INTRINSIC MOTIVATION FOR ARTIFICIAL AGENTS

PAUL ROBERTSON AND OLIVIER L. GEORGE

INTRODUCTION

- Intrinsic motivation, in the study of artificial intelligence and robotics, is a mechanism for enabling artificial agents (including robots) to exhibit inherently rewarding behaviors such as exploration and curiosity, grouped under the same term in the study of psychology.
    - Imitation of human's intrinsic motivation

- This paper reviews the state-of-the-art in intrinsic motivation research, focusing on its application in the absence of explicit goals.

SELF-DETERMINATION THEORY (SDT)

- SDT is a psychology theory about human's motivation and personality.
- SDT enumerates a number of sources of inherent satisfaction in humans as follows.
  - Autonomy, Competence, Relatedness, Curiosity and Discovery, Flow State, Creativity and Self-expression, Purpose and Meaning, Mastery of Challenges, Playfulness and Joy, Contribution and Impact, Alignment with Personal Values, Physical and Sensory Enjoyment.
- No single measure of intrinsic motivation is sufficient; rather, we aim for all measures to be present and for there to be some mechanism that decides the strength of each motivational dimension or inherent satisfaction.

EXTRINSIC MOTIVATIONS

- Although intrinsic motivations are our primary focus, it is crucial to acknowledge the influence of extrinsic factors.
  - Extrinsic motivations play a pivotal role in fostering societal cohesion and function.

- Several crucial extrinsic motivations:
  - Tangible Pursuit of External Positive Outcomes, Rational Negative Outcome Avoidance, Fear, Escaping from Negative Situations such as Discomfort or Boredom, Routine and Complying with Societal Expectations.

APPROACHES TO INTRINSIC MOTIVATION

- Curiosity-Driven Exploration
  - Pros: This approach promotes open-ended learning and discovery and is effective in sparse reward environments.
  - Cons: This approach often leads to the exploration of irrelevant or unproductive areas.

$$r_{intrinsic} = \|\widehat{y_t} - y_t\|$$

- $\widehat{y_t}$ is the agent's predicted outcome at time t.
- y_t is the result observed at time t.
- The more the predictions differ from the observations, the more interesting the situation is.

APPROACHES TO INTRINSIC MOTIVATION

- Information Gain
  - Pros: This approach promotes the development of valuable environmental models and is in good alignment with long-term learning objectives.
  - Cons: It can be computationally intensive because of the requirement for predictive modeling. In addition, it may exhibit a tendency to prioritize uncertainty over task-specific goals.

$$r_{intrinsic} = I(A;O) = H(O) - H(O|A)$$

  - $I(A;O)$ is the mutual information between the agent's actions (A) and observations (O).
  - $H(O)$ is the entropy of the observations.
  - $H(O|A)$ is the conditional entropy of the observations given the actions.

APPROACHES TO INTRINSIC MOTIVATION

- Pattern Discovery and Intellectual Curiosity
  - An intriguing variation on information gain, to focus on uncovering patterns or structures within the data, thus fostering intellectual curiosity rather than just seeking information gain.

$$r_{intrinsic} = H(P_{before}) - H(P_{after})$$

- $H(P_{before})$: Entropy of the model before discovering the pattern.
- $H(P_{after})$: Entropy of the model after integrating the discovered pattern.

APPROACHES TO INTRINSIC MOTIVATION

- Novelty Search
  - In Novelty Search, agents are rewarded for exploring new or previously unvisited states, rather than achieving predefined goals.
  - Pros: The novelty search promotes the exploration of new areas of the environment, preventing the agent from being trapped in the local optima. Since the reward is intrinsic, it is effective in environments with little or no external feedback.
  - Cons: Agents may prioritize novelty over achieving specific goals, which can reduce task efficiency. Maintaining and comparing against the set of previously visited states can be computationally expensive in large or continuous environments. Its implementation is complicated by the difficulty of defining novelty.

$$r_{intrinsic} = \min_{s_i \in S} \|s - s_i\|$$

  - s is the current state; S is the set of previously visited states.

APPROACHES TO INTRINSIC MOTIVATION

- Intrinsic Rewards via World Models
  - It uses learned models of the environment to generate intrinsic rewards based on prediction errors or discrepancies between expected and observed outcomes. This approach helps the agent develop a deeper understanding of the environment.
  - It is effective in handling complex dynamic settings but requires robust model learning and can be prone to instability if the model is inaccurate.

$$r_{intrinsic} = \max(0, \delta_t - \delta_{t-1})$$

  - δ_t is the prediction error at time t.
  - δ_{t-1} is the prediction error at the previous time step.

