LOA: Logical Optimal Actions for Text-based Interaction Games

Daiki Kimura* Subhajit Chaudhury*   Masaki Ono   Michiaki Tatsubori
Don Joven Agravante   Asim Munawar   Akifumi Wachi   Ryosuke Kohita   Alexander Gray
IBM Research AI

{daiki, subhajit, moono, mich}@jp.ibm.com,
{don.joven.r.agravante, asim, akifumi.wachi, kohi, alexander.gray}@ibm.com

Abstract

We present Logical Optimal Actions (LOA), an action decision architecture of reinforcement learning applications with a neuro-symbolic framework which is a combination of neural network and symbolic knowledge acquisition approach for natural language interaction games. The demonstration for LOA experiments consists of a web-based interactive platform for text-based games and visualization for acquired knowledge for improving interpretability for trained rules. This demonstration also provides a comparison module with other neuro-symbolic approaches as well as non-symbolic state-of-the-art agent models on the same text-based games. Our LOA also provides open-sourced implementation in Python for the reinforcement learning environment to facilitate an experiment for studying neuro-symbolic agents. Demo site: https://ibm.biz/acl21-loa, Code: https://github.com/ibm/loa

1 Introduction

Neuro-symbolic (NS) hybrid approaches have been proposed for overcoming the weakness of deep reinforcement learning (Dong et al., 2019; Jiang and Luo, 2019; Kimura, 2018; Kimura et al., 2018), including less training data with generalization, external knowledge utilization, and direct explainability of what is learned. Study of reinforcement learning (RL) in non-symbolic environments, such as those with natural language and visionary observations, would be an important step towards the real-world application of the approaches beyond classic and symbolic environments.

Under certain controls necessary for studying RL, text-based games provide complex, interactive, and a variety of simulated environments where the environmental game state observation
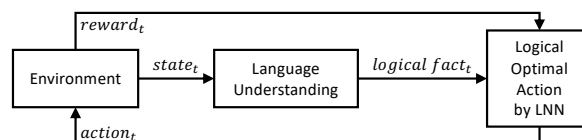
---
* denotes equal contribution

Figure 1: An architecture overview for LOA.

is obtained through the text description, and the agent is expected to make progress by entering text commands. In addition to language understanding (Ammanabrolu and Riedl, 2019; Adhikari et al., 2020), successful play requires skills such as long-term memory (Narasimhan et al., 2015), exploration (Yuan et al., 2018), observation pruning (Chaudhury et al., 2020), and common sense reasoning (Keerthiram Murugesan and Campbell, 2021). However, these studies are not using the neuro-symbolic approach which is a combination of the neural network and the symbolic framework.

A recent neuro-symbolic framework called the Logical Neural Networks (LNN) (Riegel et al., 2020) simultaneously provides key properties of both neural networks (learning) and symbolic logic (reasoning). The LNN can train the constraints and rules with logical functions in the neural networks, and since every neuron in the network has a component for a formula of weighted real-valued logics, it can calculate the probability and contradiction loss for each of the propositions. At the same time, trained LNN follow symbolic rules, which means they yield a highly interpretable disentangled representation. Using this benefit of LNN, we proposed a neuro-symbolic RL method that uses pre-defined external knowledge in logical networks, and the method successfully plays on the text-based games (Kimura et al., 2021).

In this demonstration (demo site: https://ibm.biz/acl21-loa), we present a Logical Optimal Actions (LOA) architecture for neuro-symbolic RL applications with LNN (Riegel

227

et al., 2020) for text-based interaction games. While natural language-based interactive agents are the ambitious but attractive target as real-world applications of neuro-symbolic, it is not easy to provide an environment for the agent. The proposed demonstration uses text-based games learning environment, called TextWorld (Côté et al., 2018), as a miniature of a natural language-based interactive environment. The demonstration provides a web-based user interface for visualizing the game interaction, which is including displaying the natural text observation from the environment, typing the action sentence, and showing the reward value from the taken action. The LOA in this demonstration also visualizes trained and pre-defined logical rules in LNN via the same interface, and this will help the human user understand the benefits of introducing the logical rules via neuro-symbolic frameworks. We also supply an open-sourced implementation for demo environment and some RL methods. This implementation contains our logical approaches and other state-of-the-art agents.

2 Logical Optimal Action

Our proposing LOA is an RL framework which is combining logical reasoning and neural network training. These training and reasoning are provided from functionalities of LNN (Riegel et al., 2020) that is simultaneously providing key properties of both neural networks and symbolic logic. Figure 1 shows the overview architecture for LOA. The LOA model receives the logical state value as logical fact from the language understanding component which receives raw natural language state value from the environment. The model forwards into LNN for the input to get the optimal action for it, the action goes into the environment to execute the action command, then reward is input to LOA agent. The LOA will be trained the action decision network in LNN by using the acquired reward value and chosen action from the network.

