Jiayi Weng

CST Department, Tsinghua University

February 29, 2020

Jiayi Weng

- B.Eng. in Computer Science, Tsinghua University, 2016 - 2020.
  GPA 3.75/4.00 (Rank 17/158)

- Research Assistant to Prof. Yoshua Bengio, Mila
  Jul. 2019 - now, topic: NLP, RL, Neural-Symbolic Reasoning

- Research Assistant to Prof. Jun Zhu, Tsinghua University
  Mar. 2018 - now, topic: RL

- Research Assistant to Dr. Hongwei Qin, Sensetime Inc.
  Mar. 2018 - Jun. 2018, topic: CV (HDR+, denoising)

- Research Assistant to Prof. Hailong Yao, Tsinghua University
  May 2017 - Oct. 2017, topic: VLSI (escape routing)

- GitHub 1000+ personal stars, 600+ followers

Bio
○○

Research Experience (RL)
○○●○○○○○○○○○

Future Direction
○○○

VizDoom AI Competition 2018

(a) Navigation

(b) Resources: weapon
and health bonus

(c) Shoot enemies

(d) Final goal:
open exit door

■ birth place
▪▪▪ navigation road
■ exit door

Figure 1: Example of single-player scenarios in VizDoom.

Bio
○○
Research Experience (RL)
○○○●○○○○○○○○
Future Direction
○○○

VizDoom AI Competition 2018
Challenges & Solutions

1 Visual understanding from raw images
  ⇒ Set up an object detection system (YOLO-v3[1]) aiming to efficiently identify the obstacles and enemies.

2 Visual navigation in the 3D environment
  ⇒ Incorporated the depth signal into the navigation network, which facilitated the agent to plot its paths. (+SNAIL[2])

3 Sparse and delayed reward for the agent
  ⇒ Formulated the problem within a hierarchical reinforcement learning framework by dividing the challenge into subtasks: navigation, attack, tool, and resources.

---

[1]Joseph Redmon et al. "YOLOv3: An Incremental Improvement". In: *CoRR* abs/1804.02767 (2018). arXiv: 1804.02767.

[2]Nikhil Mishra et al. "A Simple Neural Attentive Meta-Learner". In: *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. 2018.

Bio
○○

VizDoom AI Competition 2018

Research Experience (RL)
○○○○●○○○○○○○

Future Direction
○○○

VizDoom Competition Result[3]

- All teams submitted an agent to finish the game on ten unseen and more difficult maps.
- Our method achieved first place in both public-rank and private-rank leaderboards.

Team	Ours	DoomNet	VIPLAB	ddangelo
Total Time (min)	**25.34**	29.86	31.54	37.33
# of Best Record	**8**	4	3	4

Table 1: VDAIC 2018 Track(1) Competition Result (top 4 team).

---

[3]Detailed result can be found at
http://vizdoom.cs.put.edu.pl/competitions/vdaic-2018-cig/results.

Bio
○○
Relational Markov Decision Process

Research Experience (RL)
○○○○○●○○○○○○

Future Direction
○○○

1 Bio

2 Research Experience (RL)
  VizDoom AI Competition 2018
  Relational Markov Decision Process
  Rule-Transformer

3 Future Direction

Bio
○○
Research Experience (RL)
○○○○○○○●○○○○○
Future Direction
○○○
Relational Markov Decision Process

We aim to accelerate the deep reinforcement learning procedure by state abstraction (using first-order logic) and reward shaping.
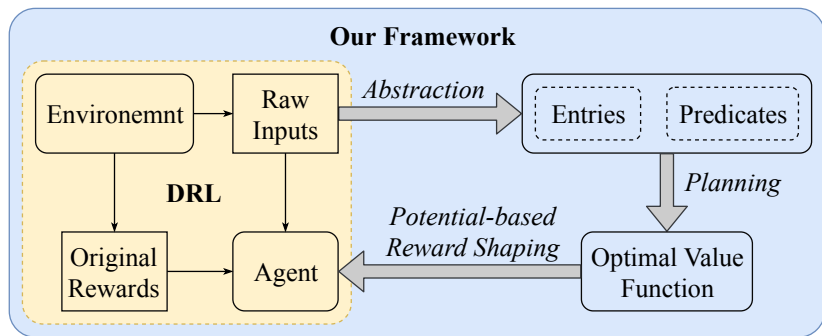

Figure 2: The working procedure of RMDP. Generate the credits for state-action pairs of the deep reinforcement learning in three stages: abstraction, planning, and potential-based reward shaping.

