The Nature of Decision-Making: Human Behavior vs. Machine Learning

Keeya Beausoleil, Taher Jafferjee[1], Nathan Wispinski[2], Scott Stone[2], Craig Chapman[3]
[1]Department of Computing Science, [2]Department of Psychology [3]Faculty of Kinesiology, Sport, and Recreation

Supported By:
Dr. Chapman
CIFAR
NSERC CRSNG
UNIVERSITY OF ALBERTA FACULTY OF SCIENCE

Introduction

How do humans and artificial agents make decisions in different environments?

- Reinforcement Learning (RL) is a branch of machine learning optimizing rewards in different environments.
- We created a grid world to create foraging tasks to be used by humans and train artificial agents.
- We can compare the RL agent to that of humans.

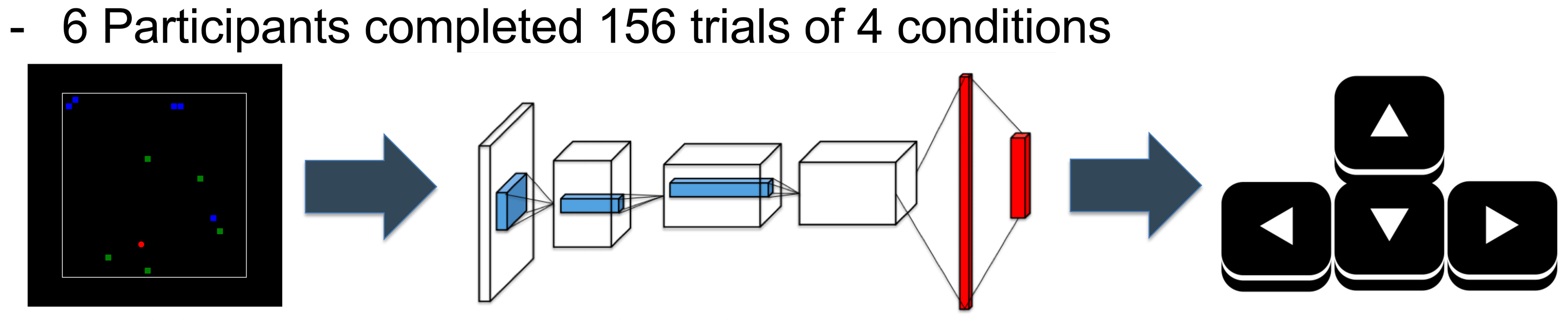
Motivation:

- Exploring this model's potential to predict human behavior in the future.
- Investigating differences that could improve the RL agent's performance.

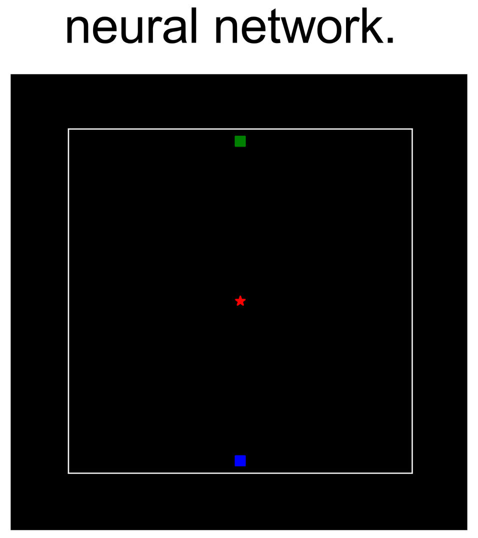
Procedure & Methods

1. Participants sit at a computer.
2. They press the arrows on the keyboard to navigate a grid world.
3. Each trial has a certain number of moves without a time limit.
4. Green = 5 points. Blue = 15 points.
5. Participants try to obtain the most reward before they run out of moves in different environments.

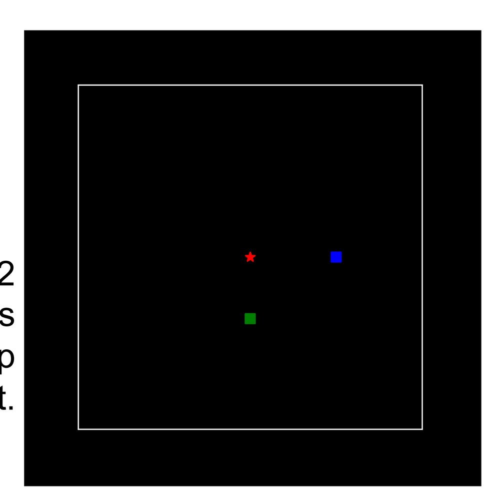
- 6 Participants completed 156 trials of 4 conditions