3 LOA Demo

The proposing web-based LOA demonstration supports two functionalities: 1) play the text-based game by human interactions, 2) visualize the trained and pre-defined LNN to increase interpretability for acquired rules.

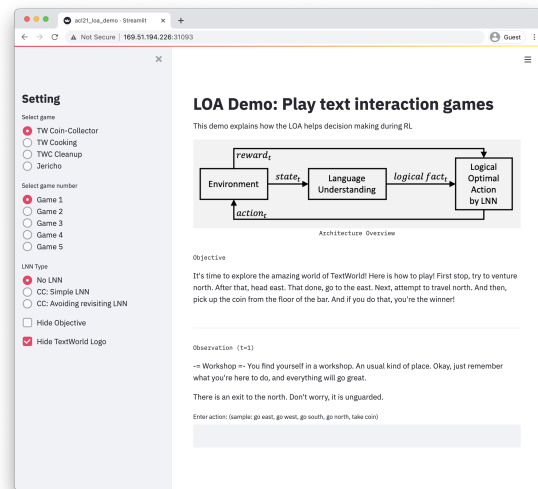For playing the games by web interface, Fig. 2 shows an initial view for the LOA demonstration.
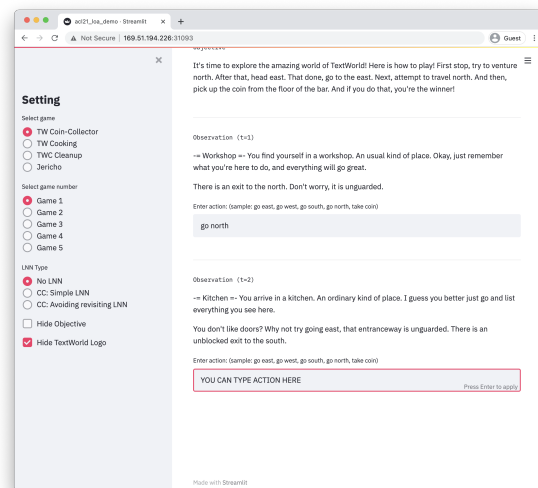

Figure 2: Initial view for LOA demo.

Figure 3: View for playing the game.

On the left-hand side, we can choose the game from some existing text-based interaction games [1], such as TextWorld Coin-Collector game (Côté et al., 2018), TextWorld Cooking game (Côté et al., 2018), TextWorld Commonsense Cleanup game (Keerthiram Murugesan and Campbell, 2021), and Jericho game (Hausknecht et al., 2019). Figure 3 shows the view for playing the TextWorld game, and Fig. 4 shows the view for another game (cleanup task). The human player can input any action by natural language then the demonstration system displays the raw observation output from the environment.

For visualizing the trained and pre-defined neuro-symbolic network in LNN, Fig. 5 and Fig. 6 show the example of the LNN output. In these

---

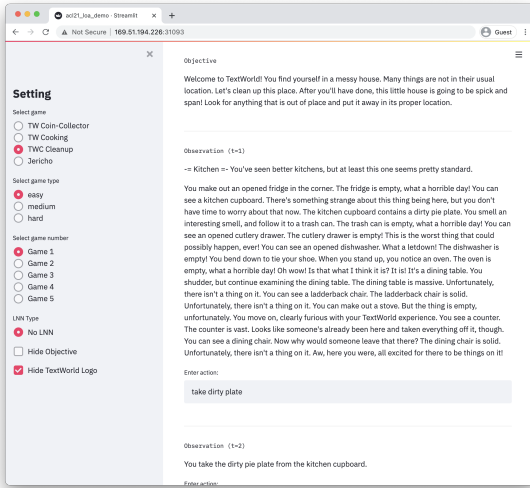[1]We are planning to add other games.
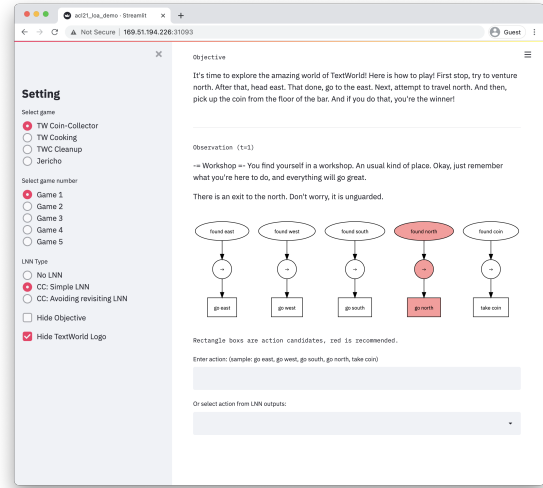
Figure 4: View for playing the cleanup game.

