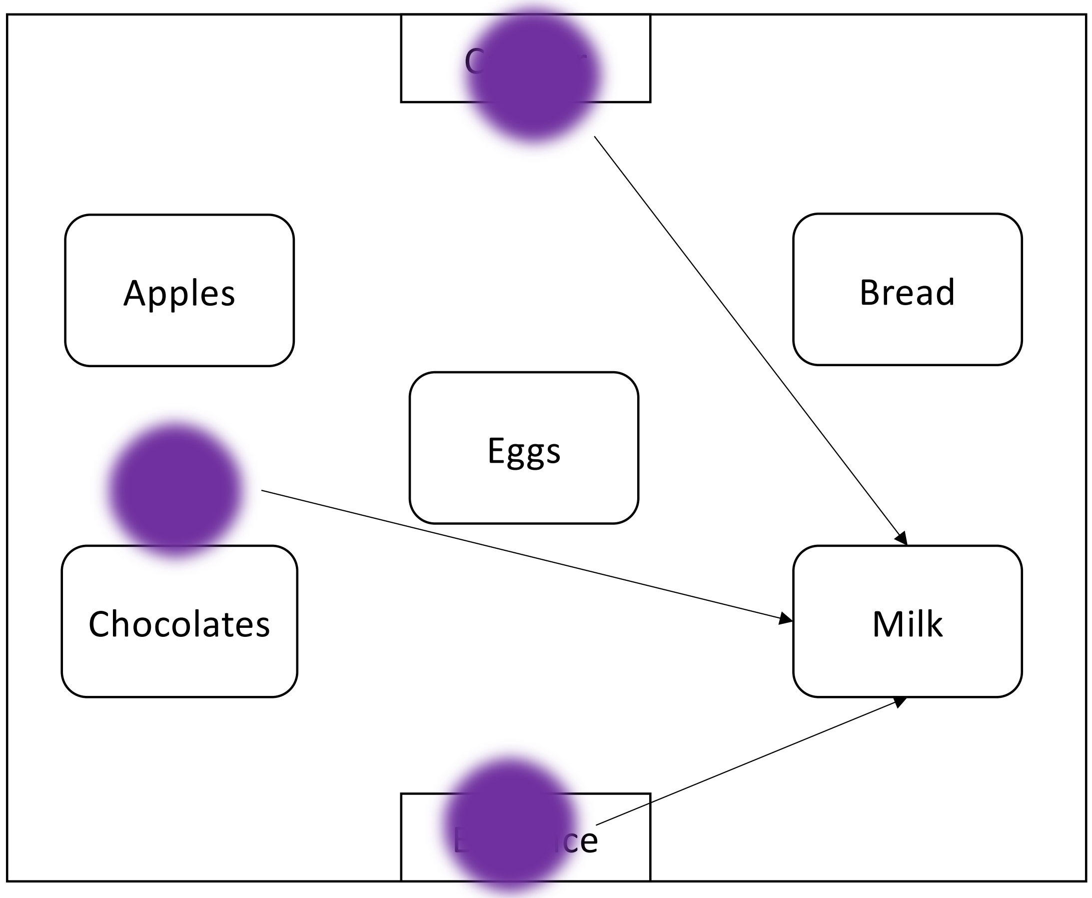
Sharma, Fiete (2022); Kumar et al. (2023) arXiv.

Navigate: Vector subtraction and choose direction to move from any location to goal



$$d(t) = g(t) - p(t)$$

*Pretrained network to learn NAVIGATE
& fix weights during task



Sarel et al. (2017); Banino et al. (2018); Kumar et al. (2023) arXiv.

Three biologically plausible schemas for one-shot navigation

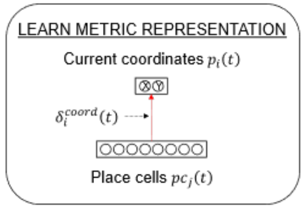
- First examples of theoretical schemas
- Implemented using biologically plausible networks and learning rules
- Possible implementation in the brain
- Each network describes the computation, algorithm and implementation levels

A) Learning a continuous metric representation of the environment (LEARN METRIC REPRESENTATION)

Algorithm 2 LEARN METRIC REPRESENTATION schema pseudocode uses place cell activity to learn a continuous X,Y coordinate metric representation in any environment for vector navigation.

Neural implementation of LEARN METRIC REPRESENTATION. Place cells synapse to X, Y coordinate cells while synapses are learned using the path integration temporal difference error modulated Hebbian plasticity rule with eligibility trace.

Initialise network weights $W^{coord} \leftarrow 0$
Initialise start coordinates $p(0) = (0,0)$
for t < T:
    Move in direction specified by policy
    Estimate coordinates using place cells (18)
    Compute path integration error (21)
    Compile history of place cell activity (22)
    Minimise path integration error (23)
End

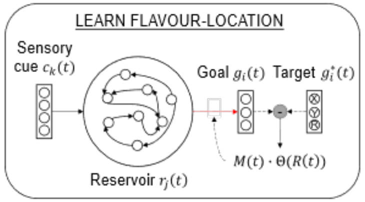


B) Associate flavour cue to location coordinates to recall goal after one-trial (LEARN FLAVOUR–LOCATION)

Algorithm 3 LEARN FLAVOUR-LOCATION schema pseudocode takes in any flavour cue as input to associate or recall goal coordinates and recall value of 1 as output.

Neural implementation of LEARN FLAVOUR-LOCATION. Cue vector passed as input to the reservoir with three readout units representing goal coordinates and recall value. Only the synapses from the reservoir to the readout units are learned using 4-factor reward gated Exploratory Hebbian rule.

Initialise associative memory
Get cue from environment
for t < T:
    Pass cue to network to recall goal (27)
    if recall value > threshold **then**
        Navigate to goal using NAVIGATE
    else
        Explore maze using random policy
        if reward obtained **then**
            Switch on plasticity (30)-(32) to associate flavour cue to current coordinates estimated using metric representation (28)
End

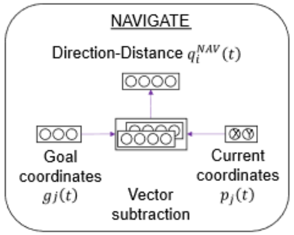


C) Move from current location to goal location (NAVIGATE)

Algorithm 1 NAVIGATE schema pseudocode takes in any current and goal coordinates with recall value as input to output direction to move.

