An Exploration of Machine Curiosity and Reinforcement Learning Using a Simple Robot

Justine Ventura
Nadia M. Ady
Patrick M. Pilarski

Department of Computing Science
Department of Medicine and Dentistry

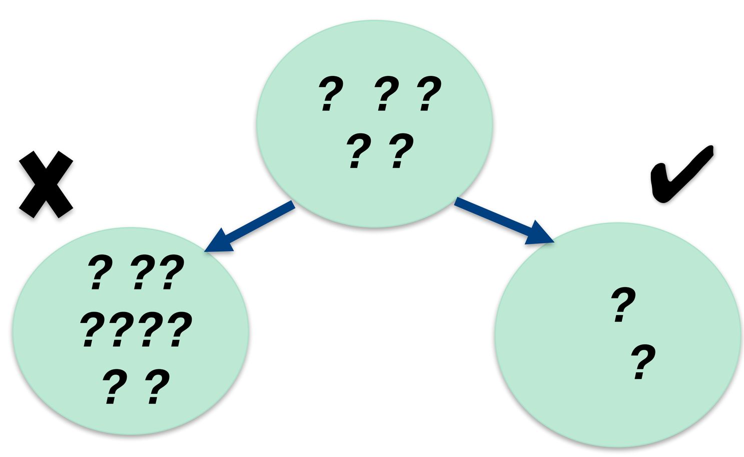
Curiosity: Desire to Learn

- Curiosity is a prevailing aspect of human life, found in stages as early as infancy
- How could we give machines their own curiosity?

- How might a curious robot's knowledge change over time?
- How does the robot behave? Is there a change over time?
- What can we learn from the robot's 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

- Reinforcement learning (RL) is a branch of Artificial Intelligence in which an agent (in this case, a robot) learns from its experience, and is encouraged to perform the actions that will maximize cumulative reward
- 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
- The agent changes its preferences for different actions based on estimates of states' values which it learns through accumulating reward