Figure 5: Displaying the simple LNN with given state.

figures, the LNN contains simple rules for the TextWorld Coin-Collector game; for example, the rule is the agent takes 'go east' action, when the agent finds the east room ("`found west`" → "`go west`"). The round box explains the proposition from the given observation inputs, the circle with a logical function means a logical function node of LNN, and the rectangle box explains an action candidate for the agent. The highlighted nodes (red node) have 'true' value, and non-highlighted nodes (white node) have 'false' value. In Fig. 5, the agent found the north exit from the given observation ("`Observation (t=1)`") by using semantic parser [2], then the going north room action ("`go north`") are activated. In Fig. 6, if the user clicks the selectable box, the LOA recommends only one action which is 'go north'. In this demonstration, we show the benefit of introducing the LNN into an RL agent, we don't prepare to automatically choose the action by LOA framework. However, if we execute the RL with LOA framework, the RL agent can converge faster than other non-symbolic and neuro-symbolic methods.

After selecting "`go north`" action at $t = 1$, next observation sentence and LNN output for next step are shown in Fig 7. In this step, the agent found two doors, which are east and south; however, the south door is connected to the previous room because the agent took going north action at the previous step. Since this LNN is simple LNN, the "`go south`" action is also recommended in

---

[2]This parser is out of our current research topic, we prepare a simple semantic parser.
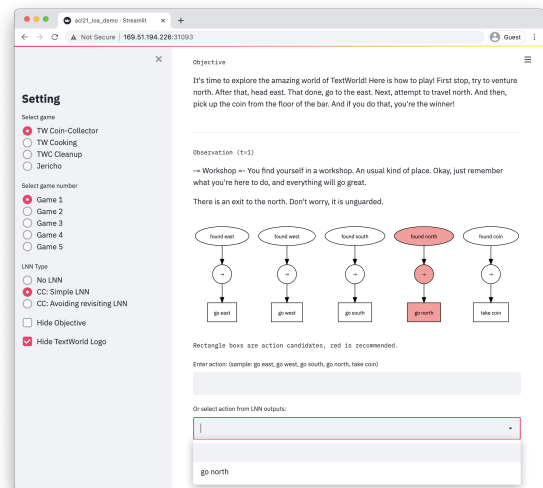

Figure 6: User can choose the recommended action.

Fig 7. Figure 8 shows the output of the complicated LNN which has functionality for avoiding revisiting the visited room. By using such the LNN, LOA can output only "`go east`" action by having contradiction loss in LNN. This is a benefit of introducing the neuro-symbolic framework, and the human user can easily understand the reason for the taken action by the agent with this interpretability by LOA.

4 Conclusion

We propose a novel demonstration (URL: https://ibm.biz/acl21-loa) which provides to play the text-based games on the web interface and visualize the benefit of the neuro-symbolic algorithm. This application helps the human user understand
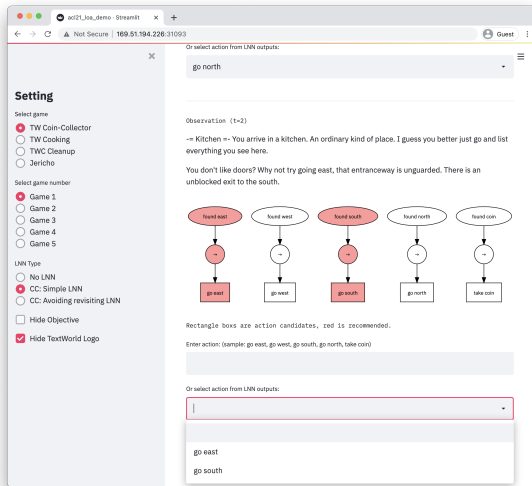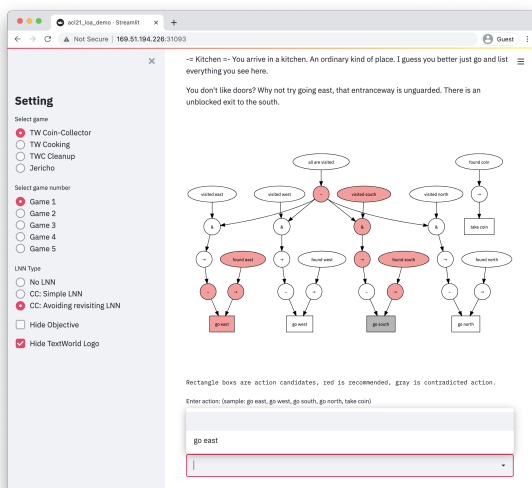
Figure 7: Result for simple LNN.