APPROACHES TO INTRINSIC MOTIVATION

- Skill Acquisition
  - Encourage the agent to learn various skills by maximizing the competence or utility of new abilities.
  - Pros: This approach promotes the development of transferable skills and promotes diversity in behaviors and solutions.
  - Cons: Skill acquisition over achieving specific tasks. Definition of valuable skill.

$$r_{intrinsic} = C_t - C_{t-1}$$

  - C_t: Competence at time t, defined as the probability of successfully completing a task.
  - C_{t-1}: Competence at the previous time step.

APPROACHES TO INTRINSIC MOTIVATION

- Constrained Intrinsic Motivation
  - Balance intrinsic and extrinsic motivation by constraining intrinsic rewards to ensure alignment with task goals.
  - This reduces the risk of irrelevant exploration and ensures that focus is on achieving task-specific objectives. The approach requires a careful design of constraints.

$$r_{total} = r_{intrinsic} - \lambda C$$

  - r_{total}: The total reward used for learning.
  - $r_{intrinsic}$: The intrinsic reward encourages exploration or curiosity.
  - C: A penalty term representing violations of constraints.
  - λ: A weighting factor that determines the influence of the penalty term.

APPROACHES TO INTRINSIC MOTIVATION

- Self-imitation Learning
  - The approach encourages sustained attention to novel skills until they reach a level of proficiency.
  - However, it may inadvertently reinforce suboptimal solutions and limit adaptability to evolving environments.

$$r_{intrinsic} = \max\big(0, Q(s_t, a_t) - Q_{mean}(s_t)\big)$$

  - $Q(s_t, a_t)$: The estimated value of taking action a_t in state s_t under the agent's current policy.
  - $Q_{mean}(s_t)$: The mean value of all actions in state s_t, representing the baseline performance.
  - This reward encourages the agent to prioritize actions that outperform the average.

APPROACHES TO INTRINSIC MOTIVATION

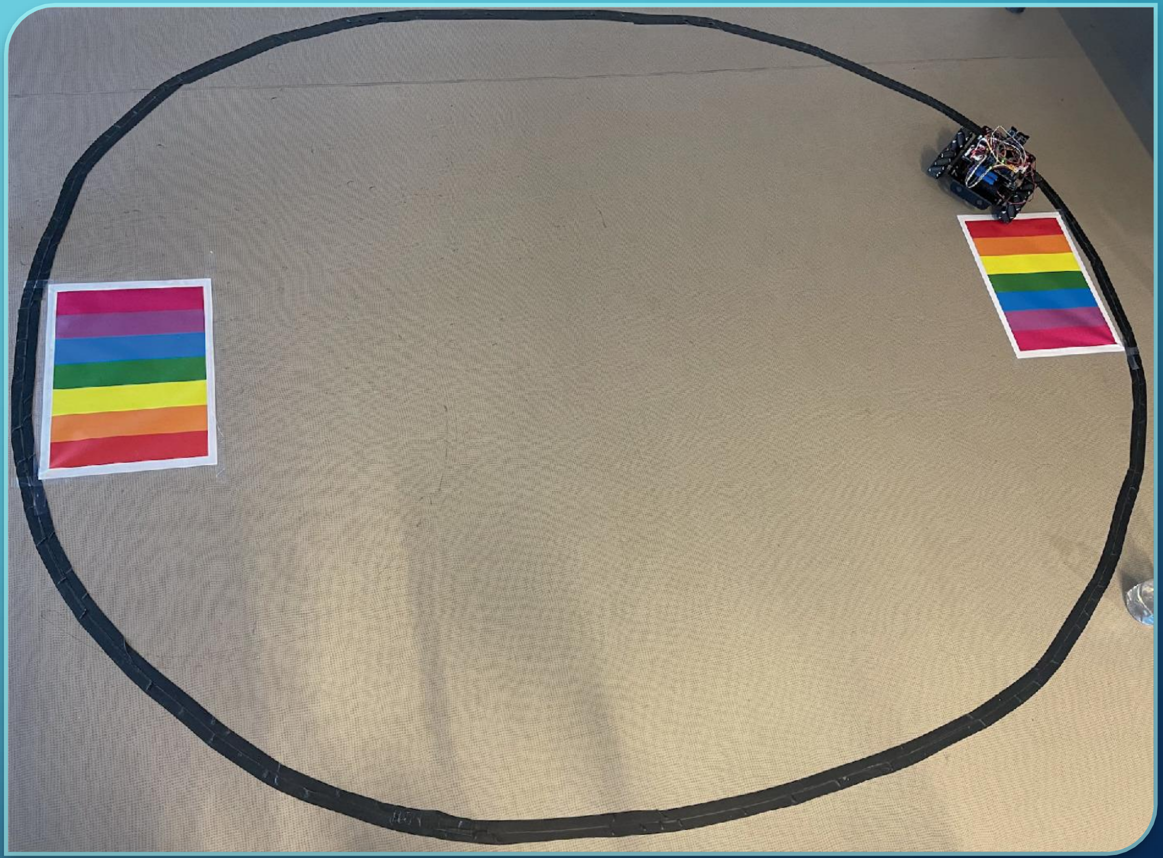
- Flocking Motivation
  - Common in animals and perhaps in all animals is the notion of flocking.
  - The Boids algorithm is developed by Craig Reynolds in 1986 which simulates the flocking behavior of birds. There are three primary forces: Separation, Alignment, and Cohesion.