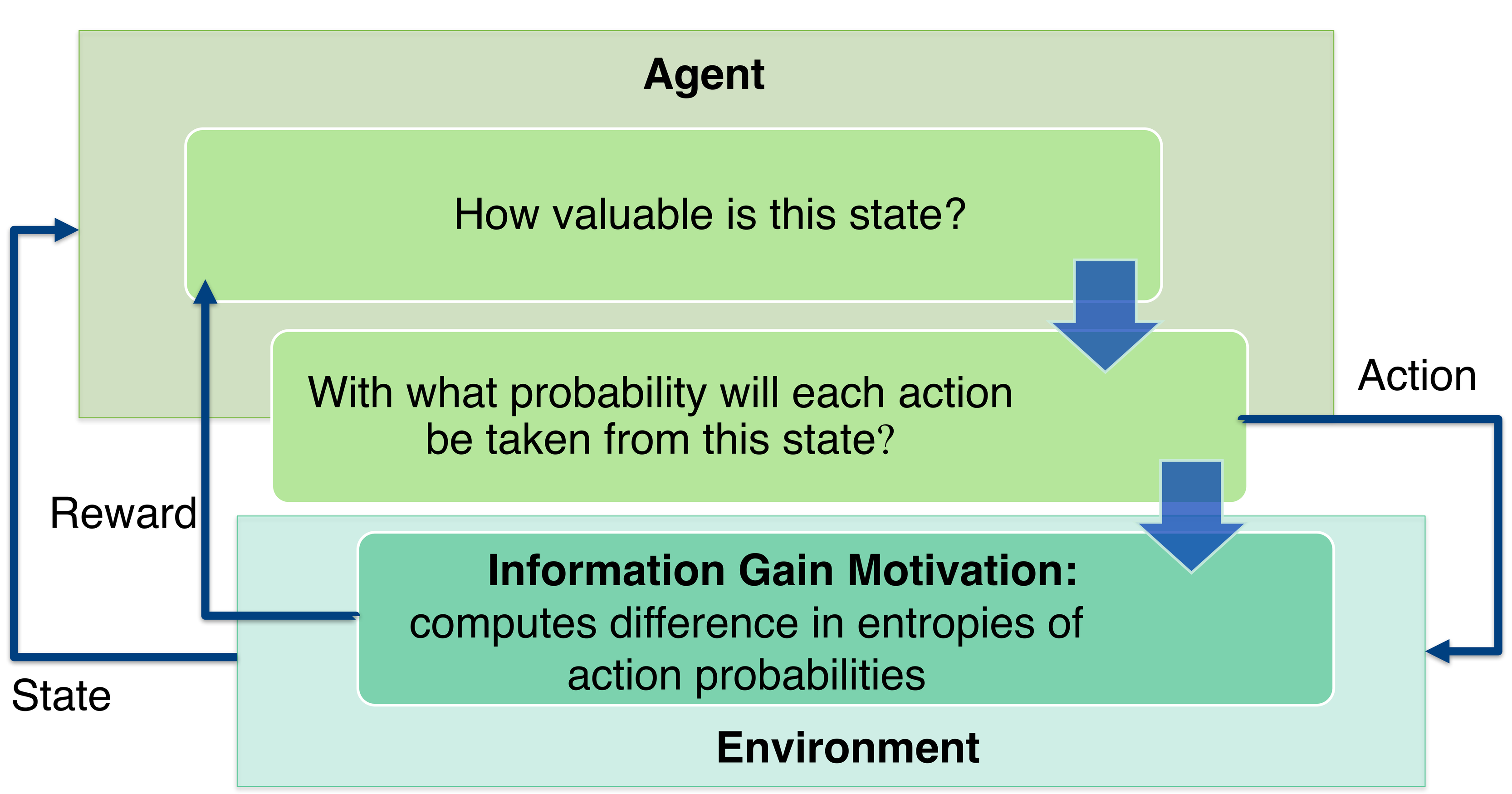
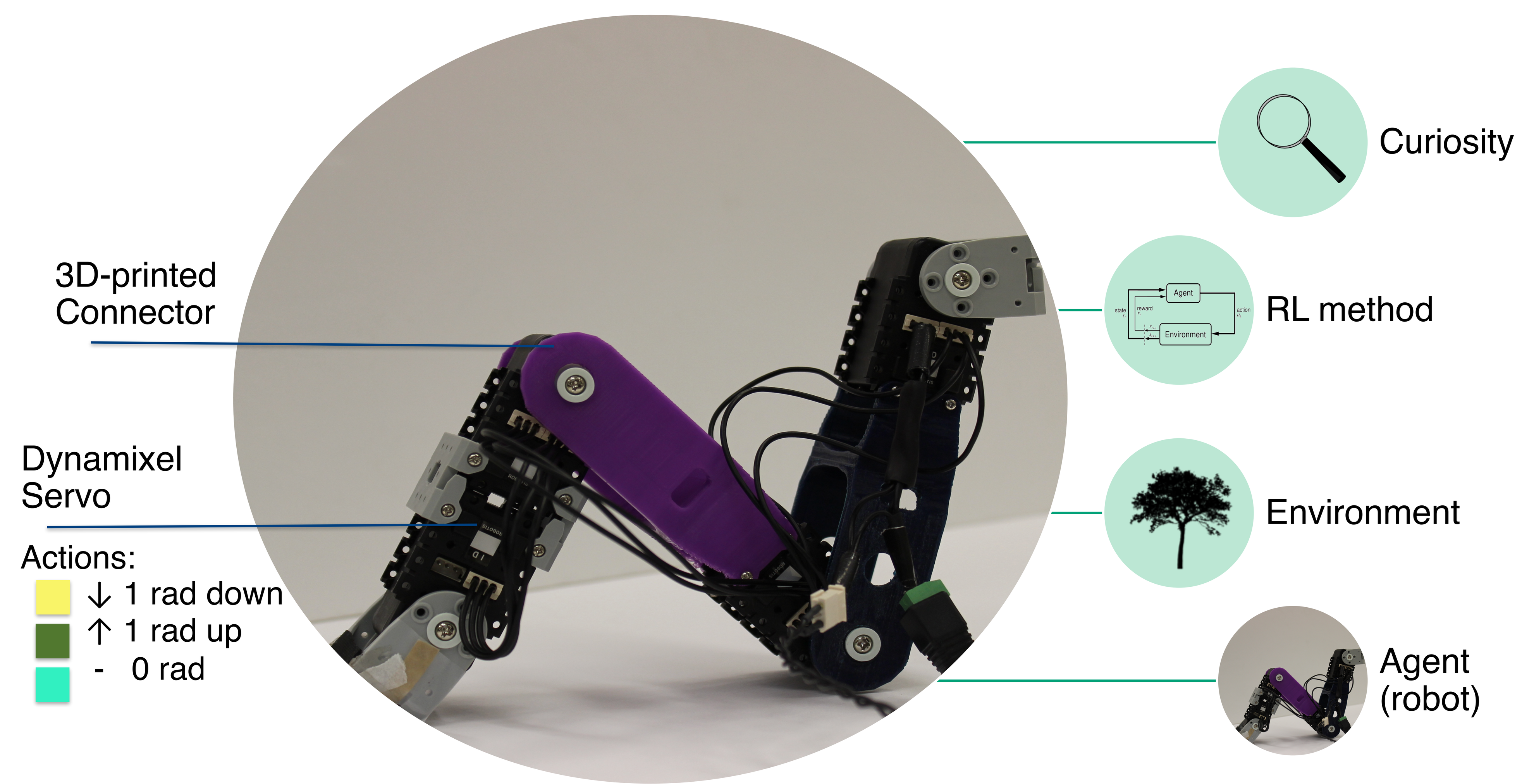