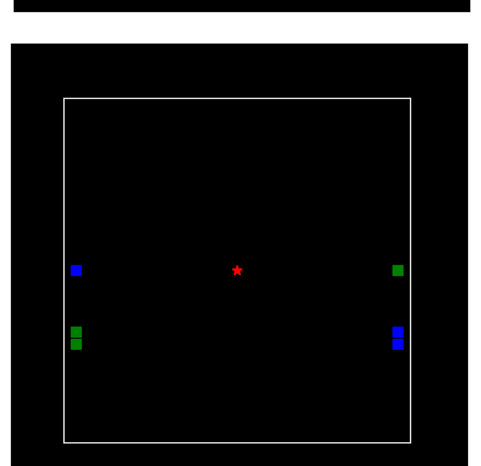
- The artificial agent was trained using a Reinforcement Learning algorithm (Deep Q Learning) through 50 million frames of random foraging task configurations. It is trained to maximize the number of points with an epsilon-greedy policy.
- The 27x27 pixel input, outputs into arrow key directions through a neural network.



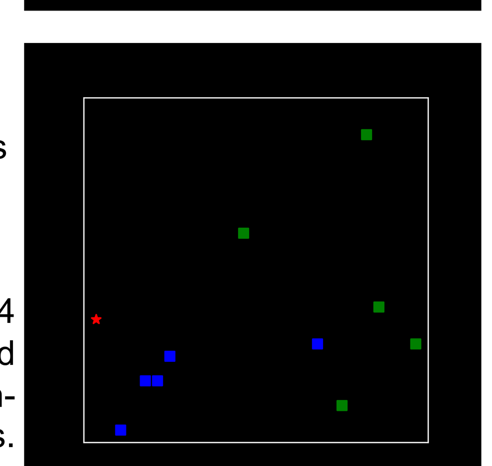
Condition 1
Limited moves meant only one reward obtainable. Tests preference.



Condition 2
"Unlimited" moves ensures both rewards are attainable. Studies relationship between reward and effort.



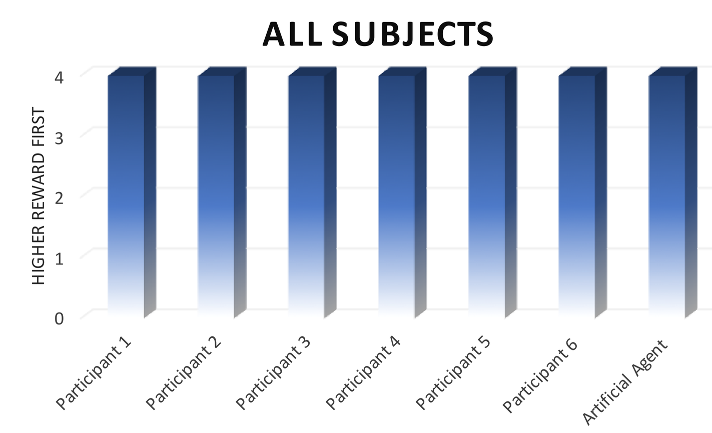
Condition 3
15 and 25 moves for one or three rewards to be achieved. Investigates long-term rewards and consideration of moves.



Condition 4
Limited number of moves. Designed randomly and rotated. Observes decision-making with environmental changes.

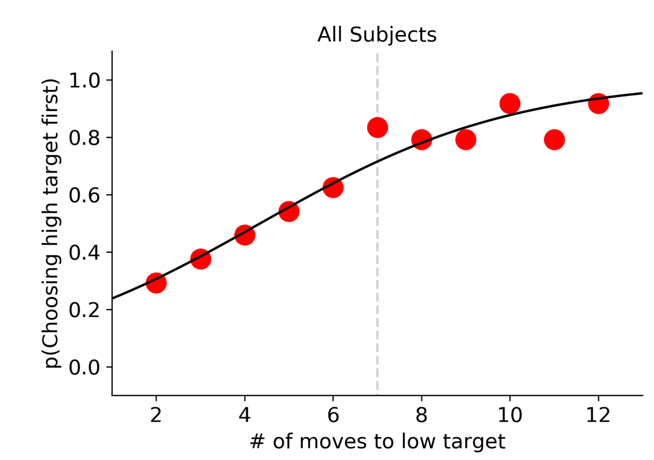
Results

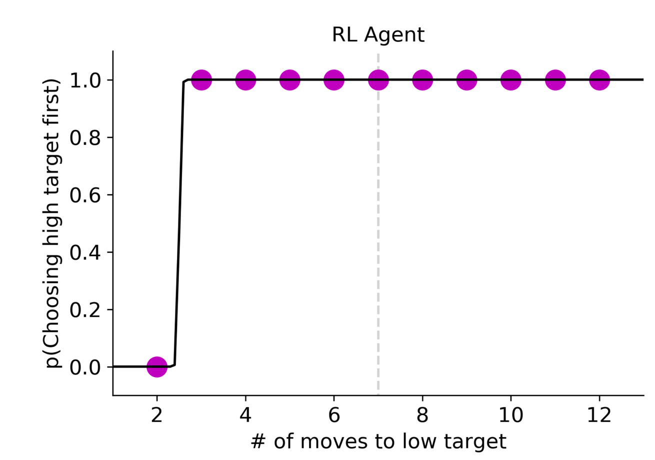
Condition #1: Which reward is preferred?

