DQN Network with Memory Learning a Sequence Task

By: Saqar Sabzali, Maria Kesa
TA: Soan Kim, Amir Mesbah

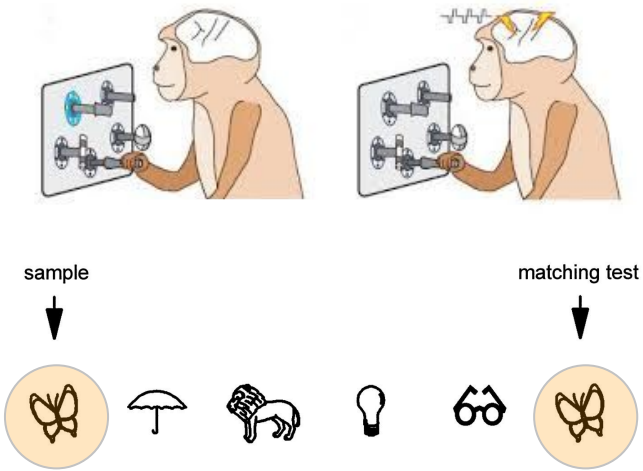
neuromatch
academy

Introduction

- We were interested in comparing artificial learning to biological learning→ We trained a RL model on a cognitive task

- Delayed Match to Sample Task

- Biological organisms obey the Matching Law:



$$\frac{C_1}{C_2} = \frac{R_1}{R_2}$$

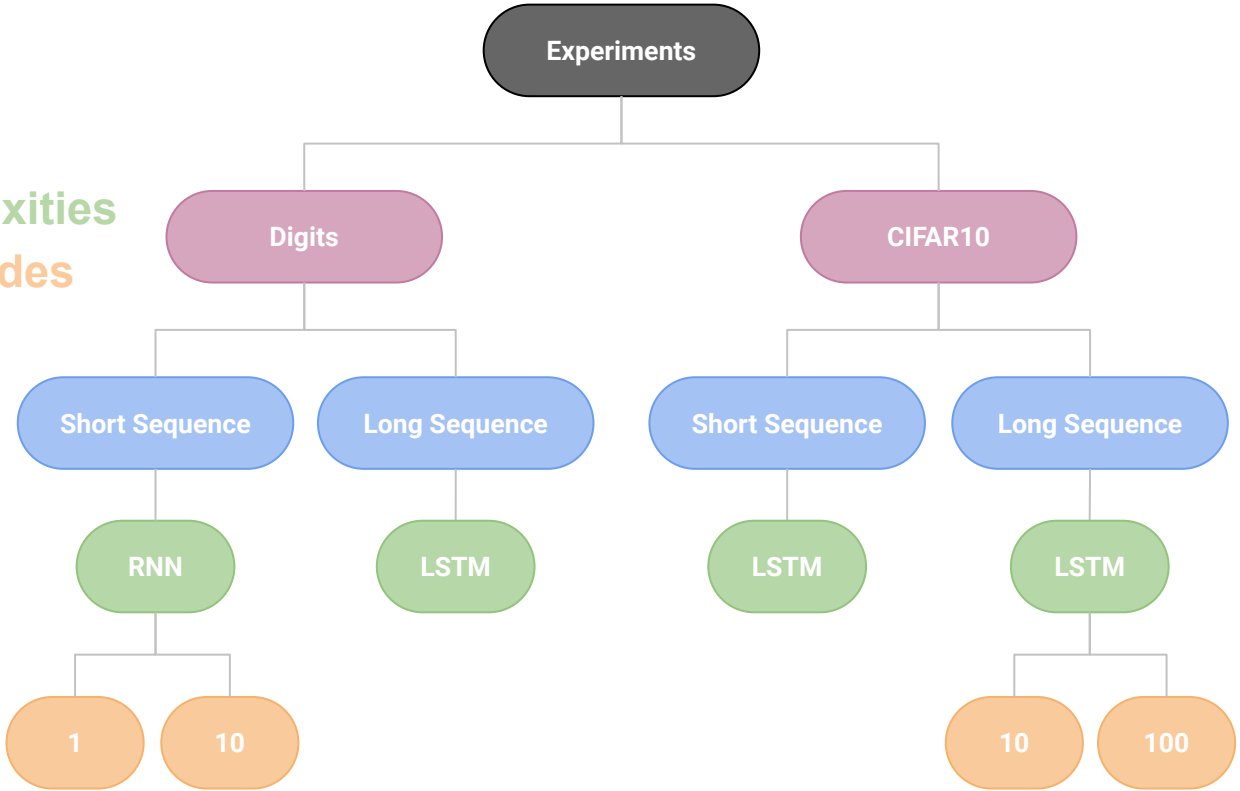
C: choice
R: reward

Q. Do artificial agents also obey the Matching Law?

Methods

- Two datasets
- Increasing difficulty
- RNN structure complexities
- Three reward magnitudes

Hidden size	64, 128
Learning rate	0.001, 0.0001
Discount factor	0.99
Batch size	32
Replay buffer size	100,000
Train steps	50,000, 200,000

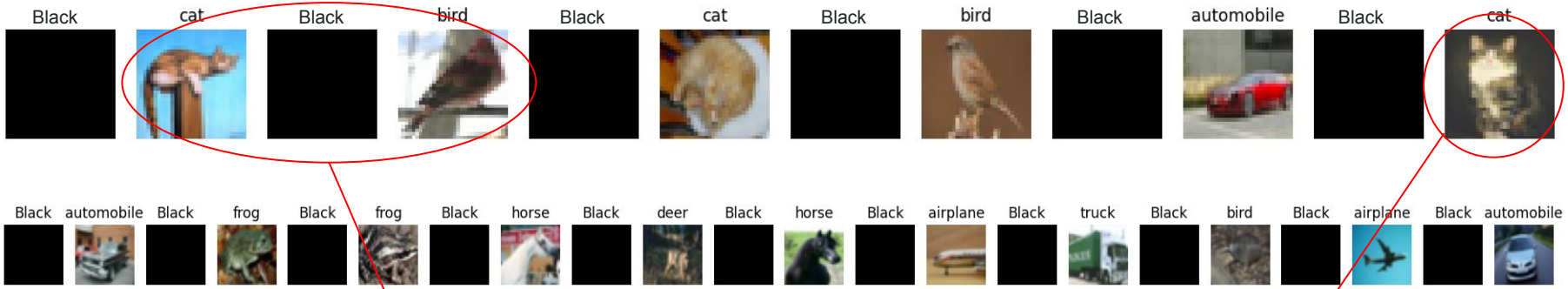


Tasks

0 - 1 - 0 - 4 - 0 - 2 - 0 - 5 - 0 - 3 - 0 - 1

0 - 1 - 0 - 5 - 0 - 3 - 0 - 6 - 0 - 3 - 0 - 4 - 0 - 9 - 0 - 2 - 0 - 8 - 0 - 7 - 0 - 1

(0, ■ : delayed period)



