Reinforcement Learning Models of Tradeoffs Between Infant Attachment and Curiosity

Xi Jia Zhou*, Frieda Rong*, Michael Frank, Nick Haber
Graduate School of Education, Department of Psychology, Department of Computer Science

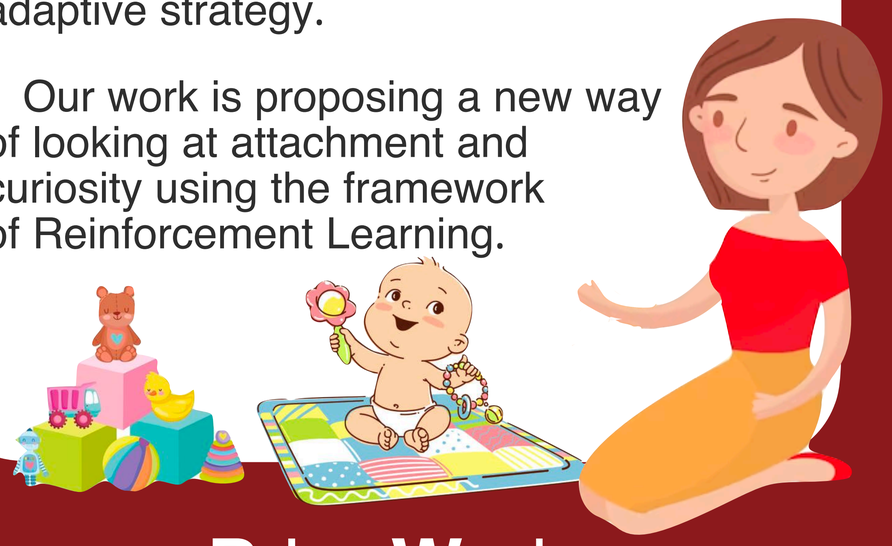
Problem Statement

Infants must balance navigating the environment and interacting with the caregiver, driven by their curiosity and attachment.

Both curiosity and attachment can be thought of as involving reward seeking from infants' immediate surroundings — curiosity as novelty- or information-seeking, and attachment as seeking proximity and comfort from a caregiver. Therefore, the decision to choose between both options involves trading off rewards from exploiting the caregiver and exploring the environment.

Can we simulate infants' various attachment styles that show as different proximity seeking and exploration patterns? Our hypothesis is that different combinations of frequencies and scales of rewards in the caregiver and the environment, including negative rewards due to distressing encounters derive different attachment behaviors—an adaptive strategy.

Our work is proposing a new way of looking at attachment and curiosity using the framework of Reinforcement Learning.



Prior Work

Attachment styles are commonly assessed using the Strange Situation Test based on observations of their play and parent-seeking behaviors.

A stranger is introduced

The child is reunited with parent after separation

Infants' were assessed across episodes in SST, a procedure of introducing infants novelty, such as a stranger and separations from the caregiver.

Computational Model & Parameters

We modeled infant's tradeoff between parent and environment as a two-armed bandit problem. We implemented the Upper Confidence Bound algorithm for infant's decision making process.