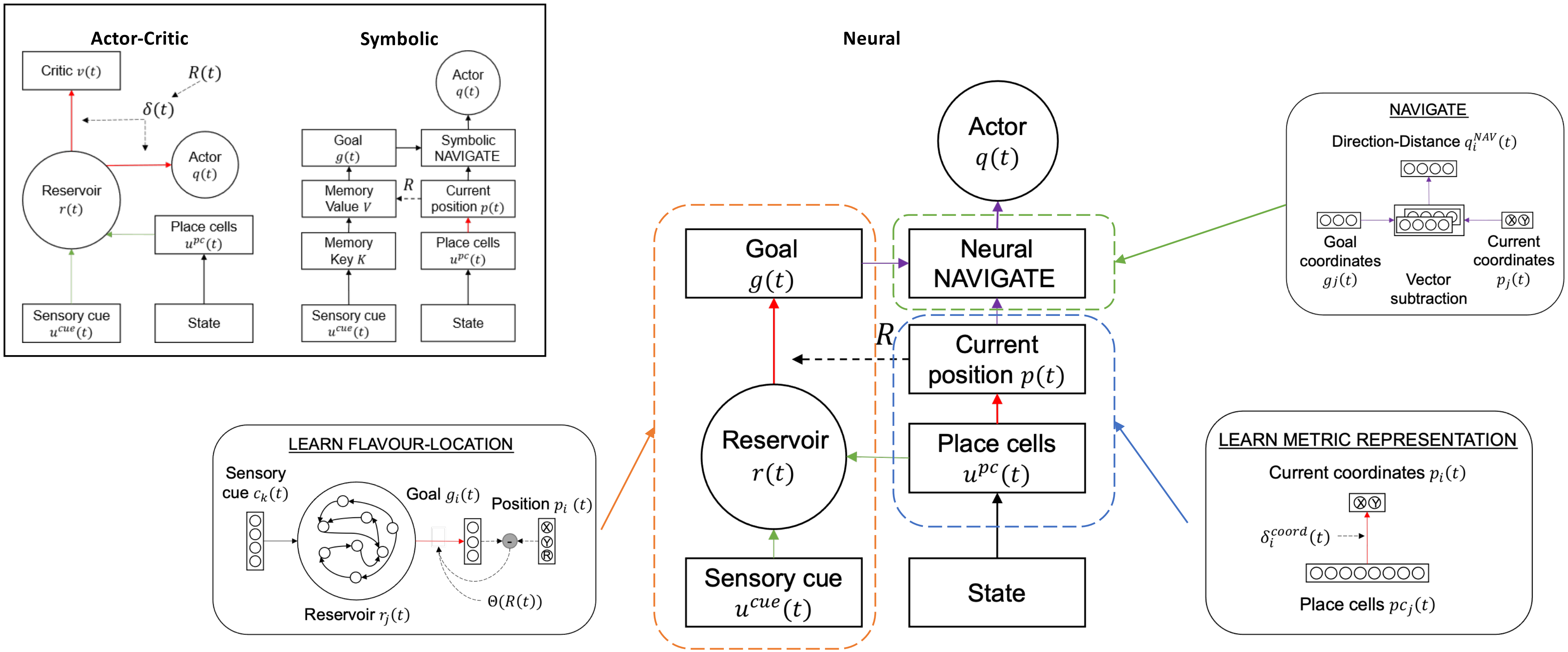
Neural implementation of NAVIGATE. Two hidden layered neural network was pre-trained by backpropagation on a dataset with different current coordinate, goal coordinate, recall value and actions. Weights were fixed during task learning.

Get agent's current coordinates $p(t)$
Get goal coordinates $g(t)$
for t < T:
    if recall value > threshold (37), **then**
        Compute vector subtraction between current and goal coordinates (35)
        Choose action based on direction specified by vector (36)
    else
        No action selected
End



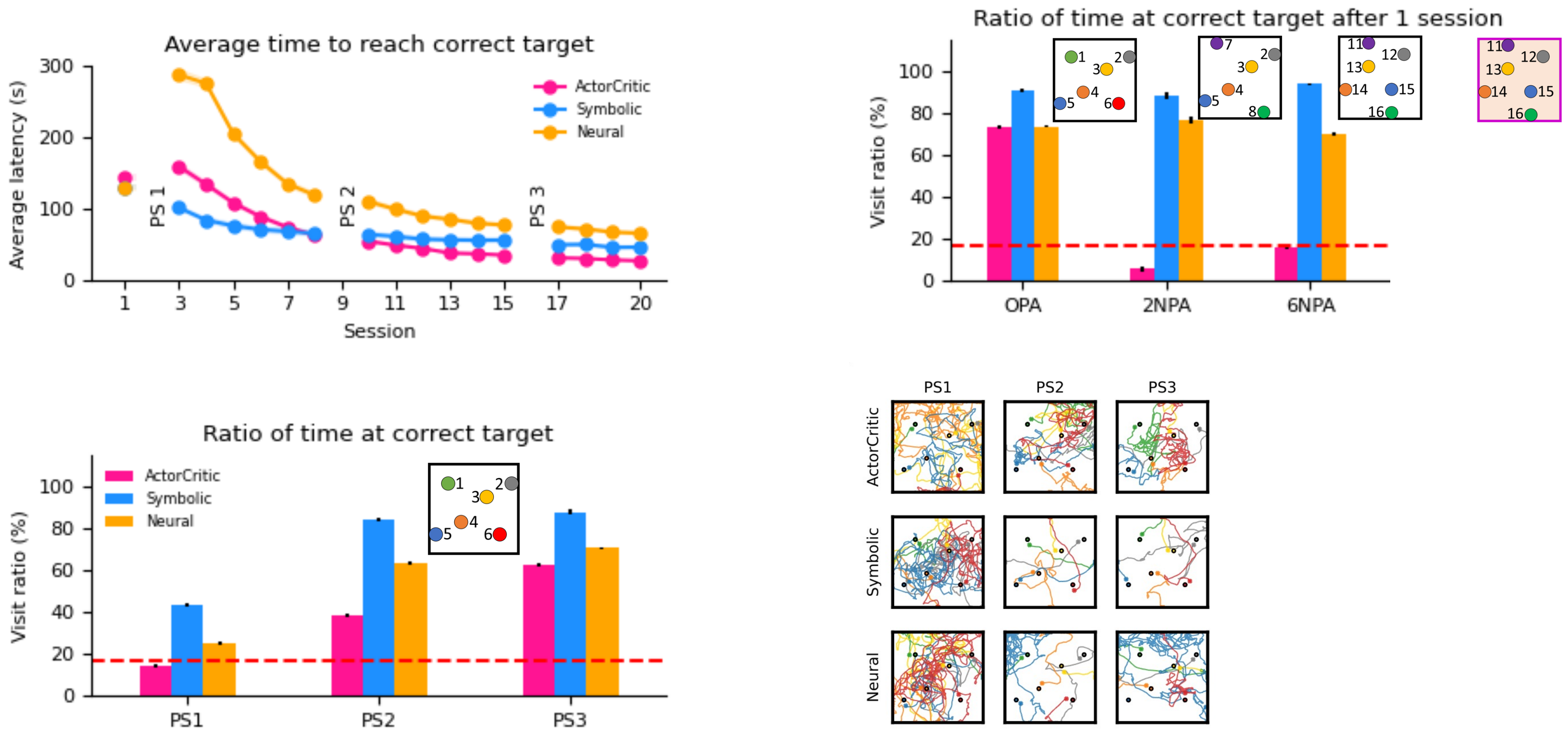
Kumar et al. (2023) arXiv.

