Journal of Artificial Intelligence Research 1 (2018) 1-15 Submitted 8/17; published 12/17

Reinforcement Learning for Natural Language Processing

Víctor Uc-Cetina UCCETINA@CORREO.UADY.MX
Computational Learning and Imaging Research Laboratory, Facultad de Matemáticas, Universidad Autónoma de Yucatán, Anillo Periférico Norte, C.P. 97119, Mérida, Mexico

Nicolás Navarro-Guerrero NAVARRO@INFORMATIK.UNI-HAMBURG.DE
Knowledge Technology Group, University of Hamburg, Department of Computer Science, Vogt-Koelln-Str. 30, 22527 Hamburg, Germany

Anabel Martin-Gonzalez AMARTING@CORREO.UADY.MX
Computational Learning and Imaging Research Laboratory, Facultad de Matemáticas, Universidad Autónoma de Yucatán, Anillo Periférico Norte, C.P. 97119, Mérida, Mexico

Xavier Hinaut XAVIER.HINAUT@INRIA.FR
Inria Bordeaux, Sud-Ouest Research Centre, 200, avenue de la Vieille Tour 33400, France

Stefan Wermter WERMTER@INFORMATIK.UNI-HAMBURG.DE
Knowledge Technology Group, University of Hamburg, Department of Computer Science, Vogt-Koelln-Str. 30, 22527 Hamburg, Germany

Abstract

Reinforcement learning algorithms do not seem to be the kind of tool that we would associate to natural language processing tasks, at least, not in a first thought. However, in the last decades many researchers have explored the use of reinforcement learning as one main component in the solution of specific natural language tasks. This paper review the state of the art of reinforcement learning methods applied to solve different problems of current interest in natural language processing. We analyzed why for certain problems reinforcement learning algorithms have provided successful solutions and for which others they are not appropriate. Finally, we point out which other problems from natural language processing might benefit from reinforcement learning in the near future.

1. Introduction

2. Analysis of Reinforcement Learning and Natural Language Processing

We analyze different natural language processing tasks and the reinforcement learning algorithms used to solve them.

3. Context and Impact

4. Research Gaps Needed to be Addressed

5. New Research Directions

6. Conclusions

(Levin, Pieraccini, & Eckert, 1997) (Singh, Kearns, Litman, & Walker, 1999) (Walker, 2000) (Oh & Rudnicky, 2000) (Singh, Kearns, Litman, & Walker, 2000) (Levin, Pieraccini,

& Eckert, 2000) (Ratnaparkhi, 2002) (Singh, Litman, Kearns, & Walker, 2002) (Schatzmann, Weilhammer, Stuttle, & Young, 2006) (Young, Gašić, Kaizer, Mairesse, Schatzmann, Thomson, & Yu, 2010) (Vogel & Jurafsky, 2010) (Dethlefs & Cuayáhuitl, 2011b) (Dethlefs & Cuayáhuitl, 2011a) (Branavan, Silver, & Barzilay, 2012) (Banchs & Li, 2012) (Gašić, Breslin, Henderson, Kim, Szummer, Thomson, Tsiakoulis, & Young, 2013a) (Gašić, Breslin, Henderson, Kim, Szummer, Thomson, Tsiakoulis, & Young, 2013b) (Young, Gašić, Thomson, & Williams, 2013) (Nio, Sakti, Neubig, Toda, Adriani, & Nakamura, 2014) (Gašić, Kim, Tsiakoulis, Breslin, Henderson, Szummer, Thomson, & Young, 2014) (Guo, 2015) (Narasimhan, Kulkarni, & Barzilay, 2015) (Li, Monroe, Ritter, Galley, Gao, & Jurafsky, 2016) (He, Chen, He, Gao, Li, Deng, & Ostendorf, 2016)

References

Banchs, R. E., & Li, H. (2012). Iris: a chat-oriented dialogue system based on the vector space model. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*.

Branavan, S. R. K., Silver, D., & Barzilay, R. (2012). Learning to win by reading manuals in a monte-carlo framework. *Journal of Artificial Intelligence Research*, *43*, 661–704.

Dethlefs, N., & Cuayáhuitl, H. (2011a). Combining hierarchical reinforcement learning and bayesian networks for natural language generation in situated dialogue. In *Proceedings of the 13th European Workshop on Natural Language Generation*.