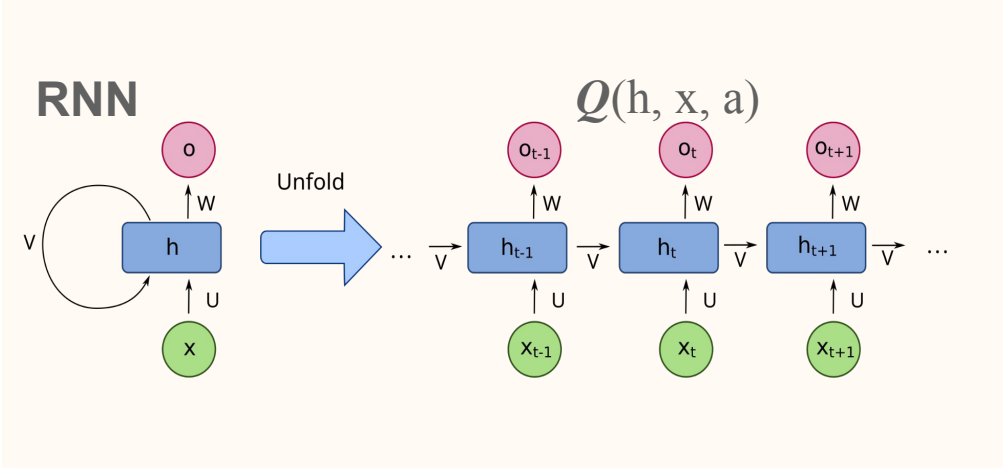
Possible actions: 0, 1, 2, 3, 4, 5, 6

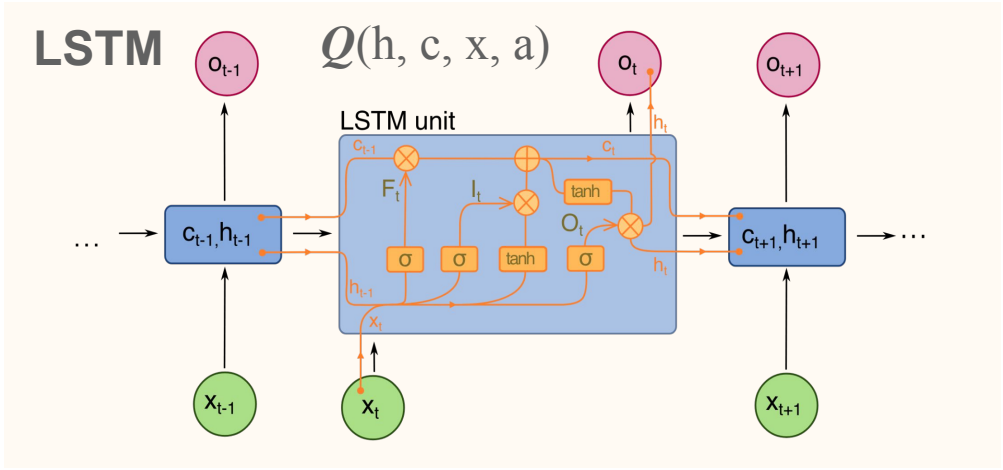
- Reward: 0
- Reward: -1

Possible actions: 0, 1, 2, 3, 4, 5, 6

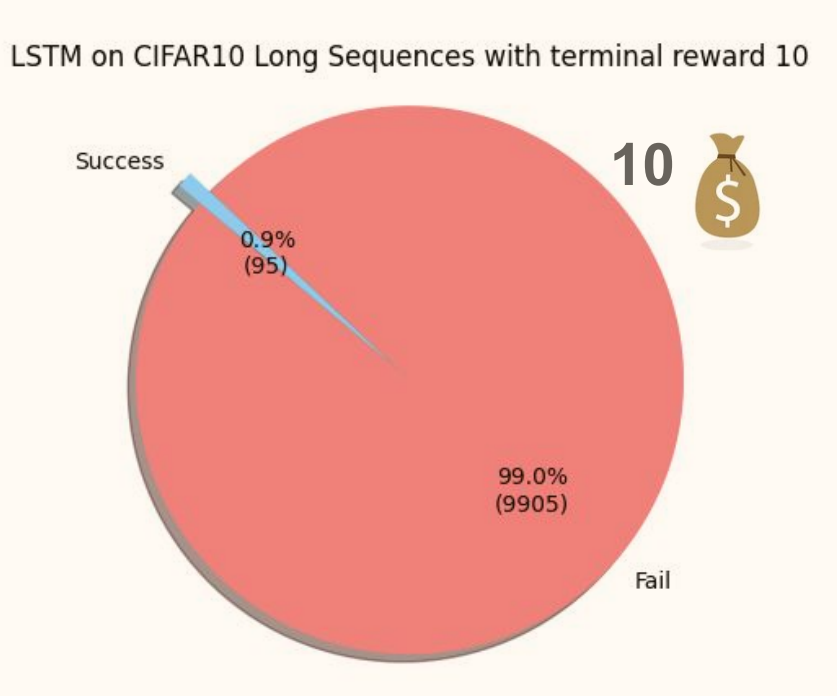
- Reward: 1, 10, 100