Compose three schemas into a fully neural agent for one-shot learning



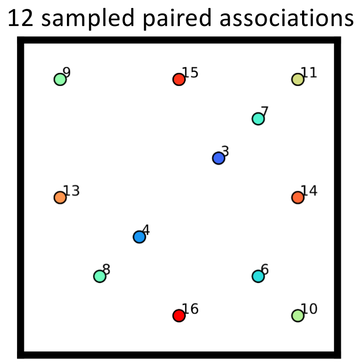
Kumar et al. (2023) arXiv.

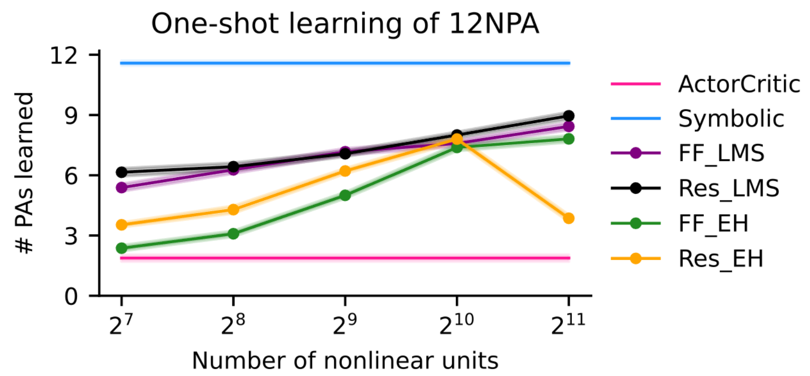
Gradual then one-shot learning of MPA

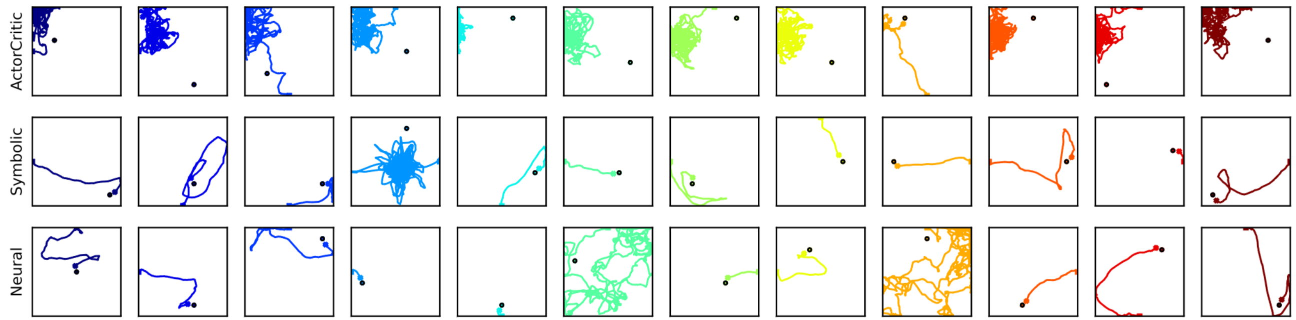


Kumar et al. (2023) arXiv.

