Reinforcement Learning for Argumentation: Describing a PhD research

Sultan Alahmari
University of York
Department of Computer Science
Deramore Lane, Heslington, York,
YO10 5GH, UK
smsa500@york.ac.uk

Tommy Yuan
University of York
Department of Computer Science
Deramore Lane, Heslington, York,
YO10 5GH, UK
tommy.yuan@york.ac.uk

Daniel Kudenko
University of York
Department of Computer Science
Deramore Lane, Heslington, York,
YO10 5GH, UK
daniel.kudenko@york.ac.uk

1 OVERVIEW

Artificial intelligence (AI) is increasingly studied in many fields such as philosophy, law and decision making. One of the approaches to AI is the use of agent and multi-agent systems. Agents are key element for building complex large-scale distributed systems[9]. In multi-agent systems, each agent interacts with the environment and communicates with other agents in order to achieve the designated goal. Communication means to share and exchange information, cooperate and coordinate with each other in order to achieve a common goal.

Argumentation is a type of communication between agents and a process attempting to form an agreement about what to believe. There has been increasing research in argumentation and dialogue systems in the past decade[23]. The agent as a dialogue participant needs sophisticated dialogue strategies in order to make high quality dialogue contributions. By reviewing the state of art literature in computerised dialogue systems (e.g.[21];[22]), it is observed that their dialogue strategies (i.e. strategic heuristics) are hardwired into the computational agent. One of the main issues with this is that an agent might be incapable of dealing with new dialogue situations that have not been coded, and indeed this is an impossible task given the dynamic nature of argumentation. It would be ideal to make an agent search for an optimal strategy by itself e.g. via trial and error, and thus the agent with the best strategy will win the argument [8].

Machine learning has an important role to play in order to meet these challenges. To make agents learn the dialogue strategies, it would be more flexible for them to make an argument through exploration (trial and error). It is believed that learning can make agents more flexible to adapt to new environments and new dialogue situations. One of the popular machine learning approaches with regards to learning agents is known as reinforcement learning (RL).

Reinforcement learning focuses on how to map an action for each state by interacting with the environment and observing the state change[15]. Sutton and Barto[15] define reinforcement learning as an agent learning what to do and how to connect each situation with an action to maximise the cumulative reward. The learner or agent is not told what action should be taken, rather the learner needs to explore a policy that yields the maximum cumulative reward by trying them out. In reinforcement learning, the agent interacts with the environment by taking an action and receiving a reward for the action taken as seen in figure 1. To make an
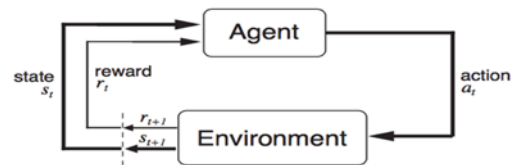

Figure 1: Reinforcement learning agent-environment interaction

agent learn to argue there is a need to identify states, actions, environment and the rewards. In this research abstract argumentation systems (AAS) is initially used [5] to represent the argumentation. Reasons are listed as follows:

(1) It has the ability to represent informal human reasoning in a way that a computer can perform calculation. In this way, argumentation bridges the gap between human and machine reasoning[11].

(2) It is easier to compute acceptable arguments in order to evaluate variance argument semantics e.g. grounded extension.

(3) It provides a great opportunity for the agent to explore the relationship between arguments.

(4) It is a powerful method to solve problems since it can be easily implemented in logic programming[5].

The classical state representation of agents in literature (e.g. [19];[4]) involves states being represented as nodes in the argumentation graph and action by the attack relation between arguments.

The main objective of our research is to investigate whether reinforcement learning agent can be used to create an argumentation AI with improved performance and efficiency comparable to state-of-the-art systems. Performance is related to how well the agent learns over time. The measurement of performance for a good argumentation; for instance, is whether argumentation can be won or lost or how many arguments from a learning agent obtains accepted against other heuristic strategy agents. Efficiency is related to whether the agent can learn within a limited or insufficient time. So the aim is to find out if the agent can learn rapidly or not. It should also ensure that the agent obtains full knowledge from the environment so as to be able to use an efficient method to find an optimal decision for each state [17].

