An Exploration of Artificial Curiosity and Reinforcement Learning in a Simple Robot

Justine Ventura
Nadia Ady
Patrick Pilarski

Department of Computing Science
Department of Medicine and Dentistry

Curiosity: Desire to Learn

- How could we give machines their own curiosity?
- A possible solution could be to integrate a curiosity model with Reinforcement Learning (RL)
- RL is a branch of Artificial Intelligence in which the agent (the robot) learns from its experience, and is encouraged to perform the actions that will maximize cumulative reward

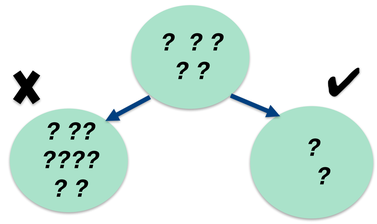
How does uncertainty change over time? How is reward affected?

How does the robot behave? Is there a change over time?

What can we learn from its behaviour?

Implementing Curiosity

- Curiosity drives us to experiences that make the world clearer
- A proposed model of this thinking is Information Gain Motivation (IGM)
- IGM rewards the decrease in uncertainty in the robot's knowledge of the world, after each action, as quantified by the decrease in *entropy*.
- Entropy is a mathematical concept which measures the uncertainty of an event.



Applying a Reinforcement Learning Method

- Most RL methods have the same model of the world: the agent in a state takes an action, and the environment gives a corresponding reward and takes it to a new state changes in predictions
- The agent changes its preferences for different actions based on estimates of states' values which it learns through accumulating reward

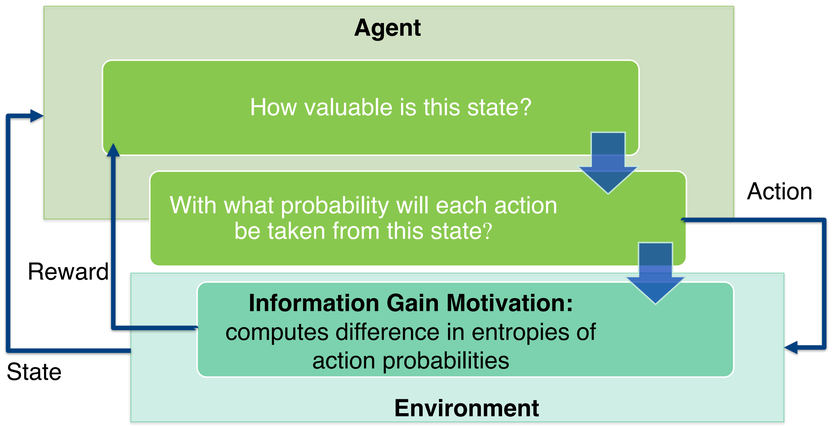
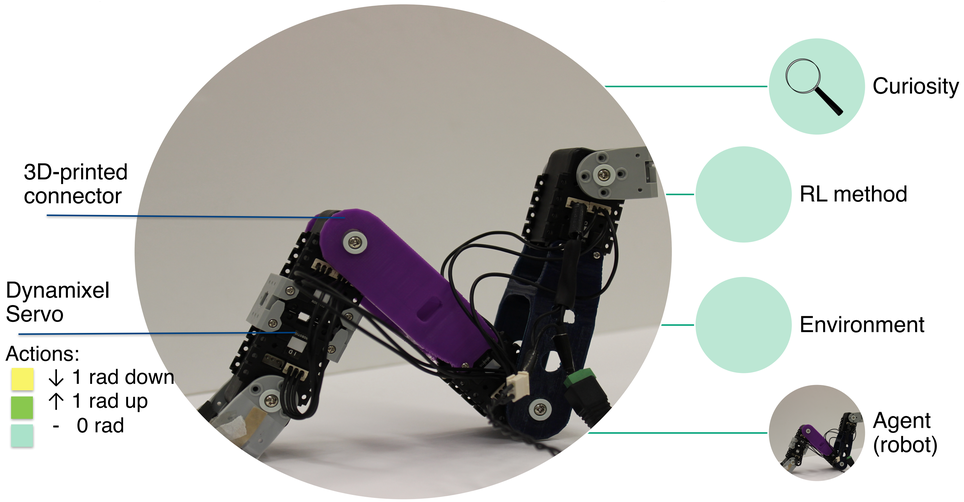


Figure 1: The agent learns to avoid actions which lead to less valuable states and to prefer actions which lead to more valuable states, gradually increasing its certainty.