- Reward: -1

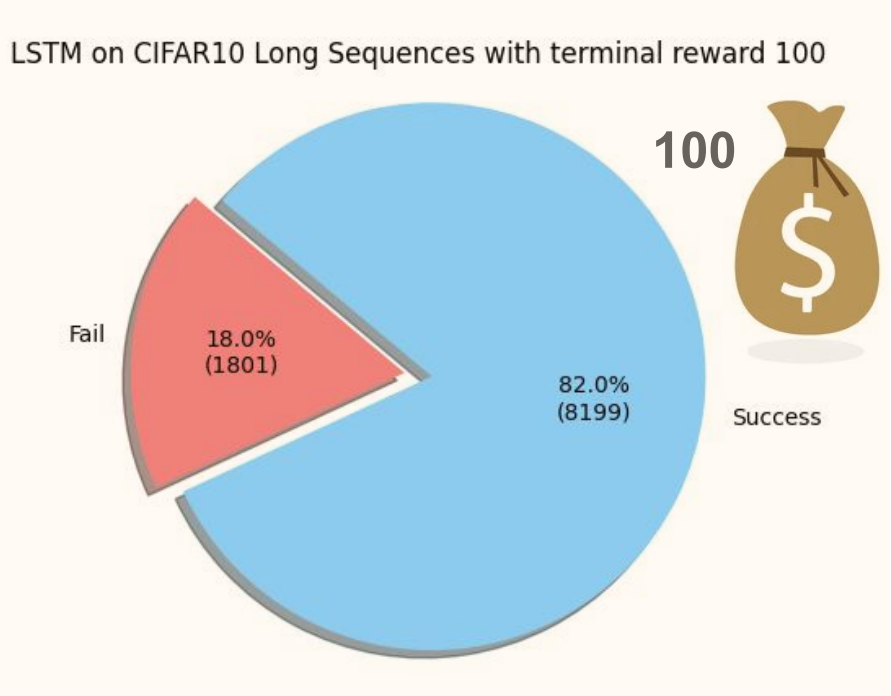
DQN


RNN
Q(h, x, a)
o
W
V
h
U
x
Unfold
o_{t-1}
o_t
o_{t+1}
h_{t-1}
h_t
h_{t+1}
x_{t-1}
x_t
x_{t+1}


LSTM
Q(h, c, x, a)
LSTM unit
c_{t-1}, h_{t-1}
c_{t+1}, h_{t+1}
F_t
I_t
O_t
tanh
σ