The light of this hypothesis, the following steps will be taken:

(1) Initially, a basic abstract argument game model is used due to its simplicity in implementing arguments. This in turn makes it possible to investigate how reinforcement learning can be applicable to a simple dialogue scenario.

(2) Evaluation of an argumentation setting with a human or another AI agent by observing learning performance over time.

(3) Investigating suitable means for reinforcement learning of a complicated dialogue scenario and studying the results in order to generalise the RL method. Complicated dialogue scenario involves more move types e.g. questions, challenges, assertion, withdrawal and moves from abstract argument level to propositional level.

This work will also investigate other different scenario such as backtracking ([14];[16]), arguments content, weight of individual argument amongst others[6]. Additionally, challenging issue such as states representation[1] as well as reward function will also be explored.

To prove the hypothesis, we have built the argumentation software to facilitate experiment for reinforcement learning agent arguing against different agents. A software testbed, *Argumento+*, named after its predecessor *Argumento* as reported in[24], has been built using the Java programming language. *Argumento+* contains the RL agent as well as three other agents namely, random, maximum probability utility and minimum probability utility agent for the sake of the evaluation. The agents play abstract argument games. RL agent plays game against them to maximise the cumulative reward by winning more games. Indeed, if RL agent win the game, it will receive rewards based on the number of acceptable arguments i.e. grounded extensions. We considered grounded extension because it contains an argument that has no doubt in comparison with other arguments [19]. Consequently, it will be a more acceptable argument.

We have performed an initial experiment to investigate whether RL agent learns to argue against baseline agents [1]. RL agent adopts a commonly used RL method, that is Q-learning algorithm. The aim of Q-learning is to allow an agent to learn through experience and map each state with an action by choosing the maximum value from the Q-table which is updated after each episode. The initial experiment and evaluation generally encourage the adopting of reinforcement learning agent in argumentation with a long term delayed reward which are considered as grounded extensions [1].

In the future, this work will attempt to suggest ways to improve the RL agent performance by carrying out further works on the initial experiment results. The state representation of the arguments still needs to be more sophisticated in order to make each one of them unique [1], and as a result this will make it easy for the agent to distinguish between states. Even though initial suggestion pointed at

making the state a combination of the current state and previous state, it is still difficult to uniquely identify each state. To sort out this issue, it will be worth investigating if this can be resolved by representing each state as: $(levelOfTree, agentID, currentState, previousState)$

Backtracking ([14];[16]) will also be considered to improve the simple argument game by developing some game rules in [18]. Moreover, to make the game more competitive and effective it is important to make the agent consider the opponent's strategy[7]. Hence, the learning agent needs to consider how to learn to argue with the opponent by expanding its knowledge base with new arguments. In addition, in complex argumentation scenario, we need to consider moving from high level abstraction to the argument contents by using propositional logic. Weighted arguments will also be considered in this research since some arguments are more important than others. We will consider choosing a suitable argument model for the complicated scenario. There are many models; for instance, Prakken's dialogue game *Persuasion with dispute* [13], Bench-Capon TDG dialogue game ([2];[3]), DC by Mackenzie [10], Utilisation by Moore in [12] and DE system (Yuan et al. [20]), all of these models will be critically reviewed.

REFERENCES

[1] Sultan Alahmari, Tommy Yuan, and Daniel Kudenko. 2017. Reinforcement learning for abstract argumentation: Q-learning approach. In *Adaptive and Learning Agents workshop (at AAMAS 2017)*.

[2] Trevor JM Bench-Capon. 1998. Specification and implementation of Toulmin dialogue game. In *Proceedings of JURIX*, Vol. 98. 5–20.

[3] Trevor J. M. Bench-Capon, T Geldard, and Paul H Leng. 2000. A method for the computational modelling of dialectical argument with dialogue games. *Artificial Intelligence and Law* 8, 2 (2000), 233–254.

