Research Topics of Reinforcement Learning

- Schedule: 6:30pm~9:30pm (Tue),
  - Location: EC114 (實體)
- Instructors:
  - I-Chen Wu (吳毅成), TEL: x31855
  - Email: icwu@cs.nycu.edu.tw
  - WWW: http://cgilab.nctu.edu.tw/~icwu/
- TAs: rl-topics@cgilab.nctu.edu.tw
  - Wei-Chen Liao(廖唯辰) + Yu-Cheng Chen (陳昱丞). TEL: x31603
- Course web site: http://e3.nycu.edu.tw
  - TA: Listed in this web site.
  - Slides: Given in this web site. (keep updating)

Goal

- Learn how to develop reinforcement learning (RL) algorithms in a wide range of applications
  - such as computer games, video games, intelligent traffic management, manufacturing scheduling, autonomous driving/racing, robotics, etc.
- Objectives:
  - To understand the basic core concepts of reinforcement learning (RL)
  - To understand many latest RL techniques for applications
  - To familiarize with tools for developing RL, such as PyTorch, etc.
  - To develop practical working systems via projects close to DeepRacer.
- Disclaimers:
  - MAY NOT go through all theorems.
  - Schedules may not be precise, since this is the first time I lecture RL research topics.

Prerequisite

- Better to have basic ideas about
  - Algorithm, Artificial Intelligence, Probability, Linear Algebra, Calculus.
- Prerequisite: (suggested)
  - Machine Learning, Deep Learning

Evaluation

- Homework: may or may not be given. Not graded.
- Program projects (~50%)
  - Design AI programs for designated games. (See the next slide.)
- Final Exam (~30%)
  - Mainly review the concepts learned from the course. (Close book!)
- Presentation (~20%) About 3-4 weeks near the end.
  - 1-3 persons per group, depending on student number.
  - List of papers to present is given in Week 8-10 and finalized in Week 11.
    - If you want to present other papers, welcome to discuss with us.
  - Time: 12-minute talk with <= 30 slides in PDF or PPT.
    - 8-minute Q & A: 4 min from 2 teams and 4 min from lecturers/TAs.
  - Submit your revised set of slides one week later.
- Note: The above percentages are subject to changes, based on the condition in the course.

Program Projects

- 4 Labs:
  - TD Learning (~week 2)
    - 2048
  - DQN (~week 4)
    - Atari games: Enduro
    - Discrete action space environment
  - PPO (~week 6)
    - Atari games: Enduro
    - Discrete action space environment
  - TD3 (~week 8)
    - Car racing gym
    - Continuous action space environment

  For each project, 12 days are allowed to implement, and normally demo on Monday.
- 1 final project:
  - Racecar Gym (~week 9).
  - Compete or demo in week 17.
- Note: The course needs a significant amount of computing resources.
  - YOU MUST PREPARE YOUR OWN MACHINE. (better above GTX 1080)

加選

- https://forms.gle/qMMws7n6dLQJX6pS9

Course Outlines (I)

Introduction to RL: (~2 weeks)

- Syllabus: (1hr.)
- Introduction to Reinforcement Learning (2 hrs.)
- Case studies: Lightweight Model (2 hrs.)
  - 2048 (Temporal Difference Learning)
  - Go Programs (with Monte-Carlo Tree Search)

  << Lab: 2048>> (1 hr.)

Core of RL: (4-5 weeks)

- Fundamentals of RL (4-5 hrs.)
  - Markov Decision Process (MDP)
  - Dynamic Programming (Tabular RL)
  - Model Free RL: TD(λ), eligibility.
  - ϵ-Greedy Exploration, Q-Learning
  - Function Approximation

Course Outlines (II)

Core of RL: (cont.)

- Value-Based Reinforcement Learning (3 hrs.)
  - DQN, DDQN (Double DQN), DRQN
  - Dueling Network (with Advantage)
  - Distributional DQN: C51, QR-DQN, IQN*, FQF*

  <<Lab: DQN>> (1 hr.)
- Policy-based Reinforcement Learning (5-6 hrs.)
  - Policy Gradient
  - Actor-Critic (Discrete actions)
  - A2C and A3C (Asynchronous Advantage Actor-Critic)
  - TRPO & PPO, GAE, Impala*, v-trace*

  <<Lab: PPO>> (1 hr.)
  - DDPG, TD3 (Continuous Actions), D4PG*
  - SAC (Soft Actor-Critic)

  <<Lab: TD3>> (1 hr.)

