Voting Mechanisms in Reinfocement Learning

Jost Alemann

Institute of intelligent & cooperating systems
Otto von Guericke University
Magdeburg, Germany
jost.alemann@ovgu.de

Abstract—This paper aims to deliver an overview over how voting mechanisms can be used in reinforcement learning. Voting mechanisms and their properties are first introduced to the reader and then explained in more detail by describing their application in related work in the field of multi-agent systems and reinforcement learning.

Index Terms—voting, reinforcement learning, multi-agent systems

I. INTRODUCTION

In a democratic society choices are not made by a dictator but by taking the preferences of the whole society into account. Therefore each member of the society is entitled to cast a vote which represents it's preferences. Votes are then evaluated and a decision is derived by a given voting scheme. Plurality where the choice with the most votes wins, might be the most common voting scheme. Still there are many different voting schemes each of which can or cannot fulfil certain properties. This lies mainly in the subject of social choice theory and therefore is only described briefly if needed in the following. The concept of considering multiple individuals' preferences by using a voting mechanism to make choices can be transferred to multi-agent reinforcement learning systems. This is motivated by the expectation that agents combine their limited perception and knowledge of the environment by deciding together. Therefore they are expected to obtain better results than agents that choose actions based on only their own perception. [1]

However designing a voting mechanism is non-trivial since the voting system has to consider different aspects like security, fairness and robustness [2]. Arrow's Impossibility Theorem even implies that no rank-order voting mechanisms can be designed to be completely fair. This will be explained later on in Section II. Ongoing research constantly tries to improve the security, robustness and fairness of voting mechanisms.

To give an overview over voting mechanisms in reinforcement learning Section II introduces basic principles of social choice theory. Section III describes related work to highlight use cases of voting mechanisms in reinforcement learning as well as ongoing research trying to improve such mechanisms.

II. BASIC PRINCIPLES

A. Different Voting Schemes

To discuss properties of voting systems we have to introduce those systems first.

B. Properties of Voting Schemes

- *Unanimity*: If every voter prefers option X over option Y, then the society prefers X over Y.
- *Independence of irrelevant alternatives*: If every voter prefers option X over option Y and option Z is removed without changing the former relation, the societies preference of X over Y also remains unchanged.
- *No dictatorship*: No single voter posseses the power to always determine the group's preference.

C. Arrow's Impossibility Theorem

Arrow's Impossibility Theorem is of great importance for the design of a voting mechanisms. It states that no rank-order voting scheme can fulfil the properties unanimity, independece of irrelevant alternatives and no dictatorship at the same time.

- Arrow's Impossibility Theorem
  - states that no rank-order electoral system can be designed to be perfectly fair matching these criterias:
  - If every voter prefers alternative X over alternative Y, then the group prefers X over Y.
  - If every voter's preference between X and Y remains unchanged, then the group's preference between X and Y also remains unchanged.
  - There is no "dictator": no single voter posseses the power to always determine the group's preference.

III. RELATED WORK

REFERENCES

[1] I. Partalas, I. Feneris, and I. Vlahavas, "A hybrid multiagent reinforcement learning approach using strategies and fusion," *International Journal on Artificial Intelligence Tools*, vol. 17, no. 05, pp. 945–962, 2008.
[2] J. Pitt, L. Kamara, M. Sergot, and A. Artikis, "Voting in multi-agent systems," *The Computer Journal*, vol. 49, no. 2, pp. 156–170, 2006.