[4] Heriberto Cuayáhuitl, Simon Keizer, and Oliver Lemon. 2015. Strategic dialogue management via deep reinforcement learning. *arXiv preprint arXiv:1511.08099* (2015).

[5] Phan Minh Dung. 1995. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial intelligence* 77, 2 (1995), 321–357.

[6] Paul E Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge. 2011. Weighted argument systems: Basic definitions, algorithms, and complexity results. *Artificial Intelligence* 175, 2 (2011), 457–486.

[7] Katie Long Genter, Santiago Ontañón, and Ashwin Ram. 2011. Learning Opponent Strategies through First Order Induction.. In *FLAIRS Conference*. 1–2.

[8] Piotr S Kośmicki. 2010. A platform for the evaluation of automated argumentation strategies. In *International Conference on Rough Sets and Current Trends in Computing*. Springer, 494–503.

[9] Ryszard Kowalczyk. 2014. Intelligent Agent Technology Research. https://www.swinburne.edu.au/ict/success/research-projects-and-grants/intelligent-agent/. (2014). [Online; accessed 06-April-2017].

[10] Jim D Mackenzie. 1979. Question-begging in non-cumulative systems. *Journal of philosophical logic* 8, 1 (1979), 117–133.

[11] Sanjay Modgil, Francesca Toni, Floris Bex, Ivan Bratko, Carlos I Chesnevar, Wolfgang Dvořák, Marcelo A Falappa, Xiuyi Fan, Sarah Alice Gaggl, Alejandro J García, and others. 2013. The added value of argumentation. In *Agreement technologies*. Springer, 357–403.

[12] David John Moore. 1993. *Dialogue game theory for intelligent tutoring systems*. Ph.D. Dissertation. Leeds Metropolitan University.

[13] Henry Prakken. 2001. Relating protocols for dynamic dispute with logics for defeasible argumentation. *Synthese* 127, 1 (2001), 187–219.

[14] Henry Prakken. 2010. Argumentation Logics: Games for abstract argumentation. http://www.staff.science.uu.nl/~prakk101/al/chongqing10.html. (2010). [Online; accessed 01-April-2017].

[15] Richard S Sutton and Andrew G Barto. 1998. *Reinforcement learning: An introduction*. Vol. 1. MIT press Cambridge.

[16] Gerard AW Vreeswik and Henry Prakken. 2000. Credulous and sceptical argument games for preferred semantics. In *European Workshop on Logics in Artificial Intelligence*. Springer, 239–253.

[17] Eric Wiewiora. 2004. *Efficient Exploration for Reinforcement Learning*. Ph.D. Dissertation. Citeseer.

[18] Michael Wooldridge. 2002. *An introduction to multiagent systems*. John Wiley & Sons.

[19] Michael Wooldridge. 2009. *An introduction to multiagent systems*. John Wiley & Sons.

[20] Tangming Yuan, David Moore, and Alec Grierson. 2003. Computational Agents as a Test-Bed to Study the Philosophical Dialogue Model" DE": A Development of Mackenzie's DC. *Informal Logic* 23, 3 (2003).

[21] Tangming Yuan, David Moore, and Alec Grierson. 2007. A human–computer debating system prototype and its dialogue strategies. *International Journal of Intelligent Systems* 22, 1 (2007), 133–156.

[22] Tangming Yuan, David Moore, and Alec Grierson. 2008. A human-computer dialogue system for educational debate: A computational dialectics approach. *International Journal of Artificial Intelligence in Education* 18, 1 (2008), 3–26.

[23] Tangming Yuan, Jenny Schulze, Joseph Devereux, and Chris Reed. 2008. Towards an arguing agents competition: Building on argumento. In *Proceedings of IJCAI2008 Workshop on Computational Models of Natural Argument*.

[24] Tangming Yuan, Viðar Svansson, David Moore, and Alec Grierson. 2007. A computer game for abstract argumentation. In *Proceedings of the 7th Workshop on Computational Models of Natural Argument (CMNA07)*.

3