

Learning Semantically Meaningful World Representations Through Embodiment

How do humans acquire a meaningful understanding of the world with little to no supervision and no semantic labels provided by the environment? I argue that our understanding and representation of the world is largely based on our interaction with it. It is commonly accepted that action requires perception. I postulate that this relation also reverses; Perception requires action. I believe that many of the shortcomings of classical neural networks, such as vulnerability to adversarial attacks, originate from the unnatural way in which they are trained and the lack of embodiment.

To investigate this, I train artificial agents in a virtual environment using deep reinforcement learning. After interaction with the environment and receiving very sparse rewards and no semantic labels I show how semantically meaningful and robust representations of the world have been learned. Furthermore, I look at how these representations compare to biological findings and if an embodied learning framework could be a more realistic way to model the brain.