

Empowerment-based learning with a physical developmental curriculum

Personnel:

PI:

Matthew Walter, TTIC

Associated PIs:

Alison Gopnik, UC Berkeley

Students: Kaylene Stocking, UC Berkeley

Google DeepMind:

Igor Mordatch

Shiry Ginosar

Abstract

We propose a tractable method for intrinsically motivated learning by coupling an empowerment-based reward with a learned physical developmental curriculum. Inspired by the self-directed curriculum of babies who independently appropriately choose which objects to manipulate and how, in every stage of their physical development, we hypothesize that a physical developmental curriculum can narrow the search space to make empowerment gain-based intrinsic motivation tractable to use in practice. Once a basic set of self-skills is learned via intrinsic motivation, it can form a basis for more complex action by recomposition through self-directed learning or imitation.

Background

Intrinsically motivated objectives enable autonomous agents to explore novel environments in domains for which reward feedback is sparse or non-existent. However, existing intrinsic objectives are typically designed specifically for the target environment and are thus difficult to generalize to different agents and domains. Recently, cognitive and machine learning researchers considered empowerment gain as a characterization of causal learning in humans and extended it to machines to realize more generalizable learning curricula [Klyubin et al., 2005; Gopnik, 2024]. An agent gains empowerment by performing actions that systematically affect the state of the world, such that varying the action predicts the variation of the outcome. Formally, empowerment-gain measures the mutual

information between actions and their consequences. By learning to act in the world through empowerment gain, an agent ends up with the ability to optimally control its environment.

Unfortunately, since empowerment gain involves computing mutual information over all possible distributions of actions that the agent might take, it is intractable to compute in practice. This intractability holds for computing one step ahead, and attempts to measure empowerment over the next k -steps of actions exacerbate it even further. While methods for approximating empowerment have led to promising results in simple grid world environments and low-dimensional dynamical systems [Mohamed and Rezende, 2015; Dai et al., 2023], these efforts have thus far yet to be scaled successfully to robust, general manipulation.

But if empowerment gain is computationally intractable, how could it characterize human learning? As numerous early learning researchers, such as Susan Carey, Elizabeth Spelke, and Alison Gopnik [Carey, 1985; Gopnik et al., 1997; Gopnik et al., 1999; Spelke, 2022], observed, human children follow a clear progression of evolving physical and mental capabilities as they grow. These evolving capabilities dictate a natural curriculum for their self-directed learning of acting in the world. In particular, the development of the infant's hand, from a tightly closed fist at birth, through controlled arm-based motion, all the way to palmer and pincer grasps, enable the human child to manipulate objects in progressively sophisticated ways from pushing, through full-hand pick-and-place, to fine-grained finger-based manipulation of small objects.

Proposed research

In this project, we adopt a unique approach by drawing inspiration from human babies' physically grounded developmental curriculum. Human children do not explore randomly driven by curiosity but follow a curriculum of self-directed learning dictated, in part, by their developing bodies and minds. We use a similar physical development trajectory as a guide to narrow the search space for empowerment gain-based learning in robots. We focus on a robot arm with a continuous state space (UR5), a camera, and a microphone. This robot learns to control its tabletop environment as it physically develops over time, gaining an increasing number of degrees of freedom. The robot aims to gain empowerment by discovering new ways to act upon itself and the objects around it. We will measure success by the array of manipulation tasks the robot learns despite not receiving an explicit or success-based reward.

The robot will start with a minimal set of available degrees of freedom to narrow the search space for possible actions. It will slowly develop more degrees of freedom as time and learning progresses. At each developmental stage, the robot will naturally

be able to perform only a limited set of tasks with a limited number of objects. Discovering this low-dimensional manifold of states where the agent can maximize its current empowerment gain becomes the goal of learning at each stage. Once the robot has mastered the current set of possible actions and objects, it becomes bored, its empowerment no longer increases, and it is ready to move to the next developmental stage. Notably, going back to explore actions that are already possible in earlier stages of development would also not be of interest to the robot, so the robot is always motivated to explore new possibilities.

The robot's environment must be rich enough to support its curriculum of intrinsically motivated learning and enable it to do more complex things as it develops. It should contain objects that allow for increasing complexity of interaction. One possibility for an adequate selection of objects is a set of toys designed for different stages of human infant development, such as those developed by Montessori for infants. These toys would provide a rich environment for the robot's development. For example, one of the first toys introduced to babies by Montessori is a stick rattle with bells on each side. Such a rattle is placed in their hand when they cannot yet reliably open their closed fists. The babies gain empowerment because they can generate sound when moving their arms. Later, when the infants start to develop pushing capabilities, a roller with bells inside makes a noise when successfully pushed.

