How rewarding can sorting cards be?

Depends on how flexible you are

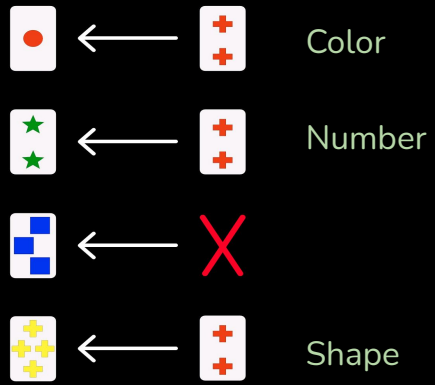
Foxy Reinforcers 🥵🦊

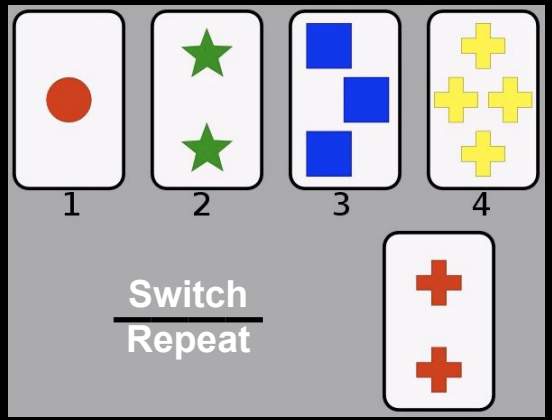
Introduction

Wisconsin Card Sorting Test (WCST)

Cognitive flexibility

Perseveration error + Set-loss errors





Question

Will a RL agent make errors similarly to healthy humans performing the WCST?

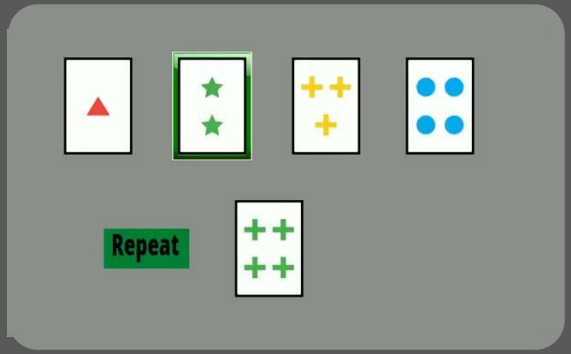
Hypothesis

Deep Q-Networks (DQN) will approximate the error distribution of human participants (Steinke et al. 2020) more closely than Naive Q-learning.

Modeling

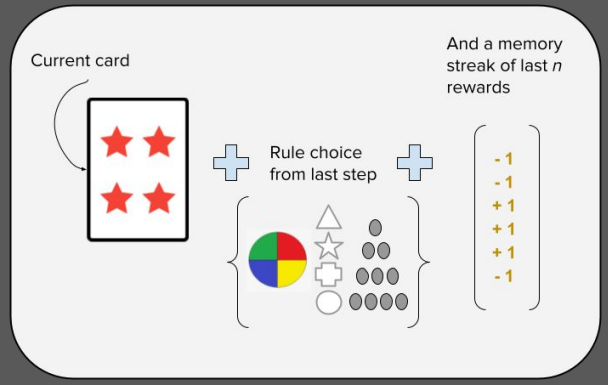
Gym Environment

- Discrete actions & observations
- Sorting rule picked at random
- Rule changes after 2-10 correct moves by the player (geometric distribution, mean 3.5)
- Reward of 1 (REPEAT) or -1 (SWITCH)



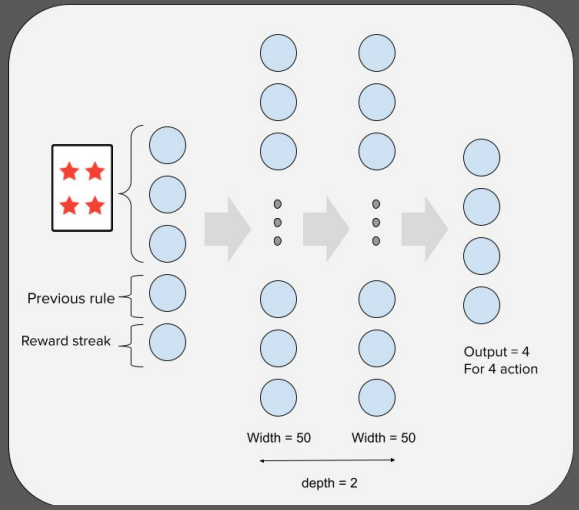
Vanilla Q-learning Agent

- Value based, off-policy, model free
- Behavior policy: ϵ-greedy
- Learning policy: TD Error min
- **Short-term memory**
- Q-table - (24 X 3 X 2n) X 4

