

Chapter 6

A Primer on Reinforcement Learning in the Brain: Psychological, Computational, and Neural Perspectives

Elliot A. Ludvig

University of Alberta, Canada

Marc G. Bellemare

University of Alberta, Canada

Keir G. Pearson

University of Alberta, Canada

ABSTRACT

In the last 15 years, there has been a flourishing of research into the neural basis of reinforcement learning, drawing together insights and findings from psychology, computer science, and neuroscience. This remarkable confluence of three fields has yielded a growing framework that begins to explain how animals and humans learn to make decisions in real time. Mastering the literature in this sub-field can be quite daunting as this task can require mastery of at least three different disciplines, each with its own jargon, perspectives, and shared background knowledge. In this chapter, the authors attempt to make this fascinating line of research more accessible to researchers in any of the constitutive sub-disciplines. To this end, the authors develop a primer for reinforcement learning in the brain that lays out in plain language many of the key ideas and concepts that underpin research in this area. This primer is embedded in a literature review that aims not to be comprehensive, but rather representative of the types of questions and answers that have arisen in the quest to understand reinforcement learning and its neural substrates. Drawing on the basic findings in this research enterprise, the authors conclude with some speculations about how these developments in computational neuroscience may influence future developments in Artificial Intelligence.

DOI: 10.4018/978-1-60960-021-1.ch006

INTRODUCTION

The last decade has seen a proliferation of research exploring the neural and psychological mechanisms of reinforcement learning (for some good reviews and perspectives, see Dayan & Daw, 2008; Doya, 2007; Maia, 2009; Niv, 2009; Rangel, Camerer, & Montague, 2008; Schultz, 2002, 2007). This flourishing area of computational neuroscience draws on the expertise and knowledge in many sub-disciplines, including psychology, neuroscience, computer science, philosophy, and economics, amongst others. This remarkable confluence of fields was catalyzed by the discovery of a close correspondence between the behaviour of dopamine neurons in classical conditioning tasks and the prediction error in the temporal-difference (TD) algorithm from reinforcement learning (Montague, Dayan & Sejnowski, 1996; Schultz, Dayan, & Montague, 1997; Sutton, 1988; Sutton & Barto, 1990; see Figure 5). The import of this finding has filtered outward from a strikingly successful model of the neural basis of a simple conditioning behavior in animals to theoretical models of human economic decision making and, in part, to an entire field of neuroeconomics (e.g., Glimcher et al., 2009; Platt & Huettel, 2008; Rangel et al., 2008; Schultz, 2009).

Our goal in this chapter is two-fold. First, we aim to provide a primer of basic introductory materials in three of the constitutive disciplines of this enterprise—psychology, computer science, and neuroscience—to facilitate access by Artificial Intelligence (AI) researchers and other computational neuroscientists into this exciting field. As our second goal, we will not directly re-tread the ground covered in detail by the many comprehensive recent reviews, but rather we use some selective examples of reinforcement-learning research and show how this multi-disciplinary enterprise has helped inform and been informed by these basic lines of inquiry.

In considering the relationship between observed behaviour, computational models, and neural mechanisms, Marr's (1982) three levels of analysis prove very instructive. Marr proposed that any information-processing system can be analyzed at three different levels: the computational or functional, the algorithmic or representational, and the implementational. At the computational level, one specifies the goals and objectives of the system. What does the system do? For example, the computational goal for classical conditioning might be the prediction of important biological events. Second, at the algorithmic level, one specifies the step-by-step procedure by which this function is accomplished. What algorithm or procedure does the system use to accomplish the computational goals? Again, for classical conditioning, this might be the Rescorla-Wagner rule (Rescorla & Wagner, 1972) or the TD algorithm (Sutton & Barto, 1990) or any other set of rules that describe how the computation happens. Finally, at the implementational level, the important details of how these different algorithms and representations can be instantiated in neural tissue or other mediums are laid out. How are these algorithms physically realized? One example would be the equating of the reward-prediction error from reinforcement learning with the burst firing of dopamine neurons (Schultz et al., 1997). A full explanation of any information-processing system would require adequate accounts at each of the three levels of analysis.

Our chapter follows Marr's proposal by dividing this introduction to neural reinforcement learning into three sections that roughly correspond to his three levels of analysis. Section 1 describes the computational problems facing creatures in simple learning and decision-making situations and summarizes some of the attempts within psychology to characterize the algorithms that may be used by animals to solve these problems. Section 2 introduces the core computational ideas of reinforcement learning (RL) and shows how ideas from RL are particularly well suited as potential

32 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/primer-reinforcement-learning-brain/49232

Related Content

Telemetry Data Mining Techniques, Applications, and Challenges

Sara Ahmed, Tarek Gaber and Aboul Ella Hassanien (2017). *Handbook of Research on Machine Learning Innovations and Trends* (pp. 915-924).

www.irma-international.org/chapter/telemetry-data-mining-techniques-applications-and-challenges/180978

A Survey of Different Approaches for the Class Imbalance Problem in Software Defect Prediction

Abdul Waheed Darand Sheikh Umar Farooq (2022). *International Journal of Software Science and Computational Intelligence* (pp. 1-26).

www.irma-international.org/article/a-survey-of-different-approaches-for-the-class-imbalance-problem-in-software-defect-prediction/301268

SOM-Based Class Discovery for Emotion Detection Based on DEAP Dataset

Aladdin Ayesh, Miguel Arevalillo-Herráez and Pablo Arnau-González (2018). *International Journal of Software Science and Computational Intelligence* (pp. 15-26).

www.irma-international.org/article/som-based-class-discovery-for-emotion-detection-based-on-deap-dataset/199014

Need for Internet of Things

V. J. K. Kishor Sontia and Sundari G. (2019). *Edge Computing and Computational Intelligence Paradigms for the IoT* (pp. 46-59).

www.irma-international.org/chapter/need-for-internet-of-things/232001

Pricing and Lot-Sizing Decisions in Retail Industry: A Fuzzy Chance Constraint Approach

R. Ghasemy Yaghin, S. M. T. Fatemi Ghomi and S. A. Torabi (2014). *Mathematics of Uncertainty Modeling in the Analysis of Engineering and Science Problems* (pp. 268-289).

www.irma-international.org/chapter/pricing-and-lot-sizing-decisions-in-retail-industry/94516