We must answer four central research questions during this investigation. First and foremost, does physical development limit the search space so that empowerment-based learning is possible in a realistic, tractable way? Second, can the physical developmental curriculum be learned (as one would imagine humans discovered it through evolution), or does it need to be prescribed somehow? Third, how should we mathematically formulate the empowerment gain objective such that it is tractable to compute in a continuous state space, given the low dimensional action space dictated by the evolving physical capabilities of the robot? Fourth, in what space should empowerment gain be measured? A visual or auditory change? A change in the number of degrees of freedom of the environment (for example, when a puzzle piece fits into place, there are fewer degrees of freedom)? Etc.

Next steps

Once we have proof that the combination of physical curriculum and empowerment gain leads to learning a meaningful set of basic skills, we plan to expand our investigation in several ways.

First, Self-directed learning with one arm on a tabletop environment can only get us so far regarding possible emergent behaviors. After an initial proof of concept, we plan to investigate more complex setups, for example, a two-armed ALOHA robot in a rich environment, such as a kitchen.

Second, Intrinsic-based empowerment learning can lead to a robust self-learned basis of capabilities. However, as we know from children, much learning also happens in a social context, such as through direct demonstrations and apprenticeships. We plan to investigate how we can integrate demonstration information into the empowerment framework. One hypothesis is that a demonstration can narrow the search space for empowerment-based learning, as the agent can infer that the demonstrated trajectory lies in a space of high empowerment gain and, therefore, only searches for nearby actions.

Third, we can measure empowerment gain not only by performing individual actions that affect the world but also by combining previously acquired skills into new capabilities. In this formulation, the agent gains empowerment by discovering new combinations of skills that affect the world in new ways, leading to more complex capabilities than those learned from scratch.

Impact

We know infants learn to manipulate and control their environment primarily self-directedly, but we have not yet managed to design artificial agents that can efficiently learn in this way. Indeed, intrinsically motivated learning, while carrying a promise of building basic capabilities, has yet to be proven effective for learning complex long-horizon actions. The main issue that stands in the way of these methods is the intractability of searching through the space of all possible actions to find rewarding or interesting ones. This project offers a leap forward toward self-learning a basic set of skills by combining empowerment-based learning or learning to control the world around us, with a physical developmental curriculum that narrows down the search space at each developmental stage to only those actions that an agent is currently capable of. Once an arsenal of basic skills is learned, they can be combined into more complex ones independently or with a combination of an imitation goal.

Why is the advantage of funding this research? What are the risks of not funding it?

Why fund the proposal: The proposal considers a long-standing problem in lifelong robot learning: the ability for autonomous agents to learn generalizable policies with little-to-no supervision through a curriculum that emerges from the core concept of

intrinsically motivated play. Motivated by the efficiency of human exploration, our proposed effort brings together researchers from robotics, computer vision, and developmental psychology to realize agents that efficiently learn to interact with their environments through information-theoretic notions of empowerment.

What are the risks of not funding this proposal: While there is a long history of work on intrinsically motivated learning objectives, little attention has been paid to empowerment-based incentives for learning and exploration. This approach, in combination with a physical learning curriculum, has the potential to enable agents to learn an arsenal of basic skill sets, on the basis of which more complex skills can be learned, for example, through imitation.

References

Carey, S. (1985). *Conceptual change in childhood*. MIT Press.

Dai, S., Xu, W., Hofmann, A., and Williams, B. (2023). *An empowerment-based solution to robotic manipulation tasks with sparse rewards*. *Autonomous Robots*, 47(5):617-633.

Gopnik, A. (2024). *Empowerment as Causal Learning, Causal Learning as Empowerment: A bridge between Bayesian causal hypothesis testing and reinforcement learning*. In: [2024] The 29th Biennial Meeting of the Philosophy of Science Association (New Orleans, LA, November 14-17 2024) URL: <https://philsci-archive.pitt.edu/id/eprint/23268> (accessed 2024-07-17).

Gopnik, A., Meltzoff, A. N. (1997). *Words, Thoughts, and Theories*. The MIT Press.

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (1999). *The scientist in the crib: Minds, brains, and how children learn*. William Morrow & Co.

A. S. Klyubin, D. Polani and C. L. Nehaniv (2005). *Empowerment: a universal agent-centric measure of control*. *2005 IEEE Congress on Evolutionary Computation*, Edinburgh, UK, 2005, pp. 128-135 Vol.1, doi: 10.1109/CEC.2005.1554676.

keywords: {Animals;Organisms;Evolution (biology);Adaptive systems;Humans;Feedback;Computer science;Educational institutions;Actuators;Evolutionary computation},

Mohamed, S. and Rezende, D. J. (2015). *Variational information maximisation for intrinsically motivated reinforcement learning*. Advances in neural information processing systems, 28.

Spelke, E. S. (2022). *What Babies Know*. Oxford University Press.