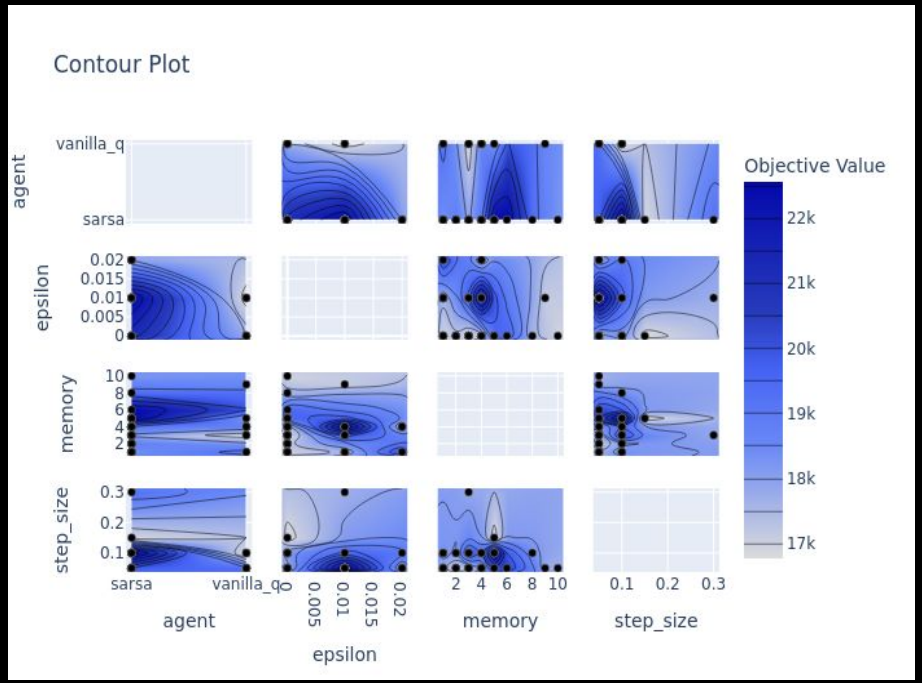


Deep Q Network

- **func(s -> a)** approximation.
- nn.Linear, nn.Relu
- 2 hidden layers
- batch size 10