INDEPENDENT ROBOTICS

- In classical robotics, the robot is given a goal, and it performs that goal.
- In developmental learning, we want our robots to learn through exploration rather than by being given explicit goals. When robots are free to choose what they want to do, they are independent of human guidance.
- In the short term, we want our robots to be endowed with intrinsic and natural extrinsic motivations to cause the robots to exhibit interesting behaviors that lead to rapid learning of new skills.

EXPERIMENT

- An experiment using physical robots on physical test beds.
  - The testbeds are simple and synthetic, but they are steppingstones to genuinely real-world scenarios.



EXPERIMENT - SCENARIO

- The robots begin their life in the simple terrain shown in figure, they begin with empty memories, but with some innate capabilities. They have an extrinsic motivation to establish a nest near a color patch. If the robots are close to the same color patch, they share the same nest. The robots explore their terrain and play together. As their battery level decreases, the robots lose interest in playing, and they focus their attention on exploration in which they map their terrain. They have an extrinsic motivation to remove objects from their terrain. Finally, they return to their nest to sleep and have their batteries replaced.

EXPERIMENT - MOTIVATIONS

- The intrinsic and extrinsic motivations enable the robot to choose what to do.
- Motivations:
  - Establishing a Home
  - Return to the Nest Periodically
  - Return to the Nest When the Batteries Are Low
  - Maintain a Map of the Terrain
  - Play with Other Robots
  - Remove Objects Found in the Terrain (which beyond the boundary of the terrain)
- With these intrinsic motivations, robots can already learn simple collaboration, and by adding sounds and the recognition of sounds, they can learn to communicate with each other to guide their learned collaboration skills.

CASE STUDY

- This paper review the work by Georgeon, et al. to demonstrate how intrinsic motivation can be employed in a schema-based constructivist learning framework.
  - The model in their work, named the intrinsically motivated schema mechanism, aims to address three intertwined issues: intrinsic motivation, autonomously constructed internal state, and adaptive learning.
  - The model is based on three main mechanisms: reinforcement learning with intrinsic rewards, a hierarchical schema mechanism inspired by Piagetian constructivist epistemology, and episodic memory inspired by trace-based reasoning.
  - The results show that the agent can autonomously construct a representation of its situation and learn to use certain schemas to inform its perceptions and determine subsequent actions.

CONCLUSION

- Intrinsic motivation mechanisms demonstrate significant potential in driving AI agents to learn, explore, and adapt in the absence of explicit goals.
- However, the challenge of designing a universal mechanism capable of operating effectively in all circumstances and addressing the full spectrum of inherent satisfactions remains unresolved.
- One critical issue is ensuring that an intrinsically motivated agent does not fall into counterproductive states, such as idleness or excessive distraction. In addition, intrinsic motivation systems cannot be completely isolated from other considerations, such as task requirements, environmental constraints, and the level of expertise of the agent.
- In biological systems, motivation is intricately linked to attention, physiological states, and emotions. It seems improbable that a single, straightforward intrinsic motivation mechanism will meet all the demands of general AI.

THANK YOU