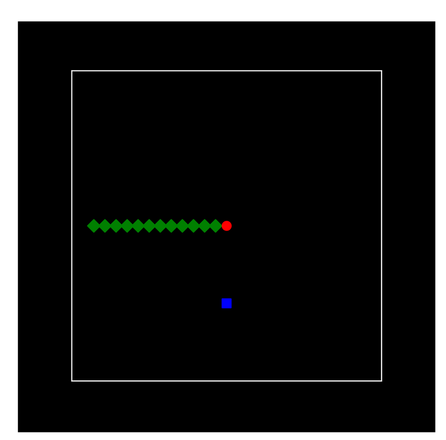


- Higher targets are clearly preferred by both human and artificial agents.
- The artificial agent behaves similarly to humans.

Condition #2: How much are humans and artificial agents willing to trade off rewards for movement?

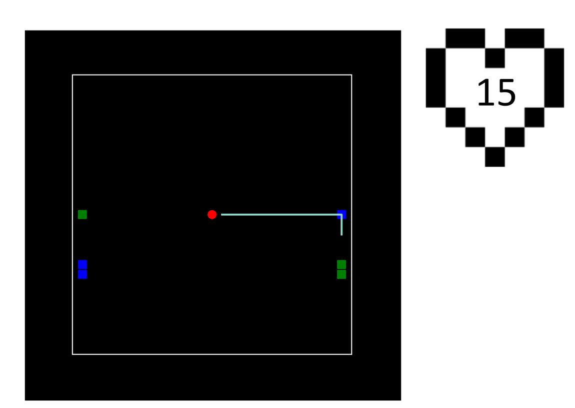


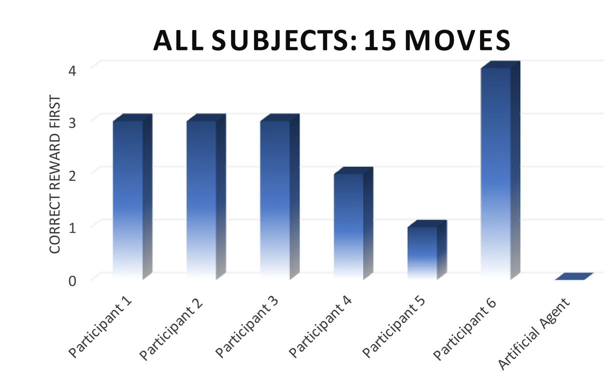


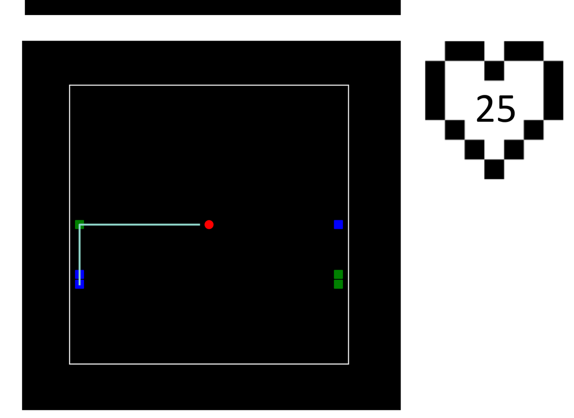


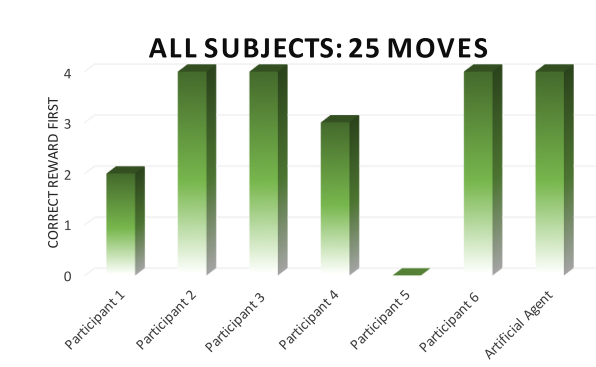
- Humans agents are indifferent when green target deviates approx. 4.5 moves.
- Artificial agent is indifferent when green target deviates approx. 3 moves.
- The slopes are different, showing human data has more noise.
- The RL agent cares about movement cost more.

Condition #3: Do humans and artificial agents plan multiple moves in advance? Do they consider the amount of moves?