Matching Law Effect on Accuracy



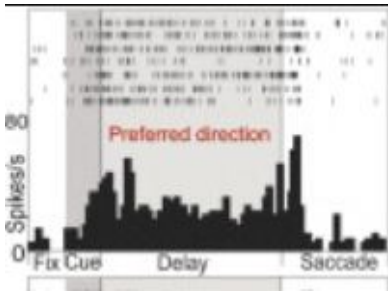


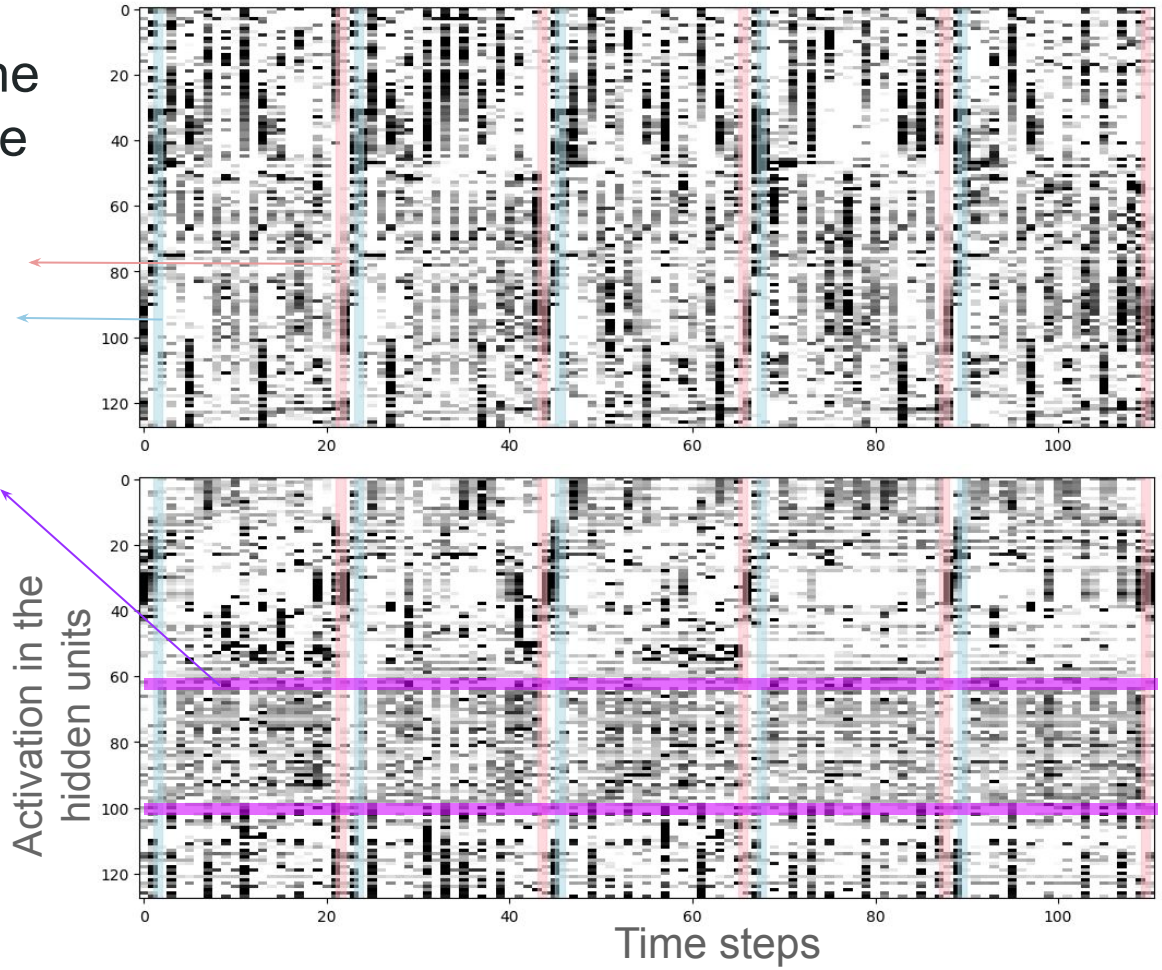
Raster map of the activations of the hidden units

Target stimulus

First stimulus

Working Memory State





10

100

Discussion

- **Similarities between natural & artificial intelligence:** Biological learning and artificial learning can show convergence in response to the same reinforcement principles

- **Experiments beyond the laboratory:** It is possible to scale sequence tasks in artificial agents beyond what laboratory animals are capable of

- **Future directions**
  - Tasks with an irregular sequence length
  - Other RL algorithms
  - Transfer learning to other tasks
  - Transformers instead of RNN’s

Thank you!