Making a robot stop a penalty

van Heusden, R.; Visser, A.

Publication date
2018

Document Version
Final published version

Published in
30th Benelux Conference on Artificial Intelligence

Citation for published version (APA):
van Heusden, R., & Visser, A. (2018). Making a robot stop a penalty. In M. Atzmueller, & W. Duivesteijn (Eds.), *30th Benelux Conference on Artificial Intelligence: BNAIC 2018 Preproceedings : November 8-9, 2018, Jheronimus Academy of Data Science (JADS), 's-Hertogenbosch, The Netherlands* (pp. 89-90). (BNAIC; Vol. 30). Jheronimus Academy of Data Science.

30th Benelux Conference on Artificial Intelligence

BNAIC 2018 Preproceedings

November 8-9, 2018

Jheronimus Academy of Data Science (JADS), 's-Hertogenbosch, The Netherlands

Editors:
Martin Atzmueller
CSAI/JADS, Tilburg University
Wouter Duivesteijn
Eindhoven University of Technology

Making a Robot Stop a Penalty

Ruben van Heusden Supervisor : dhr. dr. A. Visser

Intelligent Robotics Lab, Faculty of Science, University of Amsterdam

Abstract. This research will focus on is the question of how to train a humanoid robot to stop a penalty using Q Learning combined with Transfer Learning on a neural network. The Transfer Learning technique proved allow the robot to learn stop the robot without domain knowledge in the form of an intermediate reward.

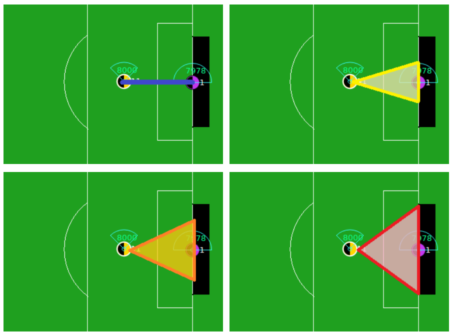
Keywords: Reinforcement Learning · Transfer Learning · RoboCup

Context This is a short extract of my thesis [4], which addresses the problem of decision making for an autonomous goalkeeping robot in a penalty shootout. A solution would contribute to the goal of having robot soccer players outperform human soccer players by 2050, which is still an open challenge in artificial intelligence [1]. The problem is complex as the amount of information provided to the goalkeeper is limited and the goalkeeper only has a limited time frame within it has to make a decision. Previous approaches often involved a human coding complex behavioral rules for the robots, which is very inflexible and the process of constructing behavioral rules and discovering the right parameters can be very time consuming [3].

Method To allow a simulated robot to efficiently learn an optimal strategy by itself, a sensible reward function should be used, which requires domain knowledge. Instead, the domain knowledge is used to present the robot a succession of tasks, increasing in difficulty and using the experience gained in previous tasks to guide the robot in the following tasks [5]. In this research, 4 tasks increasing in difficulty (possible shooting angles for the penalty kicker) were used as shown in Figure 1, where the agent progressed to the next task when it was able to stop 30 penalties in a row. By using this approach, the agent would not necessarily need an intermediate reward, as it might be able to learn the basic concepts of the game in a simple environment and build on that knowledge to have success in more difficult tasks.

Results The approach presented in this thesis uses Q learning combined with a deep neural network [2] applied to discrete state and actions spaces and transfer learning to learn a behavioral policy. With intermediate rewards the algorithm is able to successfully stop a penalty in $\sim$96% of the trials, with an average training length of 813 trials. Without intermediate rewards and using transfer learning, the agent was able to stop a penalty with an accuracy of $\sim$65% after an average of 7476 trials.

Fig. 1. Increasing difficulty in the tasks; allowing a wider shooting range to the penalty kicker.



Conclusion In the experiments, two different agent were trained, one that used transfer learning, and one that only uses the neural network function approximator. The results showed that the agent that was trained using intermediate rewards was able to reliably catch the ball with the penalty kicker at full capacity. The intermediate reward was based on the distance of the goal keeper to the expected trajectory of the ball, which is a feature containing a lot of domain knowledge. In the case of the agent that was trained using transfer learning, the agent was able to progress through all the training stages, and learn to implicitly estimate the most likely trajectory from the position of the penalty kicker behind the ball. The implicit estimate was not perfect: the goalkeeper developed a 'weak' side. Video's of the learned behavior are publicly available[1]. For future work, experiments can be conducted on the size and shape of the neural network used, as this seemed to have quite a big influence on the performance of the algorithm.

References

1. Gerndt, R., Seifert, D., Baltes, J.H., Sadeghnejad, S., Behnke, S.: Humanoid robots in soccer: Robots versus humans in robocup 2050. IEEE Robot. Automat. Mag. **22**(3), 147–154 (2015)
2. Hausknecht, M., Stone, P.: Deep reinforcement learning in parameterized action space. arXiv preprint arXiv:1511.04143 (2015)
3. Lagrand, C.G.: Learning a robot to score a penalty: Minimal reward reinforcement learning. Bachelor's thesis, Universiteit van Amsterdam (July 2017)
4. van Heusden, R.J.: Making a robot stop a penalty using q learning and transfer learning. Bachelor's thesis, Universiteit van Amsterdam (June 2018)
5. Watkinson, W., Camp, T.: Training a robocup striker agent via transferred reinforcement learning. In: RoboCup 2018: Robot Soccer World Cup XXII. Lecture Notes in Artificial Intelligence, Springer (2019)

[1] https://www.youtube.com/watch?v=MlHNmP-_R3g