One-shot learning capacity of multiple goals

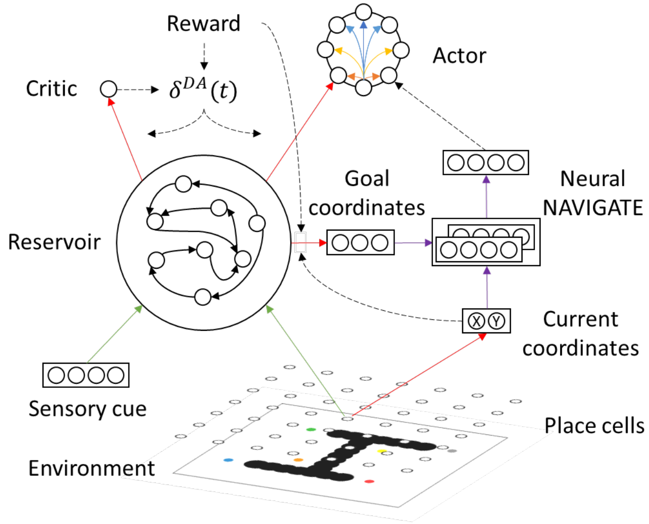


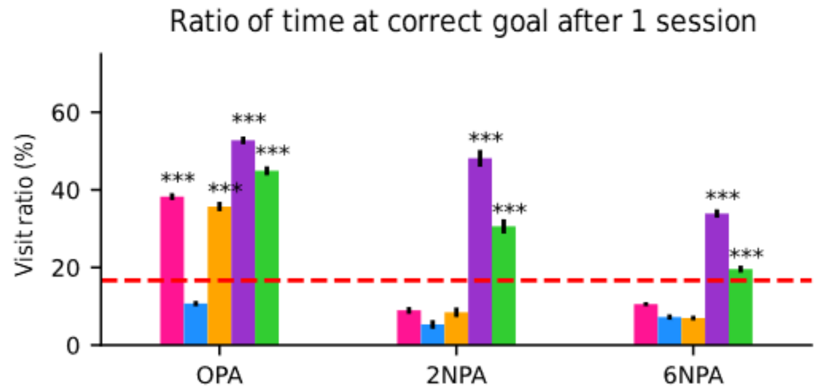


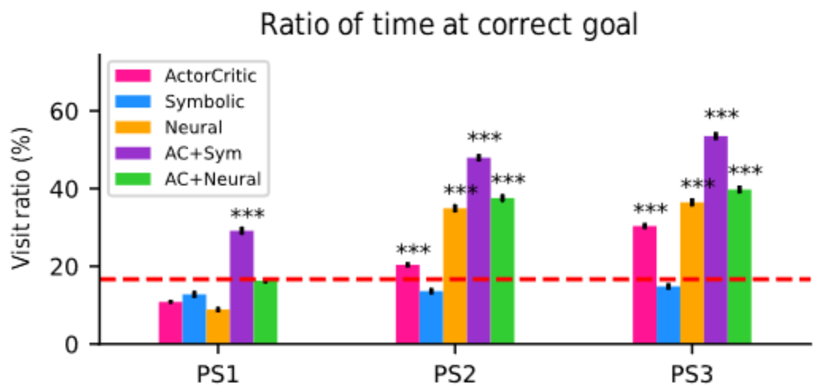


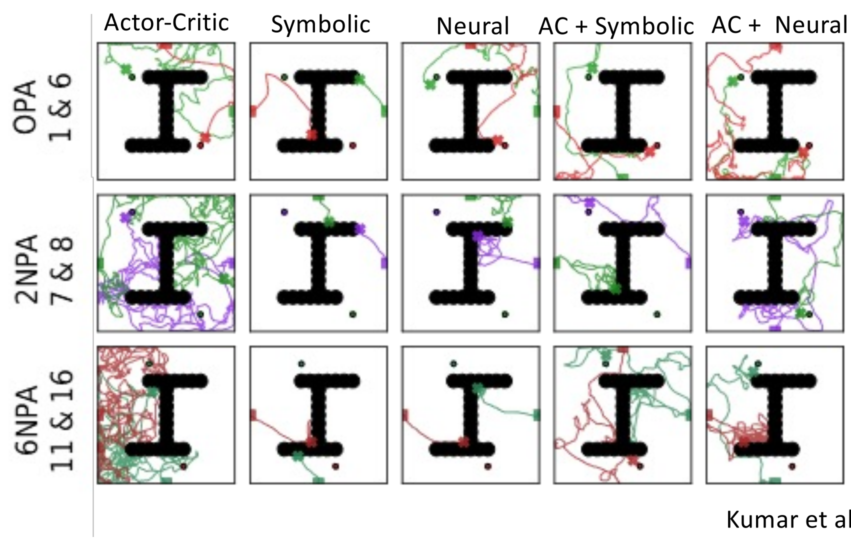
Kumar et al. (2023) arXiv.

One-shot navigation of new PAs past obstacles



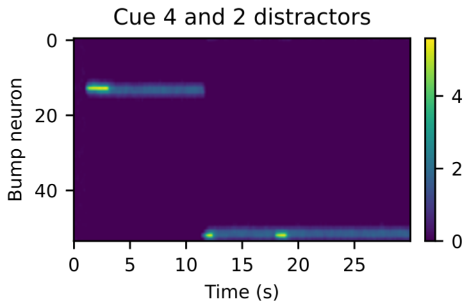




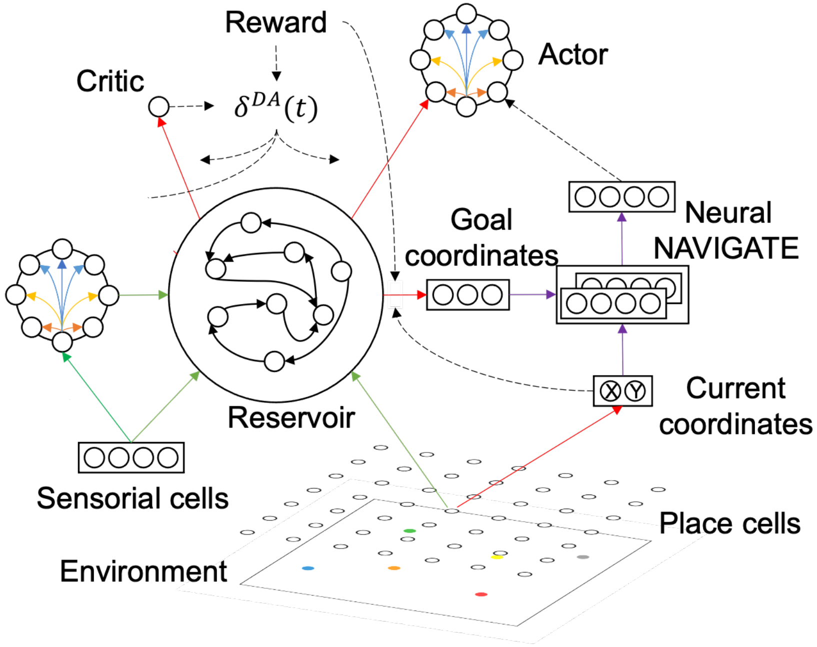


Kumar et al. (2023) arXiv.