Course Outlines (III)

Advanced Topics of RL: (6-7 weeks)

- Planning: (4-6 hrs.)
  - Multi-Arm Bandits, UCB, Sequential Halving,
  - Monte-Carlo Tree Search (MCTS), AlphaGo, AlphaZero, MuZero,
  - Path Consistency, Abstraction*, EfficientZero*, Gumbel MuZero*, Stochastic MuZero*.
- Applications: (3 hrs.)
  - DeepRacer: Augmentation, RL-cycleGAN, DrQ.

  <<Lab: SAC/TD3>> (1 hr.)

  <<Final Project: TBA>> (1 hr.)
  - Solving Rubik Cube*, RL for optimization* (JSP/TSP),
  - AlphaStar*, AlphaTensor*, RLHF*, ITM*

(*: could be topics for student presentation)

Course Outlines (IV)

Advanced Topics of RL: (cont.)

- Exploration vs. Exploitation (3 hrs.)
  - ICM, RND, Never Give Up, Agent57, Go-Explore*.
- Model-based RL from demonstration (3-4 hrs.)
  - PER, Ape-X, R2D2*
  - DQfD, R2D3, Dreamer*
  - Behavior Cloning*, inverse RL*, GAIL*, Implicit Q-learning*, Offline RL*, IQ Learn*
- Multi-Agents RL (MARL) (6-8 hrs.)
  - Introduction: CTDE
  - Value-based MARL: VDN, QMIX.
  - Policy-based MARL: COMA, MAA2C, MAPPO, HATRPO/HAPPO, MAT

週次	上課日期	課程進度、內容、主題
1	9/12	Introduction to Reinforcement Learning
2	9/19	Case studies of lightweight model applications: 2048 and Go (Lab about 2048)
3	9/26	Fundamentals: Markov Decision Process (MDP), Dynamic Programming (Tabular RL), ϵ-Greedy Exploration, Q-Learning, Function Approximation.
4	10/3	Value-Based Reinforcement Learning: DQN, DDQN, Dueling Network, Distributional DQN. (Lab about DQN)
5	10/10	
6	10/17	Policy-based Reinforcement Learning: Policy Gradient, Actor-Critic (Discrete actions), A2C and A3C (Asynchronous Advantage Actor-Critic)
7	10/24	Policy-based Reinforcement Learning: TRPO & PPO, GAE, DDPG, TD3 (Continuous Actions), SAC (Soft Actor-Critic) (Lab about PPO)

8	10/31	Exploration vs. Exploitation: Multi-Arm Bandits, UCB, Sequential Halving, Planning: Dyna, Monte-Carlo Tree Search (MCTS), AlphaGo, AlphaZero, MuZero, Path Consistency, Abstraction
9	11/7	Advanced Exploration: ICM, RND; Experience Reply: PER, Ape-X (Lab about TD3)
10	11/14	Applications: DeepRacer: Augmentation, RL-cycleGAN, DrQ. (Final Project, …)
11	11/21	RL from demonstration: DQfD, R2D3; Multi-Agents RL (value-based): VDN, Q-mix,
12	11/28	Multi-Agents RL (policy-based): COMA, MAT, etc.
13	12/5	Presentation
14	12/12	Presentation
15	12/19	Presentation
16	12/26	Final exam
17	1/2	Project Demo (Final competition)

References

- Slides: http://e3.nctu.edu.tw
  - Not all released at the very beginning.
- References:
  - R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, Nov. 2017. http://webdocs.cs.ualberta.ca/~sutton/book/ebook/the-book.html
  - David Silver, Online Course for Deep Reinforcement Learning. http://www.cs.ucl.ac.uk/staff/D.Silver/web/Teaching.html
  - Papers and slides.
- Other resources: Most top conferences, such as ICLR, NeurIPS, ICML, AAAI, IJCAI, UAI, ICRA, IROS, etc.
- Acknowledgements:
  - Many slides are modified from Silver's online course (above).
  - Many slides of this course are designed and modified from many of my students,
    - mainly including the two TAs, 廖唯辰、陳昱丞, and
    - those who I marked at the beginning of slide files.
    - …

Claim

- THE SLIDES ARE USED FOR THIS COURSE ONLY, NOT IN PUBLIC. LINKS TO THESE SLIDES ARE ONLY POSTED IN THE DESIGNATED BOARD, e3.nycu.edu.tw. STUDENTS IN THIS COURSE OR SEARCH ENGINES ARE FORBIDDEN TO PUBLICIZE THESE SLIDES OR LINKS OVER THE INTERNET.