In: 2023 IEEE Conference on Artificial Intelligence (CAI 2023), 5-6 June 2023, Santa Clara, USA, pp. 85-86
DOI: 10.1109/CAI54212.2023.00044

Explainability of AI-Driven Air Combat Agent

1st Emre Saldiran
School of Aerospace, Transport and Manufacturing
Cranfield University
Cranfield, United Kingdom
emre.saldiran@cranfield.ac.uk

2nd Mehmet Hasanzade
School of Aerospace, Transport and Manufacturing
Cranfield University
Cranfield, United Kingdom
mehmet.hasanzade@cranfield.ac.uk

3rd Gokhan Inalhan
School of Aerospace, Transport and Manufacturing
Cranfield University
Cranfield, United Kingdom
inalhan@cranfield.ac.uk

4rd Antonios Tsourdos
School of Aerospace, Transport and Manufacturing
Cranfield University
Cranfield, United Kingdom
a.tsourdos@cranfield.ac.uk

***Abstract*—In safety-critical applications, it is crucial to verify and certify the decisions made by AI-driven Autonomous Systems (ASs). However, the black-box nature of neural networks used in these systems often makes it challenging to achieve this. The explainability of these systems can help with the verification and certification process, which will speed up their deployment in safety-critical applications. This study investigates the explainability of AI-driven air combat agents via semantically grouped reward decomposition. The paper presents two use cases to demonstrate how this approach can help AI and non-AI experts to evaluate and debug the behavior of RL agents.**

***Index Terms*—explainable, reinforcement learning, reward decomposition, air combat**

I. Introduction

Using AI-driven Autonomous Systems (ASs) with highly complex and novel behavior in the real world requires the explainability of the model to increase users, developers, and policymakers' trust [1]. Explainability is also envisioned to play an important role in the verification, certification, and adaptability of AI-driven ASs in safety-critical applications [2]. The role of explainability is to present the users with the rationale behind AI actions to enable them to build higher-quality mental models.

Explainable AI methods have been applied to other fields such as the medical domain, Judicial System, and banking/financial domain [3]. However, no study applies explainability methods to AI-driven air combat agents. Air combat is a highly dynamic and challenging problem in which the pilot has to make split-second sequential decisions. This problem can be tackled with an AI method called Deep Reinforcement Learning (DRL) as shown during DARPA's AlphaDogFight Trial [4].

This paper introduces explainability to the air combat RL agent via reward decomposition. The RL agent is trained on a 3DoF air combat simulator. A User-friendly interface is developed to help with visualization. Two example use cases tailored for both AI and non-AI experts are shown. This paper's main contributions include exploring the explainability of an AI-driven air combat agent through reward decomposition, utilizing the resulting explanations to debug issues in the RL training process, and demonstrating the application of the decomposed rewards to provide a warning when the agent's behavior deviates from the user's expectations.

This work is funded by BAE Systems.

II. Methodology

In traditional RL, the reward signal is fed as a scalar value. In reward decomposition, different Deep Q-Networks (DQNs) are trained for each reward type. The formulation and derivation of reward decomposition are given in the existing literature [5].

The air combat environment is represented by reduced order 3DOF aircraft dynamic as follows,

$$\begin{aligned} \dot{V_c} &= \Delta V_c & \dot{V} &= K_V\left(V_c - V\right) & \dot{x} &= V\cos\chi \\ \dot{\chi_c} &= \Delta\chi_c & \dot{\chi} &= K_\chi\left(\chi_c - \chi\right) & \dot{y} &= V\sin\chi \end{aligned}$$

where ΔV_c is $\Delta\chi_c$ is the commanded delta velocity and heading angle. K_v and K_χ are the velocity and heading angle gains. V and χ are the velocity and heading angle states, respectively. Eight discrete actions are formed by combining maximum and minimum speed and heading angle commands, with the addition of a "do nothing" action to create a total of nine actions. Observation space consists of ATA, AA, relative heading, and LOS vector information. Following observation vector used during training.

$$s_t = \left[LOS_x, LOS_y, \|LOS\|, \chi^{red} - \chi^{blue}, ATA, AA\right]^T$$

To group the rewards into semantically meaningful types, the reward components are chosen as Antenna Train Angle (ATA), Aspect Angle (AA), and Line-Of-Sight (LOS) as shown in Fig. 1.

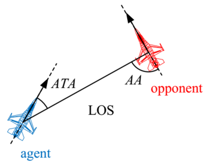

Fig. 1. ATA, AA, and LOS geometric representation.

ATA, AA, and LOS rewards are calculated using an exponential formula to provide a continuous reward signal which helps with gradient descent during training.

$$r_{ATA} = e^{-|ATA|}, \quad r_{AA} = e^{-|AA|}$$
$$r_{LOS} = min(1, e^{-LOS/1000-1})$$

Finally, reward components are multiplied by a constant depending on their importance to create a reward vector.

$$\vec{r} = [r_1, r_2, r_3]^T = [0.4, 0.4, 0.2]^T \odot [r_{ATA}, r_{AA}, r_{LOS}]^T$$

III. Experiment

In this section, we investigated how to apply explainability to evaluate why an agent took a certain action for AI experts and non-AI experts.