Learning to gate working memory using TD-HL

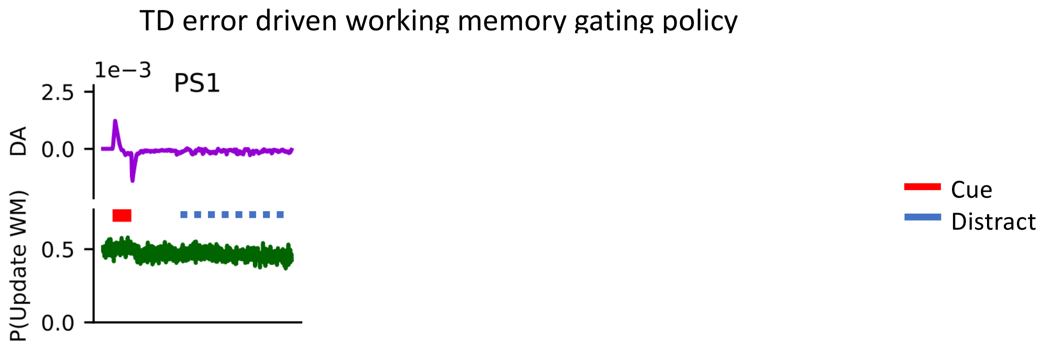


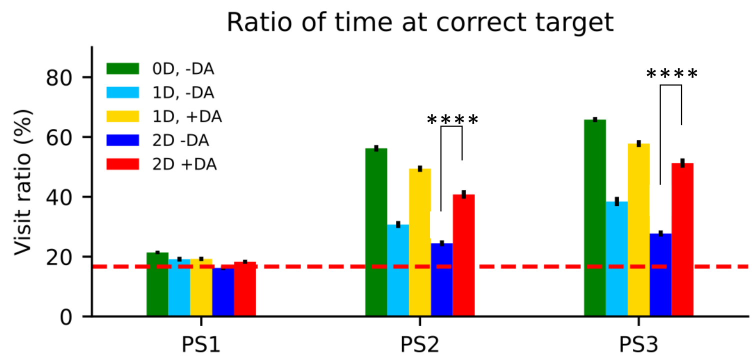
$$\boxed{\Delta W_{ij}^{gate} = \underbrace{r_j(t)}_{\text{Presynaptic}} \cdot \overbrace{\pi_i(t)}^{\text{Postsynaptic}} \cdot \underbrace{\delta^{DA}(t)}_{\text{Scalar signal}}}$$

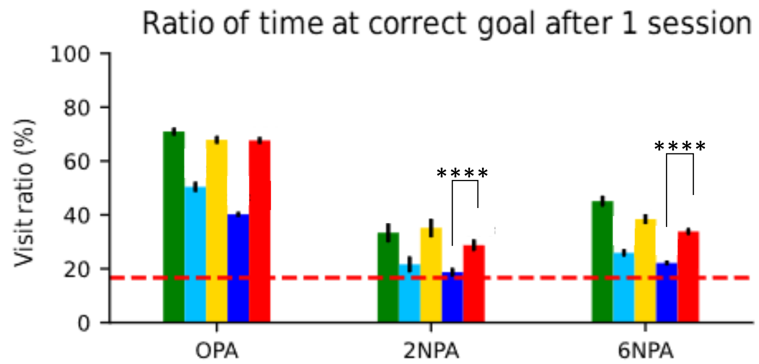


O'Reilly & Frank (2006); Lloyd et al. (2012); Kumar et al. (2023) arXiv.

Gradually learned gating policy generalises to new PAs



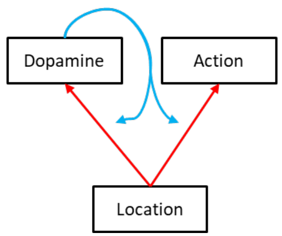




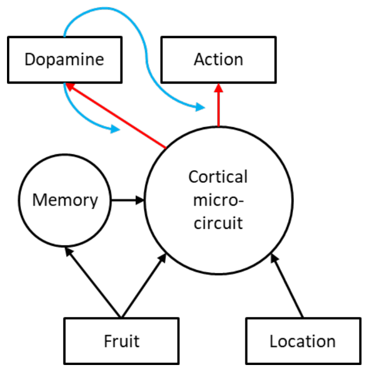
Kumar et al. (2023) arXiv.

Interim conclusion 2

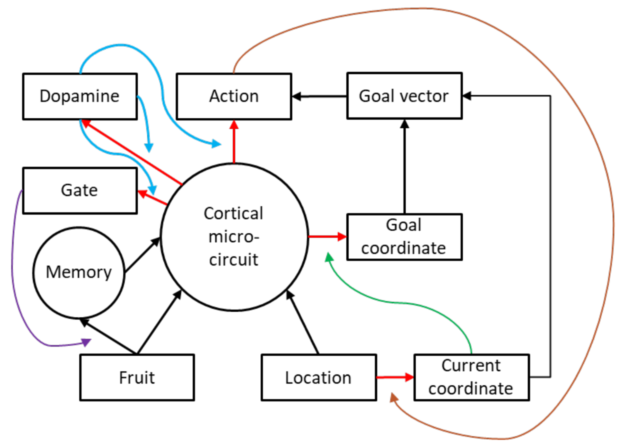
- Each modular schema requires different architecture and biologically plausible learning rules
- Composing 3 biologically plausible schemas achieve one-shot learning of multiple new PAs
- Combining schema agents with actor-critic and working memory gating components increased its robustness to other conditions



Gradually learn single targets



Gradually learn multiple targets



Rapidly learn multiple targets

Limitations & future directions

Biological plausibility

- Leaky rate coded neurons and Hebbian rule are only approximations
- Spiking neurons, Spike Time Dependent Plasticity, Neuromodulators etc.

Anatomical mapping

- Not constrained to neural data or brain regions
- Loosely mapped to Prefrontal – Hippocampus – Striatum
- Insufficient experimental evidence

Generalizability to other tasks

- Sequential navigation to multiple subgoals
- Hierarchical task structure instead of paired associations
- Cognitive control tasks such as rule reversal and task switching

Memory consolidation

- Hippocampal episodic memories to semantic memories in the cortex was not shown
- No established algorithm for consolidation

DetermiNet: A Large-Scale Dataset for Complex Visually-Grounded Referencing using Determiners

Clarence Lee*, M Ganesh Kumar*, Cheston Tan

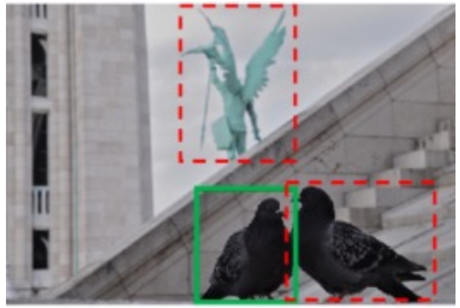
Centre for Frontier AI Research (CFAR)

Agency for Science, Technology and Research (A*STAR)