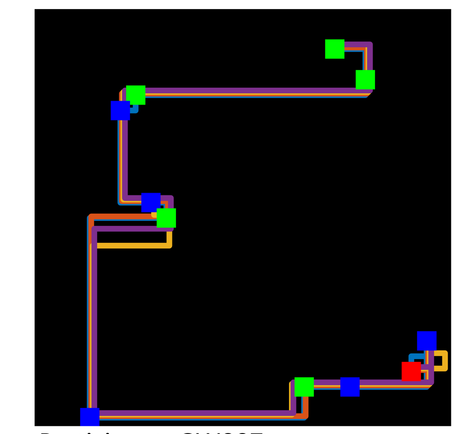




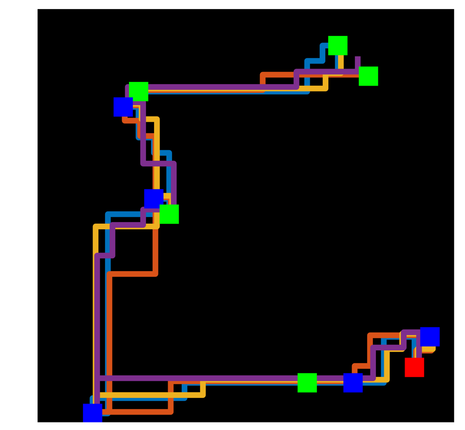


- Both humans and the RL agent plan multiple moves in advance.
- Humans are likely to consider moves when deciding their path, usually taking at least one trial to realize this. The RL agent disregards the difference in moves between 15 and 25.
- Humans illustrate the interplay between moves and long-term reward, something the RL agent is missing.

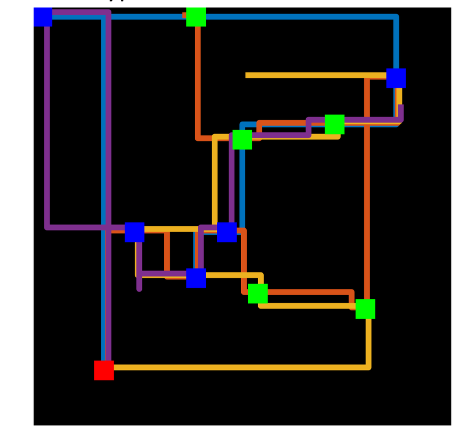
Condition #4: When environments are altered, do humans and agents change their paths?



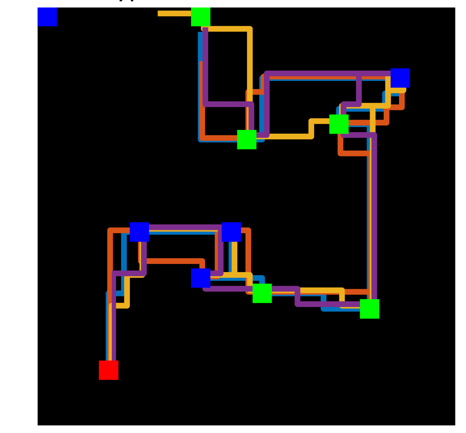
Participant: GW007
Trial Type: 6



Artificial Agent
Trial Type: 6



Participant: GW001
Trial Type: 22



Artificial Agent
Trial Type: 22

- Every agent completed the same arrangement 4 times, rotated.
- Unexpected that the RL agent does not appear deterministic.
- The RL agent is less variable than humans.
- Within one trial type, a human participant could have completely different paths, while others had minimal deviation.

Conclusions

Similarities:
Both humans and machine agents prefer higher rewards, trade off effort with reward, and plan multiple moves in advance.

Differences:
Humans and RL agents trade off movement and reward differently. Humans consider both future movements and number of moves, whereas the RL agent only favors long term reward. Both human's and the RL agent's path deviated differently with environmental changes. RL agent acted unexpectedly stochastic, yet did not resemble the human's randomness.

Future:
The similarities between human and artificial agents shows the potential RL agents have of being good models of predicting human behavior. The differences between the human and artificial agents suggests potential improvements of Reinforcement Learning methods, bridging the separation between artificial intelligence and human minds.

Acknowledgements

Thank you to Dr. Matina Kalcounis-Rueppell (Dean) and Dr. Ali Declan (Assistant Dean, Researcher) of the Faculty of Science for your support.
We would like to thank Hadeel Mohamed for her help with data collection!

References:
Wu et al. (2018). *Nature Human Behaviour.*
Goldstone, R. L., & Ashpole, B. C. (2004). *Psychonomic Bulletin & Review.*
Willke et al. (2019). *bioRxiv.*
Doshi-Velez, F., & Ghahramani, Z. (2011). *In Proceedings of the Annual Meeting of the Cognitive Science Society.*
Mnih et al. (2015). *Nature.*