In the first use case, AI experts use semantically meaningful reward components to evaluate and debug RL agent behavior, identifying areas of underperformance or unexpected behavior. RL agents trained with different hyperparameters and different initial randomization ranges can have very similar episodic returns, but completely different behavior during evaluation. However, a comparison of decomposed mean Q-values for each reward type shows that the mean Q-value of LOS reward is lower than the other components as shown in Fig. 2. This could be due to the agent only seeing the target far away at the beginning of the episode, since the goal is to get close, which results in low data distribution for far-away scenarios.

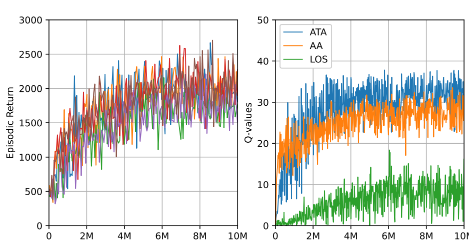

Fig. 2. Left: Episodic return of multiple training. Right: Comparison of component Q-values of multiple training.

In the second use case, decomposed rewards are utilized as a warning mechanism when the agent's behavior does not align with the user's expectations. For example, in Fig. 3 despite the bandit aircraft being far away and pointing right, the RL agent to chose a speed-up turn left action. The expected behavior would be to choose to speed up and turn right action and have a higher LOS Q-value to justify it. The misalignment

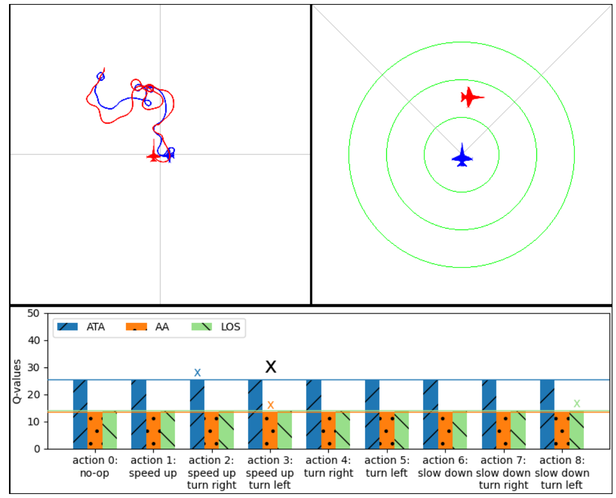

Fig. 3. Trained RL agent tracking performance. The upper left plot is a bird's eye view. The upper right figure is perspective from the blue agent body frame and each green circle has a 1,000 m radius. The lower figure is decomposed reward bar chart, the black x sign represents the chosen action while the other x signs represent the action associated with the maximum expected reward of each DQN for their respective reward type.

both in the selected action and the reasons for that action can be used to stop using this RL agent. By presenting the rewards in a semantically meaningful way, users can better evaluate the RL agent action and make informed decisions about how to interact with it. This increased transparency can help to build trust between the user and the agent, as well as provide a mechanism for users to intervene or adjust the behavior of the agent when necessary.

IV. Conclusion

In this paper, the explainability of AI-driven air combat agent is studied. Two different use cases are demonstrated for both AI and non-AI experts. The first use case identifies and debugs training process shortcomings by using semantically decomposed reward types. The second use case uses decomposed rewards as a warning mechanism when there is a misalignment between non-AI expert and RL agent expectations. Future work will combine hierarchical methods with behavioral reward types instead of geometrical ones and conduct a user study to measure changes in people's trust in RL agents in the air combat tactic generation problem.

References

[1] Davinder Kaur, Suleyman Uslu, Kaley J Rittichier, and Arjan Durresi. Trustworthy artificial intelligence: a review. *ACM Computing Surveys (CSUR)*, 55(2):1–38, 2022.

[2] David Gunning, Eric Vorm, Yunyan Wang, and Matt Turek. Darpa's explainable ai (xai) program: A retrospective. *Authorea Preprints*, 2021.

[3] Nadia Burkart and Marco F Huber. A survey on the explainability of supervised machine learning. *Journal of Artificial Intelligence Research*, 70:245–317, 2021.

[4] Alphadogfight trials foreshadow future of human-machine symbiosis, 2020.

[5] Zoe Juozapaitis, Anurag Koul, Alan Fern, Martin Erwig, and Finale Doshi-Velez. Explainable reinforcement learning via reward decomposition. In *IJCAI/ECAI Workshop on explainable artificial intelligence*, 2019.

2023-08-02

Explainability of AI-driven air combat agent

Saldiran, Emre

IEEE

Saldiran E, Hasanzade M, Inalhan G, Tsourdos A. (2023) Explainability of AI-driven air combat agent. In: 2023 IEEE Conference on Artificial Intelligence (CAI 2023), 5-6 June 2023, Santa Clara, USA, pp. 85-86

https://doi.org/10.1109/CAI54212.2023.00044