Figure 8: Result for avoiding revisiting LNN.

the trained network and the reason for taken action by the agent. We also extend more complicated LNN for other difficult games on the demo site. At the same time, we open the source code for the demonstration (URL: https://github.com/ibm/loa).

References

Ashutosh Adhikari, Xingdi Yuan, Marc-Alexandre Côté, Mikuláŝ Zelinka, Marc-Antoine Rondeau, Romain Laroche, Pascal Poupart, Jian Tang, Adam Trischler, and William L. Hamilton. 2020. Learning dynamic belief graphs to generalize on text-based games. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*.

Prithviraj Ammanabrolu and Mark Riedl. 2019. Playing text-adventure games with graph-based deep reinforcement learning. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 3557–3565.

Subhajit Chaudhury, Daiki Kimura, Kartik Talamadupula, Michiaki Tatsubori, Asim Munawar, and Ryuki Tachibana. 2020. Bootstrapped q-learning with context relevant observation pruning to generalize in text-based games. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, pages 3002–3008.

Marc-Alexandre Côté, Ákos Kádár, Xingdi Yuan, Ben Kybartas, Tavian Barnes, Emery Fine, James Moore, Matthew J. Hausknecht, Layla El Asri, Mahmoud Adada, Wendy Tay, and Adam Trischler. 2018. Textworld: A learning environment for text-based games. In *Computer Games - 7th Workshop, CGW 2018, Held in Conjunction with the 27th International Conference on Artificial Intelligence, IJCAI 2018, Stockholm, Sweden, July 13, 2018, Revised Selected Papers*, pages 41–75.

Honghua Dong, Jiayuan Mao, Tian Lin, Chong Wang, Lihong Li, and Denny Zhou. 2019. Neural logic machines. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*.

Matthew Hausknecht, Prithviraj Ammanabrolu, Côté Marc-Alexandre, and Yuan Xingdi. 2019. Interactive fiction games: A colossal adventure. *CoRR*, abs/1909.05398.

Zhengyao Jiang and Shan Luo. 2019. Neural logic reinforcement learning. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, pages 3110–3119.

Pavan Kapanipathi Pushkar Shukla Sadhana Kumaravel Gerald Tesauro Kartik Talamadupula Mrinmaya Sachan Keerthiram Murugesan, Mattia Atzeni and Murray Campbell. 2021. Text-based RL Agents with Commonsense Knowledge: New Challenges, Environments and Baselines. In *Thirty Fifth AAAI Conference on Artificial Intelligence*.

Daiki Kimura. 2018. Daqn: Deep auto-encoder and q-network. *arXiv preprint arXiv:1806.00630*.

Daiki Kimura, Subhajit Chaudhury, Ryuki Tachibana, and Sakyasingha Dasgupta. 2018. Internal model from observations for reward shaping.

Daiki Kimura, Subhajit Chaudhury, Akifumi Wachi, Ryosuke Kohita, Asim Munawar, Michiaki Tatsubori, and Alexander Gray. 2021. Reinforcement learning with external knowledge by using logical neural networks. *arXiv preprint arXiv:2103.02363*.

Karthik Narasimhan, Tejas D. Kulkarni, and Regina Barzilay. 2015. Language understanding for text-based games using deep reinforcement learning. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015, Lisbon, Portugal, September 17-21, 2015*, pages 1–11.

Ryan Riegel, Alexander G. Gray, Francois P. S. Luus, Naweed Khan, Ndivhuwo Makondo, Ismail Yunus Akhalwaya, Haifeng Qian, Ronald Fagin, Francisco Barahona, Udit Sharma, Shajith Ikbal, Hima Karanam, Sumit Neelam, Ankita Likhyani, and Santosh K. Srivastava. 2020. Logical neural networks. *CoRR*, abs/2006.13155.

Xingdi Yuan, Marc-Alexandre Côté, Alessandro Sordoni, Romain Laroche, Remi Tachet des Combes, Matthew J. Hausknecht, and Adam Trischler. 2018. Counting to explore and generalize in text-based games. *CoRR*, abs/1806.11525.