Bio
○○

Research Experience (RL)
○○○○○○○○●○○○○

Future Direction
○○○

Relational Markov Decision Process

- Our method can accelerate the convergence of value-based RL method.
- RMDP is more robust under harder games.
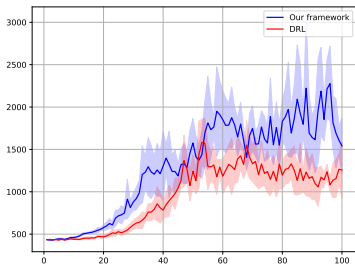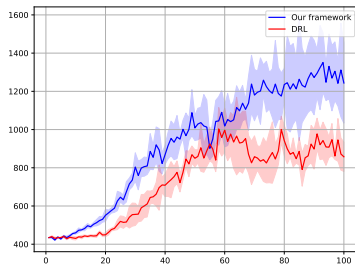
(a) Easy Map   (b) Hard Map

Figure 3: Test RMDP on Health Gathering scenarios. Vertical axis indicates the accumulated reward per episode through training phase.

Bio
○○

Research Experience (RL)
○○○○○○○○○○●○○

Future Direction
○○○

Rule-Transformer

- We aimed to develop a system with Consciousness Prior[4], which integrated the advantage of both neural and symbolic models.
- To make Transformers more like reasoning with rules, instead of a fixed rule per layer, allow to also pick a rule, at each level, and share that set across levels.
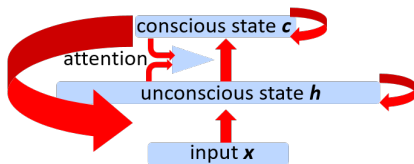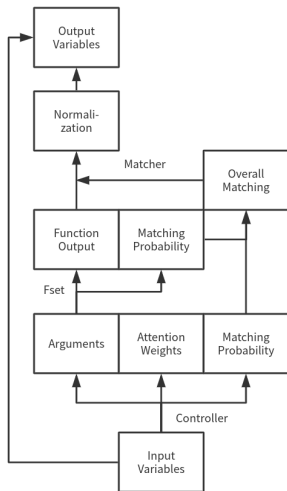

Figure 4: The Consciousness Prior, credit[5].

---

[4]Yoshua Bengio. "The Consciousness Prior". In: *CoRR* abs/1709.08568 (2017). arXiv: 1709.08568.

[5]Yoshua Bengio. *From System 1 Deep Learning to System 2 Deep Learning*. 2019.

Bio
○○
Research Experience (RL)
○○○○○○○○○○●○

Future Direction
○○○

Rule-Transformer

Figure 5: Architecture of
Rule-Transformer.

- Allow picking a rule at each level
- Sharing knowledge across the level
- Update different variables with different functions
- The rule could potentially be dynamically defined

Rule-Transformer

What we have tried and are trying

- x Machine Translation
- x Language Modeling
- = GLUE Benchmark
- = Multitask Imitation Learning
- x Mathematical Dataset
- v LeetCode-style algorithmic problems
- ? Brainfuck programming language explainer (ongoing)

1 Bio

2 Research Experience (RL)

3 Future Direction

- Generally, I would like to continue my research in the field of reinforcement learning.
- I am interested in addressing the challenges of sample complexity (incorporating human-prior knowledge for better disentanglement) / robustness and generalization (combining with meta-learning, game-theory).
- I am also keen to work on the real-world application of reinforcement learning: learning-based control system, self-driving cars, etc.

Thank you!