Figure 1: The agent learns to avoid actions which lead to less valuable states and to prefer actions which lead to more valuable states, gradually increasing its certainty.

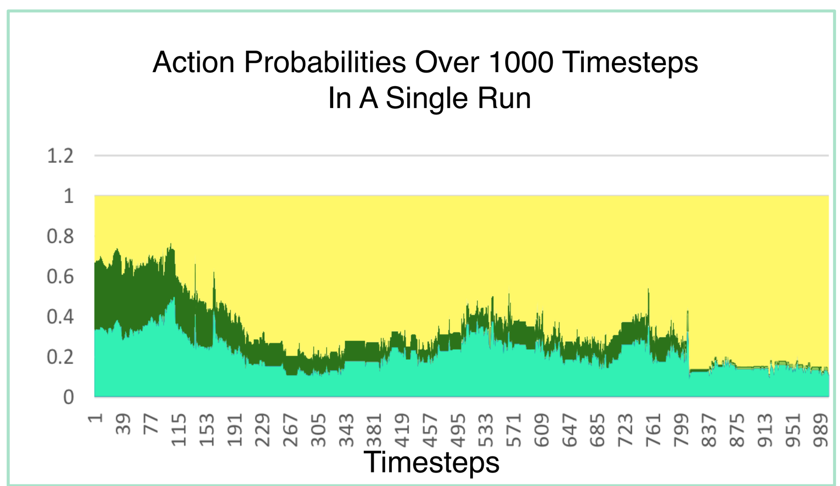
Robot in Action

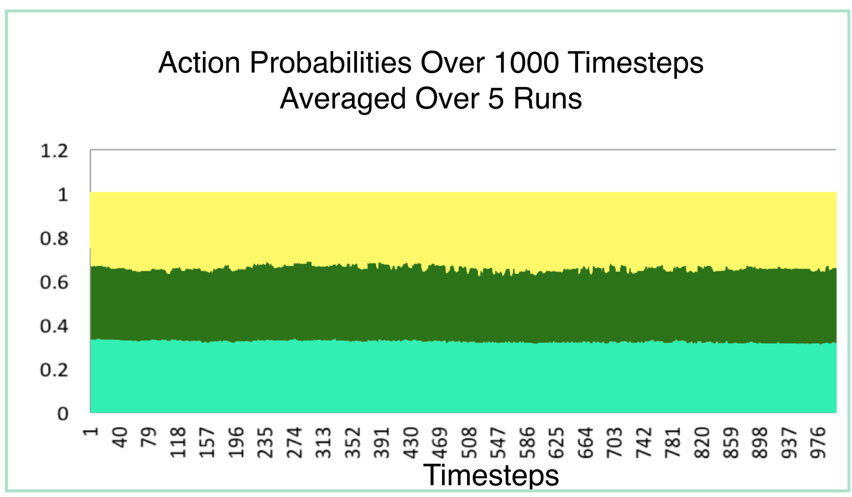
- The robot observes the positions of its motors, used to represent its state
- At the start of learning, as the robot has no prior knowledge, it has equal preference towards all actions, so it behaves randomly



