Intrinsic Motivation in Dynamical Control Systems

Stas Tiomkin[a], Ilya Nemenman[b,c,d], Daniel Polani[e], and Naftali Tishby*[f,g]

[a] *Computer Engineering Department, Charles W. Davidson College of Engineering San Jose State University, CA, 95192,* [b] *Department of Physics,* [c] *Department of Biology,* [d] *Initiative in Theory and Modeling of Living Systems, Emory University, Atlanta, GA 30322, USA ,* [e] *Adaptive Systems Research Group, University of Hertfordshire, Hatfield, UK,* [f] *The Rachel and Selim Benin School of Computer Science and Engineering,* [g] *Edmond and Lilly Safra Center for Brain Sciences (ELSC), Hebrew University of Jerusalem, 96906 Israel*

Abstract

Biological systems often choose actions without an explicit reward signal, a phenomenon known as intrinsic motivation. The computational principles underlying this behavior remain poorly understood. In this study, we investigate an information-theoretic approach to intrinsic motivation, based on maximizing an agent's empowerment (the mutual information between its past actions and future states). We show that this approach generalizes previous attempts to formalize intrinsic motivation, and we provide a computationally efficient algorithm for computing the necessary quantities. We test our approach on several benchmark control problems, and we explain its success in guiding intrinsically motivated behaviors by relating our information-theoretic control function to fundamental properties of the dynamical system representing the combined agent-environment system. This opens the door for designing practical artificial, intrinsically motivated controllers and for linking animal behaviors to their dynamical properties.

Keywords— information capacity | sensitivity gain | stabilization | predictive information

Introduction

Living organisms are able to generate behaviors that solve novel challenges without prior experience. Can this ability be explained by a single, generic mechanism? One proposal is that novel, useful behaviors can be generated through *intrinsic motivation* [1], which is defined informally as a set of computational algorithms that are derived directly from the intrinsic properties of the organism-environment dynamics and not specifically learned.

Increasingly, there is a move away from reinforcement learning and its extrinsically specified reward structure [2, 3] in the theory and practice of artificial agents, robots, and machine learning more generally [4–20]. A specific class of such intrinsic motivation algorithms for artificial systems is known as *empowerment maximization*. It proposes that agents should maximize the mutual information [21] between their potential actions and a subsequent future state of the world [22]. This corresponds to maximizing the diversity of future world states achievable as a result of the chosen actions, potentiating a broader set of behavior options in the future. Intrinsically motivated synthetic agents develop behaviors that are atypical for inanimate engineered systems and often resemble those of simple living systems. Interestingly, potentiating future actions is also a key part of the success of modern reward-based training algorithms [8, 23, 24].

Despite the successes of empowerment maximization, it remains unclear how well it can be used as a general intrinsic motivation principle. There are many different versions of intrinsic motivation related to empowerment, and their relation to each other is unknown [20, 23, 25]. Additionally, most work on empowerment maximization has relied on simulational case studies and ad hoc approximations, and analytical results are scarce. In order to gain insight, it is important to link empowerment to other, better-understood characterizations of the systems in question. Finally, calculating the mutual information between two interlinked processes in the general case is a challenging task [26, 27], which has so far limited the use of empowerment maximization to simple cases.

In this work, we unify different versions of intrinsic motivation related to the empowerment maximization paradigm. Here our main contribution is in showing analytically that empowerment-like quantities are linked to the sensitivity of the agent-environment dynamics to the agent's actions. This connects empowerment maximization to well-understood properties of dynamical systems. Since highly sensitive regions of the dynamics potentiate many diverse future behaviors, the connection to dynamical systems also explains why empowerment-based intrinsic motivations succeed in generating behaviors that resemble those of living systems.

The analytical results allow us to develop a practical computational algorithm for calculating empowerment for complex scenarios in the continuous time limit, which is the second major contribution of the paper. We apply the algorithm to standard benchmarks used in intrinsic motivation research [14, 28, 29]. Specifically, a controller based on the efficient calculation of empowerment manages to balance an inverted pendula without extrinsic rewards. This opens the door for designing complex robotic intrinsically motivated agents with systematically computed — rather than heuristically estimated — empowerment.

**Prof. Naftali Tishby passed away when this work was in development. This project began under his leadership when Stas Tiomkin was a PhD student in his group. The rest of the authors agree that he should be a senior author on this manuscript, but his consent for this was not obtained.*