Reward from Parent

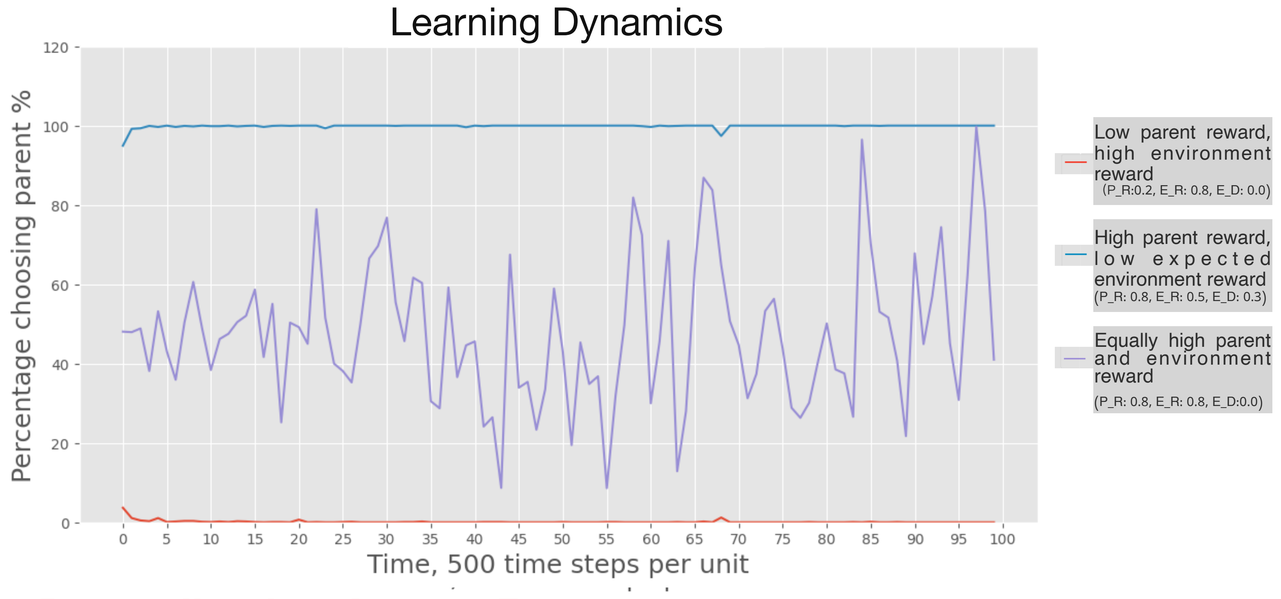
Reward/ Distress from Environment

Experimental Parameters:
Parent: Probability of Reward.
Environment: Probability of Reward & Probability of Distress.

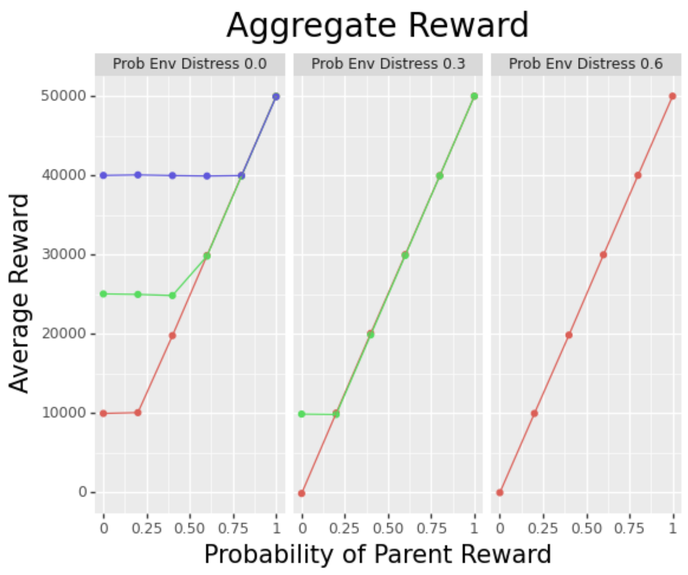
(Probability of ignorance = *1-reward* for parent, and similarly, Probability of no-response = *1-reward -distress* for environment).

Simulating Different Attachment Styles

Simulation Outcomes



Generalization Across Parameters

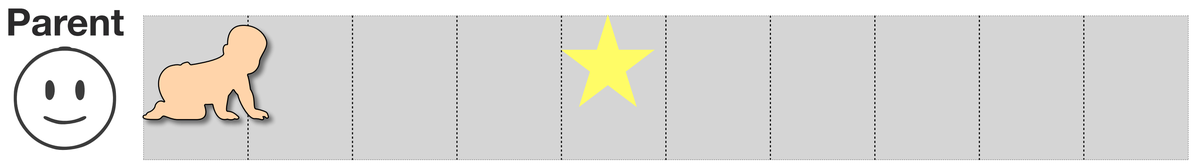


Total run time = 50000
Iterations = 3

As expected, in our bandit set-up, infant chooses either parent or environment following maximum expected rewards.

Current Work & Next Steps

Parent



We generated different behaviors of avoiding and approaching the parent from our simulation and an infant favoring maximum expected reward. For next steps, we are increasing the complexity of the environment and the infant agent. We are implementing the child as an agent crawling in a 1-d grid world with caregiver on one side and new grids for exploration on the other, from which we can measure distance between infant and the caregiver. We are also adding states to our infant agent in this new environment.