Ergodicity breaking in Reinforcement Learning: When expected values are not the value you expect

Bert Verbruggen
Vrije Universiteit Brussel, DataLab
e-mail: bert.verbruggen@vub.be

Arne Vanhoyweghen
Vrije Universiteit Brussel, DataLab

Vincent Ginis
Vrije Universiteit Brussel, DataLab

Keywords: Reinforcement Learning, Decision Theory, Ergodicity Economics

Imagine being in a casino with the challenge to choose between two bets: one that guarantees a slight, fixed increase in wealth and another with a stochastic outcome where your wealth can increase significantly, or you can lose all of it. How would you explain your decision? This classic example highlights the complexities of human decision-making and has been widely used to illustrate theories on choice dynamics.

In traditional economic models, decisions are often based on expected outcomes, calculated as the average outcome for a group of identical, independent agents playing the same game. However, the mathematical foundations of most theories underlying these economic models may need to be completed. These foundations are based on psychological interpretations and arguments and often involve finding the highest expected return on an outcome and assigning a mapping function to the data. The theory assumes optimizing toward the maximum expected group average is the best solution. However, this approach may only sometimes be appropriate, as it relies on the assumption of ergodicity, which assumes that time averages and ensemble averages are equivalent. Many processes in the real world do not follow this assumption, breaking ergodicity and leading to consequences that traditional economic models do not accurately capture.

Ergodicity Economics is a field of study that examines the impact of these dynamics, where optimizing toward expected values may not be the optimal strategy. By considering the dynamics of time growth and the breaking of ergodicity, precise predictions can be made with a solid mathematical foundation. This concept can also be applied to human decision-making, which often plays a central role in the dynamics described by economic science.

Furthermore, this issue has implications in the Reinforcement Learning (RL) field, which focuses on training agents to make optimal decisions and translate them into policies. The Bellman equation, a fundamental concept in RL, heavily relies on expected values. However, human decision-making only sometimes aligns with optimizing toward expected values, as predicted by ergodicity economics. In processes where time growth differs from a group average, using the group average as an expected value estimate can lead to inaccurate results.

In our work, we employ standard techniques in RL to investigate whether agents can be trained to optimize policies based on time-growth averages rather than expected values. We apply the thought experiment from the beginning of this abstract and translate it into the realm of RL in various conceptual cases.

Our research highlights the importance of considering time growth in RL, as it can result in significantly different outcomes compared to traditional approaches that solely optimize toward expected values. Our findings demonstrate that agents can learn and optimize their policies based on time-growth averages. By incorporating the breaking of ergodicity into decision-making processes, we can gain a more comprehensive understanding of human decision-making and improve the accuracy of economic models and RL algorithms.