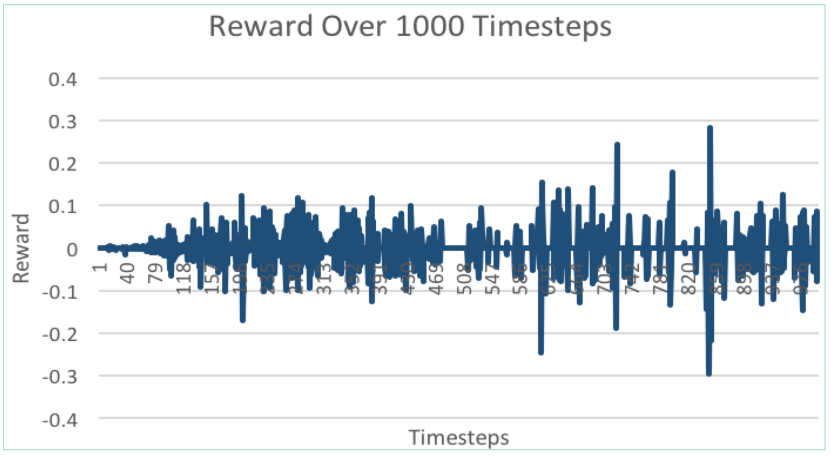
Observations

- 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 further 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 entropies between states became larger over time, and thus reward increased in magnitude
- Where entropy did not change, reward was zero

Evaluating IGM

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

- 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

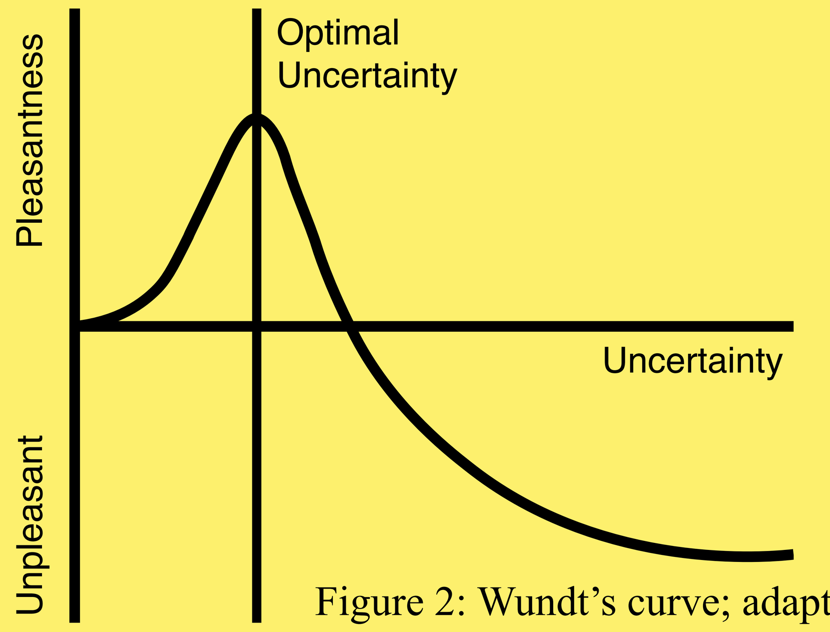


Figure 2: Wundt's curve; adapted from O'Neil and Dillings

Moving Forward

- An improvement on IGM could be to reward the absolute difference in entropies rather than the signed difference
- This modification could encourage the robot to learn about and explore its environment

- With RL being more prevalent than ever, machine curiosity also holds great potential to be applied to technologies.
- For example, in medicine, curiosity could be applied to a prosthetic arm so it could perform a task not thought of by the arm's designers

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, 2016.
Oudeyer, P. Y., & Kaplan, F. (2007). What is motivation? A typology of computational approaches. Frontiers in neurorobotics, 1.
O'Neil, Harold F., and Michael Dillings. "Curiosity and Exploratory Behaviour." Motivation: Theory and Research, Erlbaum, 1994.