Submitted to International Conference on Computer Vision (ICCV) 2023

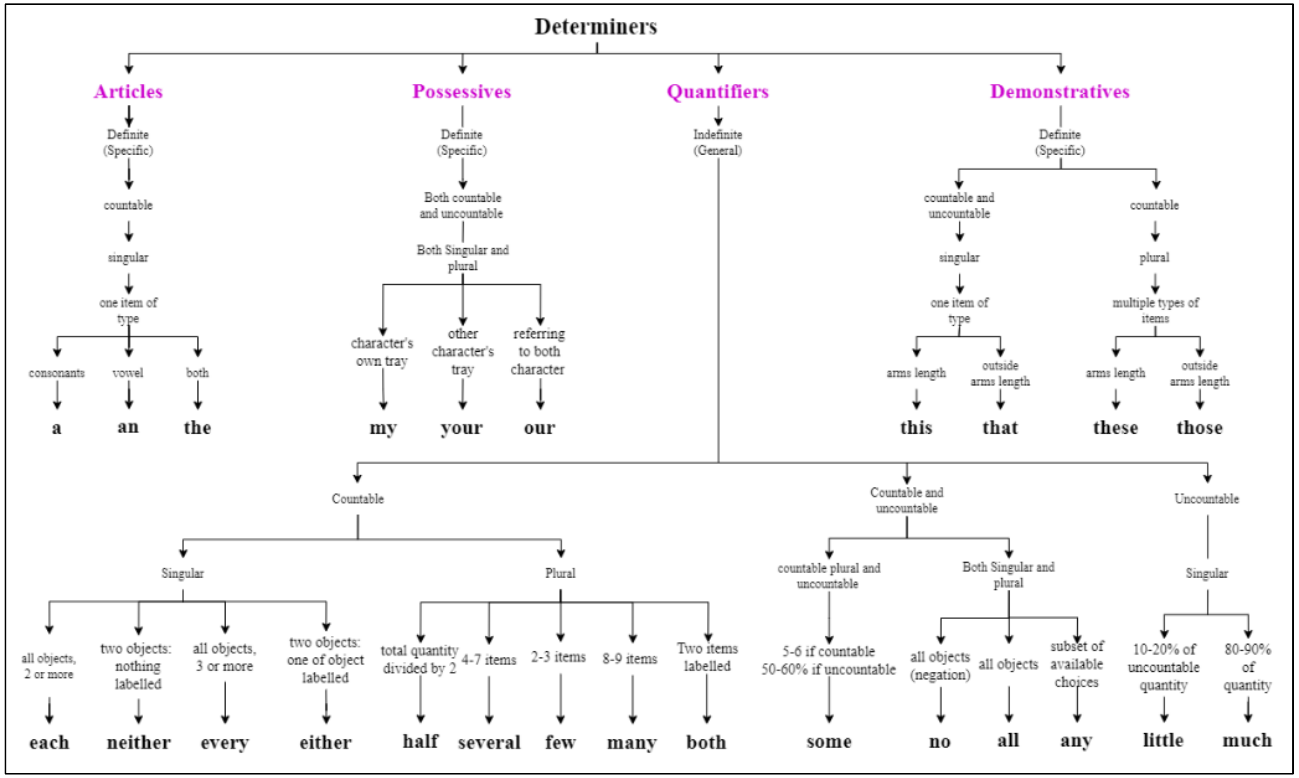
The need for Determiners

- Referring Expression Comprehension (REC) task:
    - Models identify the correct object given a caption-based query
    - “Bird below the blue statue”
    - MDETR (100 % on CLEVR-Ref+), OFA (94.03 % on RefCOCO)
- However, natural language vocabulary used in computer vision datasets is limited.



Lee, Kumar, Tan (2023) under review.

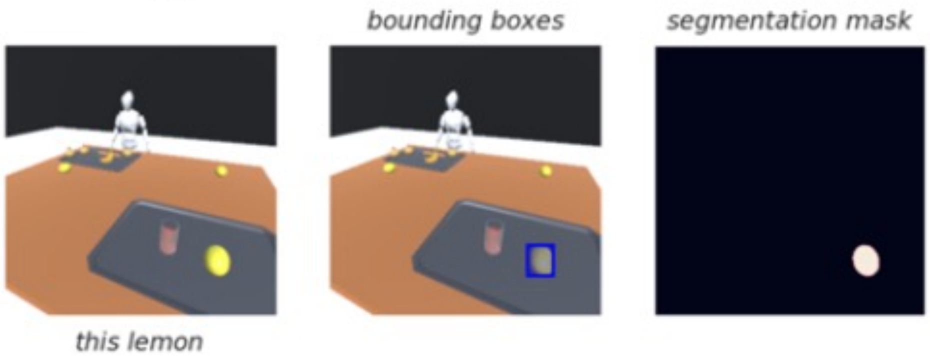
Organisation of determiners



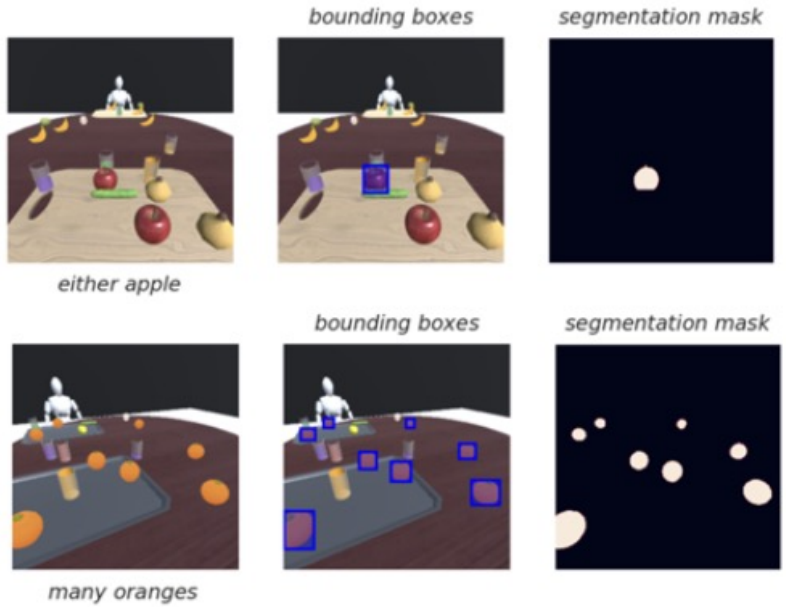
Narrowed down to 25 determiners from 44 determiners found in British National Corpus (57% coverage)

Lee, Kumar, Tan (2023) under review.