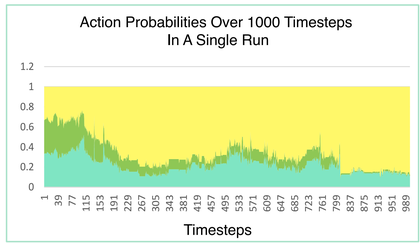
Robot in Action

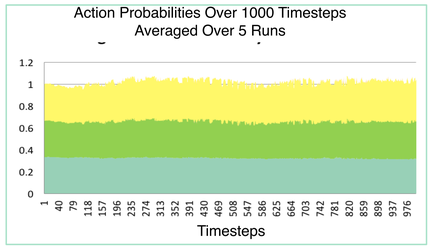
- The robot observes the positions of its motors, representative of its state
- At the start of each run, as the robot has no prior knowledge, it has equal preference towards all actions, so it behaves randomly



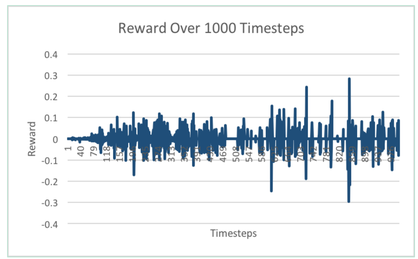
Observations

- In each individual run, the agent eventually developed a preference for taking a non-zero action, resulting in the robot taking the same action consecutively
- More specifically, it learned to remain stationary at the edge of its range of motion
- By observing its preferences, we found that at the edge of its range of motion, the robot increased its certainty more quickly than in the middle of its range
- The robot learned to prefer remaining still and learning nothing over futher exploration because the increase in uncertainty experienced by leaving edge states results in negative value





- The preferred non-zero action varied from run to run. Averaging over five runs, we found the probabilities of all actions at each timestep to be equal



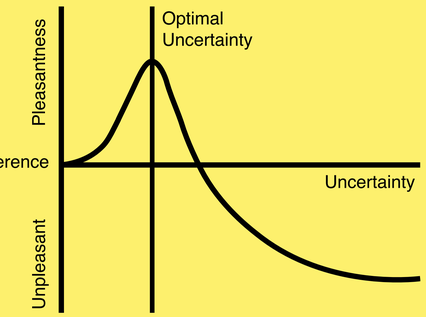
- Since the robot visited some states more frequently than others, the difference in state entropies became larger over time, and thus reward increased in magnitude
- Where entropy did not change, reward was zero

Evaluating Curiosity

- The robot's tendencies to behave in a non-exploratory manner and remain in the same position is non-ideal for curiosity-driven behaviour.
- By simply rewarding decrease in uncertainty, it implies that no exploration is better than receiving negative reward.

Moving Forward

- To improve upon IGM, it could be modified to reward the absolute difference in entropies rather than the true difference.
- This would encourage the robot to learn about and explore its environment.
- The prevailing theory of curiosity presents the idea that a positive amount of uncertainty is optimal
- However, the aim of IGM is to reduce uncertainty, which is not fully compatible with this theory



Fig, 2: Wundt's curve

- With RL being more prevalent than ever, artificial curiosity also holds great potential to be applied to technologies.
- For example, in the field of medicine, curiosity could be applied to a prosthetic arm such that it can perform a task in a manner not previously known

Acknowledgements

The author would like to thank:
- Nadia Ady, Supervisor
- Dr. Patrick Pilarski, Principal Investigator
- The BLINC Lab
- Process Solutions
- WISEST

Literature Cited

Sutton, Richard S., and Andrew G. Barto. Reinforcement Learning: An Introduction. Cambridge, MA: MIT, 2012. Print.

Oudeyer, P. Y., & Kaplan, F. (2007). What is motivation? A typology of computational approaches. Frontiers in neurorobotics, 1.

Magnifying Glass Clipart.

O'Neil, Harold F., and Michael Dillings. "Curiosity and Exploratory Behaviour." Motivation: Theory and Research, Erlbaum, 1994.