Model fitting, Architecture, Fine tuning



NN architectures we tried

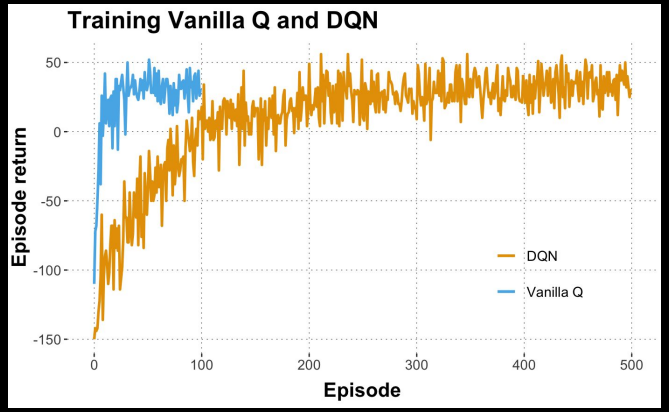
- 5 → 50 → 50 → 4
- 5 → (32*n)*k → 4

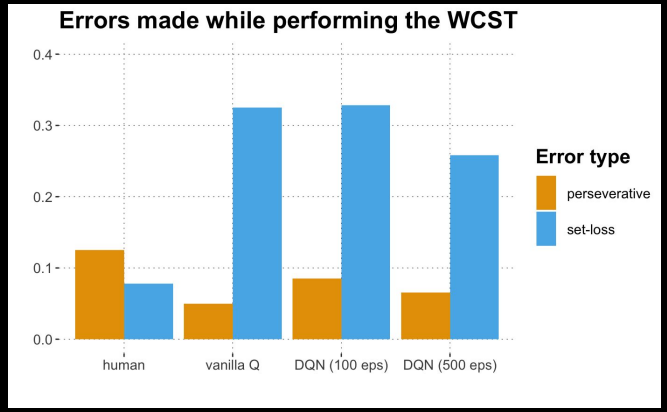
Best performance

- 5 → 64 → 96 → 32 → 4
- epsilon = 0.01, LR = 0.002

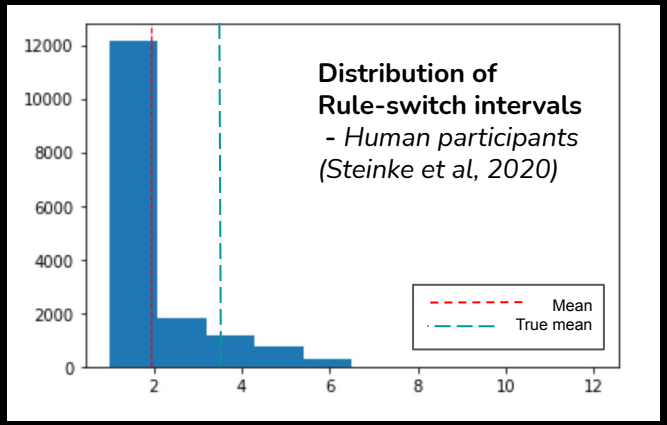
Results / Findings

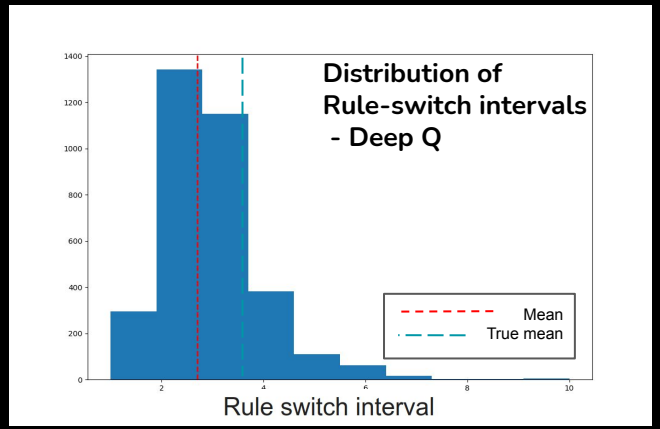
Comparison among models





Human vs. Model performance





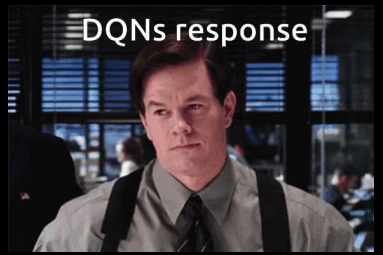
Discussion

What we can conclude

- Models switched their sorting rule more frequently than humans
- Models minimized perseverative errors at the cost of set-loss errors. After training for longer, DQN made fewer set-loss errors
- None of the models had a similar error distribution to humans

Limitations & future directions

- Policy-based methods - *Policy gradient*
- Models designed for short-term memory - *LSTM, GRU, LSTM-DQN*
- Model-based methods, direct modelling of rule switch probability - *HMM*
- Maybe it is a low dimensional problem - and we don't require a deep net at all?

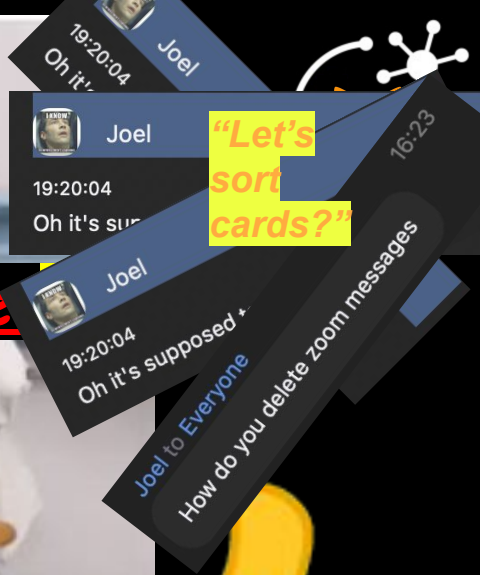


DO WE NEED AN EPILEPSY WARNING BEFORE THIS SLIDE??

Refefefefefererencecees

DO WE WANNA DO LUNCH? or any time

Check out our scripts at: https://tinyurl.com/ebonyfoxes



TOP 10 Q-LEARNING MODELS DEEPMIND DOESN'T WANT YOU TO KNOW ABOUT!!!

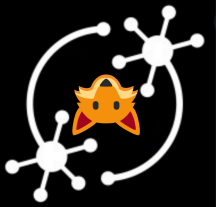
ϵRstepp$_u$

Thumbs Up Sign Emoji - Emojipedia." https://emojipedia.org/thumbs-up/. Accessed 16 Aug. 2021.

Kiitos ihan vitusti ধন্যবাদ..

“Fitting will fit great here feat Foxy Reinforcers 🥵🦊 *“ - Thank you!*

Mercy Bouquet!!



Thanks to Sanjukta, Hanifa.

Thanks to Radhimas, Jin and Michael.

And big thanks to *Neuromatch*!

Who we are (where, when, and why we are)