Complexities in DetermiNet



Multiple correct solutions



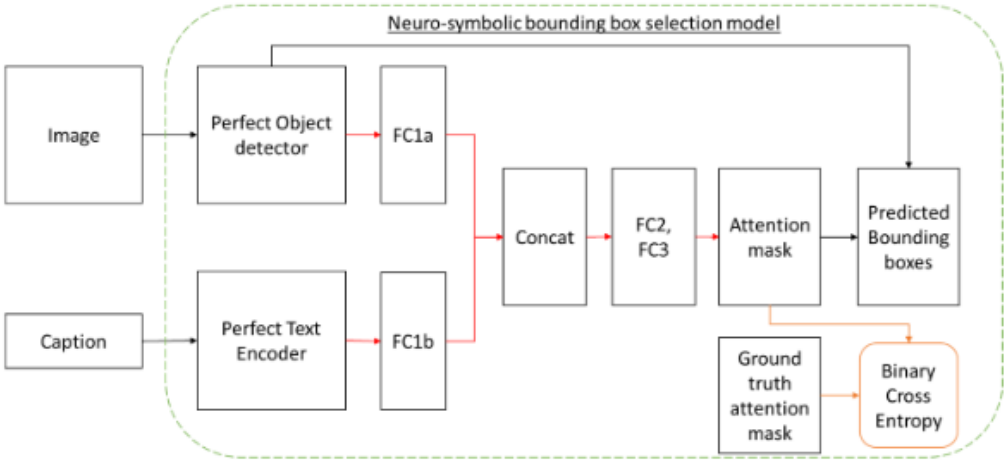
- Human point-of-view at workbench environment with robot
- 15 objects (countable vs uncountable), 3 POV, 3 tray positions, etc.
- **10,000** variations per determiner X **25** determiners = **250,000** samples

- Ground truth correction function modifies ground truth bounding boxes according to determiner rules to handle multiple correct solutions

Lee, Kumar, Tan (2023) under review.

Oracle neuro-symbolic model

- Oracle (Upper bound)
  - Perfect object detector extracts all bounding boxes of all N objects in image
  - Perfect text encoder encodes determiners and nouns as two sets of one-hot vectors
  - Neural network selects n bounding boxes by learning an attention mask



Lee, Kumar, Tan (2023) under review.

SOTA models struggle to solve DetermiNet

Table 4. Model performance after correcting ground truth annotations. *OFA only predicts one bounding box.

Models	AP@IoU=0.5:0.95
Random	9.8
Oracle	92.5
OFA [28]	20.5*
GLIP [31]	52.0
MDETR [14]	70.7

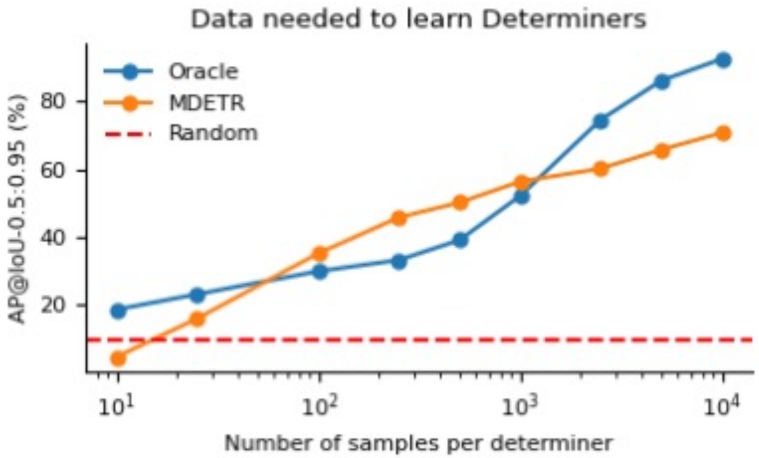
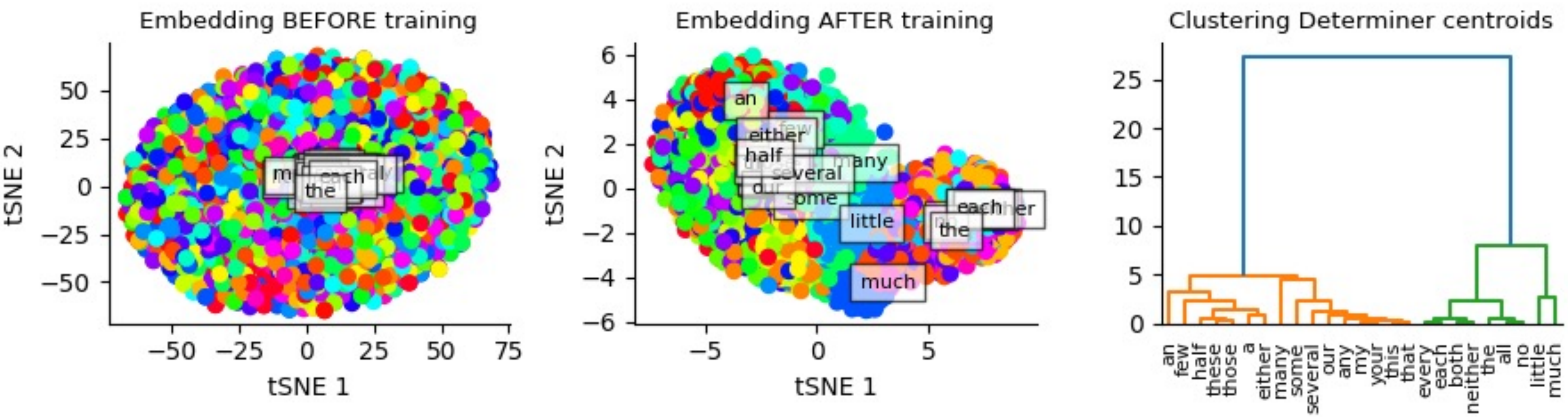


Figure 6. Dataset efficiency to learn determiner rules

Lee, Kumar, Tan (2023) under review.