Dethlefs, N., & Cuayáhuitl, H. (2011b). Hierarchical reinforcement learning and hidden markov models for task-oriented natural language generation. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*.

Gašić, M., Breslin, C., Henderson, M., Kim, D., Szummer, M., Thomson, B., Tsiakoulis, P., & Young, S. (2013a). On-line policy optimisation of bayesian spoken dialogue systems via human interaction. In *Proceedings of the IEEE ICASSP*.

Gašić, M., Breslin, C., Henderson, M., Kim, D., Szummer, M., Thomson, B., Tsiakoulis, P., & Young, S. (2013b). Pomdp-based dialogue manager adaptation to extended domains. In *Proceedings of the SIGDIAL Conference*.

Gašić, M., Kim, D., Tsiakoulis, P., Breslin, C., Henderson, M., Szummer, M., Thomson, B., & Young, S. (2014). Incremental on-line adaptation of pomdp-based dialogue managers to extended domains. In *Proceedings of the InterSpeech*.

Guo, H. (2015). Generating text with deep reinforcement learning. In *Proceedings of the NIPS Conference*.

He, J., Chen, J., He, X., Gao, J., Li, L., Deng, L., & Ostendorf, M. (2016). Deep reinforcement learning with a natural language action space. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*.

Levin, E., Pieraccini, R., & Eckert, W. (1997). Learning dialogue strategies within the markov decision process framework. In *Proceedings of IEEE Workshop on Automatic Speech Recognition and Understanding*.

Levin, E., Pieraccini, R., & Eckert, W. (2000). A stochastic model of human-machine interaction for learning dialog strategies. *IEEE Transactions on Speech and Audio Processing*, *8*(1), 11–23.

Li, J., Monroe, W., Ritter, A., Galley, M., Gao, J., & Jurafsky, D. (2016). Deep reinforcement learning for dialogue generation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

Narasimhan, K., Kulkarni, T. D., & Barzilay, R. (2015). Language understanding for text-based games using deep reinforcement learning. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*.

Nio, L., Sakti, S., Neubig, G., Toda, T., Adriani, M., & Nakamura, S. (2014). Developing non-goal dialog system based on examples of drama television. In *Proceedings of the Natural Interaction with Robots, Knowbots and Smartphones Conference*.

Oh, A. H., & Rudnicky, A. I. (2000). Stochastic language generation for spoken dialogue systems. In *Proceedings of the 2000 ANLP/NAACL Workshop on Conversational Systems*.

Ratnaparkhi, A. (2002). Trainable approaches to surface natural language generation and their application to conversational dialog systems. *Computer Speech and Language*, *16*, 435–455.

Schatzmann, J., Weilhammer, K., Stuttle, M., & Young, S. (2006). A survey of statistical user simulation techniques for reinforcement-learning of dialogue management strategies. *The Knowledge Engineering Review*, *21*(02), 97–126.

Singh, S., Kearns, M., Litman, D., & Walker, M. (1999). Reinforcement learning for spoken dialogue systems. In *Proceedings of NIPS Conference*.

Singh, S., Kearns, M., Litman, D. J., & Walker, M. A. (2000). Empirical evaluation of a reinforcement learning spoken dialogue system. In *Proceedings of the 17th National Conference on Artificial Intelligence*.

Singh, S., Litman, D., Kearns, M., & Walker, M. (2002). Optimizing dialogue management with reinforcement learning: experiments with the njfun system. *Journal of Artificial Intelligence Research*, *16*, 105–133.

Vogel, A., & Jurafsky, D. (2010). Learning to follow navigational directions. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*.

Walker, M. A. (2000). An application of reinforcement learning to dialogue strategy selection in a spoken dialogue system for email. *Journal of Artificial Intelligence Research*, *12*, 387–416.

Young, S., Gašić, M., Kaizer, S., Mairesse, F., Schatzmann, J., Thomson, B., & Yu, K. (2010). The hidden state model: a practical framework for pomdp-based spoken dialogue management. *Computer Speech and Language*, *24*, 150–174.

Young, S., Gašić, M., Thomson, B., & Williams, J. D. (2013). Pomdp-based statistical spoken dialogue systems: a review. In *Proceedings of the IEEE*.