Accelerating Multi-agent Reinforcement Learning with Dynamic Co-learning

Steve Homer[1], Fabian Perez[1], Quinten Rosseel[1] and Matthias Humt[1]

[1]{steven.homer, fabian.perez, quinten.rosseel, matthias.humt}@vub.be

Abstract

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Introduction

In the introduction a clear specification of the problem is provided, including related work. Look for additional but relevant references which you can cite in this part.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Methods

In the Methods section you specify how the model is constructed.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Results and Discussion

In the Results section you provide a descriptions of the simulations you have done and their results. Also include the information concerning the parameter settings.
In the Discussion section you provide a summary and an explanation of your work.

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

Conclusion

Related work. All articles cited in the original paper.

(Carroll and Seppi, 2005), (Garant et al., 2015),
(Ghavamzadeh et al., 2006), (Gmytrasiewicz and Doshi, 2005),
(Guestrin et al., 2002), (Kitano et al., 1999),
(Lazaric et al., 2008), (Littman, 2001),
(Nair et al., 2005), (Nedic and Ozdaglar, 2009),
(Oliehoek et al., 2008), (Price and Boutilier, 2003),
(Rényi et al., 1961), (Taylor and Stone, 2009),
(Vickrey and Koller, 2002), (Witwicki and Durfee, 2010),
(Zhang et al., 2010), (Zhang and Lesser, 2013)

References

Carroll, J. L. and Seppi, K. (2005). Task similarity measures for transfer in reinforcement learning task libraries. In *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*, volume 2, pages 803–808. IEEE.

Garant, D., da Silva, B. C., Lesser, V., and Zhang, C. (2015). Accelerating multi-agent reinforcement learning with dynamic co-learning. Technical report, Technical report.

Ghavamzadeh, M., Mahadevan, S., and Makar, R. (2006). Hierarchical multi-agent reinforcement learning. *Autonomous Agents and Multi-Agent Systems*, 13(2):197–229.

Gmytrasiewicz, P. J. and Doshi, P. (2005). A framework for sequential planning in multi-agent settings. *J. Artif. Intell. Res.(JAIR)*, 24:49–79.

Guestrin, C., Koller, D., and Parr, R. (2002). Multiagent planning with factored mdps. In *Advances in neural information processing systems*, pages 1523–1530.

Kitano, H., Tadokoro, S., Noda, I., Matsubara, H., Takahashi, T., Shinjou, A., and Shimada, S. (1999). Robocup rescue: Search and rescue in large-scale disasters as a domain for autonomous agents research. In *Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on*, volume 6, pages 739–743. IEEE.

Lazaric, A., Restelli, M., and Bonarini, A. (2008). Transfer of samples in batch reinforcement learning. In *Proceedings of the 25th international conference on Machine learning*, pages 544–551. ACM.

Littman, M. L. (2001). Value-function reinforcement learning in markov games. *Cognitive Systems Research*, 2(1):55–66.

Nair, R., Varakantham, P., Tambe, M., and Yokoo, M. (2005). Networked distributed pomdps: A synthesis of distributed constraint optimization and pomdps. In *AAAI*, volume 5, pages 133–139.

Nedic, A. and Ozdaglar, A. (2009). Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1):48–61.

Oliehoek, F. A., Spaan, M. T., Whiteson, S., and Vlassis, N. (2008). Exploiting locality of interaction in factored dec-pomdps. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1*, pages 517–524. International Foundation for Autonomous Agents and Multiagent Systems.

Price, B. and Boutilier, C. (2003). Accelerating reinforcement learning through implicit imitation. *Journal of Artificial Intelligence Research*, 19:569–629.

Rényi, A. et al. (1961). On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. The Regents of the University of California.

Taylor, M. E. and Stone, P. (2009). Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, 10(Jul):1633–1685.

Vickrey, D. and Koller, D. (2002). Multi-agent algorithms for solving graphical games. In *AAAI/IAAI*, pages 345–351.

Witwicki, S. J. and Durfee, E. H. (2010). Influence-based policy abstraction for weakly-coupled dec-pomdps. In *ICAPS*, pages 185–192.

Zhang, C. and Lesser, V. (2013). Coordinating multi-agent reinforcement learning with limited communication. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 1101–1108. International Foundation for Autonomous Agents and Multiagent Systems.

Zhang, C., Lesser, V., and Abdallah, S. (2010). Self-organization for coordinating decentralized reinforcement learning. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 739–746. International Foundation for Autonomous Agents and Multiagent Systems.