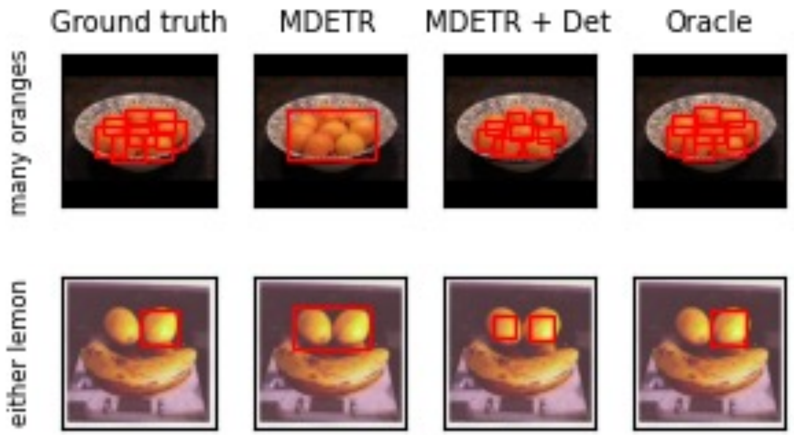
Neural networks can learn most determiners



Visualization of determiner representation in FC3 of the oracle model using t-SNE

Lee, Kumar, Tan (2023) under review.

Determiners transfer from synthetic to real images



100 real image-caption samples from COCO dataset

Table 3. Zero-shot evaluation on real-image dataset

Models (Tasks pretrained on)	**AP@IoU=0.5:0.95**
Neuro-symbolic oracle	73.2
MDETR (RefCOCO)	10.5
MDETR (RefCOCO + DetermiNet)	19.5

Lee, Kumar, Tan (2023) under review.

Interim conclusion 3

- Neural networks learn determiner rules based on the dataset design
- Rule representation can be used to make inference on different datasets
- SOTA models trained end to end are not sample efficient

Points to consider…

1. Neural networks are great to learn representations, but current models are not sample efficient
   - Activation functions: Static nonlinear vs Adaptive thresholds e.g. spikes
   - Architectures: Feedforward vs Recurrent
   - Representations: Static features vs Dynamics + attractor topology

2. Backpropagation is the best optimization technique, but it is not biologically plausible to understand neural computations
   - Error modulated Hebbian plasticity – Supervised learning
   - Hebbian plasticity – Unsupervised/associative learning
   - Anti-Hebbian plasticity/Sanger's rule – Sparse/orthogonal representation & stability
   - Feedback alignment – Train multiple layers
   - Spiking networks + STDP – biologically plausible computations

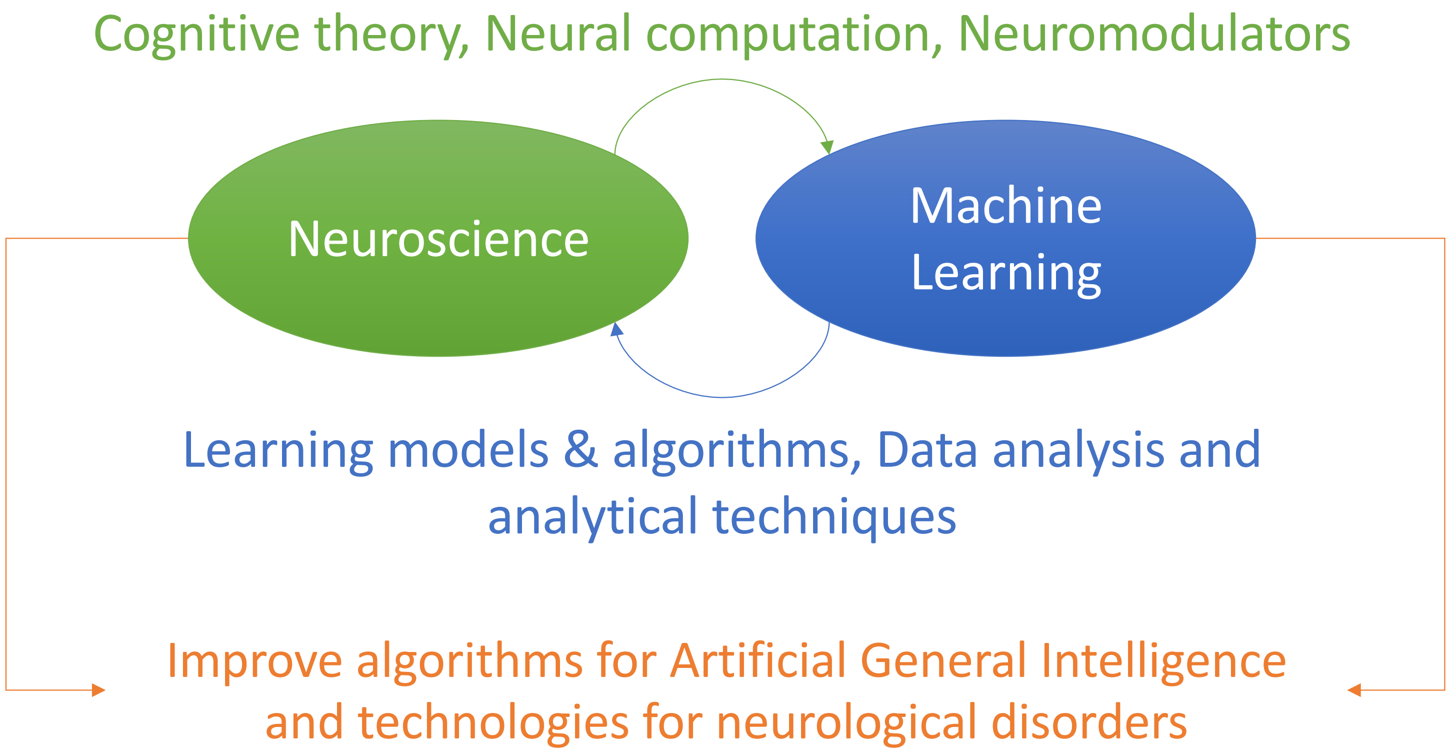
3. Biologically plausible models can solve complex tasks
   - Modular single-layer networks
   - Different learning rules: Reinforcement, supervised, self-supervised
   - Analysis of computation can be tractable

Questions of interest

1. How does the brain represent abstract ideas?
   - How are inductive biases, associations, rules, schemas represented in neural circuits?
   - Are these representation similar between biological, biologically plausible and artificial circuits?

2. How can computation be flexible and generative?
   - How is new information transformed to fit into pre-learned representations quickly?
   - How do networks use a single schema to generate new combinations for planning and reasoning?

Research strategy



Questions?

m_ganeshkumar@u.nus.edu M Ganesh Kumar mgkumar138 @Ganeshk92