yandexdataschool / **Practical_RL**

Watch 114 Star 1,079 Fork 220

Code Issues 3 Pull requests 0 Projects 0 Wiki Insights

A course in reinforcement learning in the wild

reinforcement-learning course-materials deep-learning deep-reinforcement-learning git-course mooc theano lasagne tensorflow

342 commits **1** branch **0** releases **18** contributors MIT

Branch: master New pull request Create new file Upload files Find file Clone or download

justheuristic Merge branch 'master' of https://github.com/yandexdataschool/practica… Latest commit 857c728 Jul 5, 2017

docker	Typo fixed in run_jupyter.sh	May 5, 2017
week0	add url to post about ES	Mar 31, 2017
week1	minor monitor fix	Jul 5, 2017
week2	add readme tags	Mar 8, 2017
week3.5	alternative tensorflow assignment	Jun 12, 2017
week3	move mcts to bonus week	Apr 25, 2017
week4	add urls	Mar 31, 2017
week5	Update README.md	Apr 15, 2017
week6.5	fix seq2seq int8->int32	Apr 23, 2017
week6	add tensorflow version	Apr 8, 2017
week7	add video	Apr 24, 2017
week8	better task descriptions, get_score->get_distance	May 17, 2017
week9	fix descr	Jun 5, 2017
yet_another_week	minor formatting	May 20, 2017
.gitignore	Lectures removed	Feb 17, 2017
Amazon GPU howto.md	Update Amazon GPU howto.md	Mar 24, 2017
Dockerfile	tmp comment tf	Feb 24, 2017
LICENSE.md	add MIT license	Feb 12, 2017
README.md	clarification	Jun 12, 2017
xvfb	add lecture1	Jan 24, 2017
youtube_dl_lectures.sh	add urls	Mar 22, 2017

README.md

Practical_RL

A course on reinforcement learning in the wild. Taught on-campus in HSE and Yandex SDA (russian) and maintained to be friendly to online students (both english and russian).

Manifesto:

- **Optimize for the curious.** For all the materials that aren't covered in detail there are links to more information and related materials (D.Silver/Sutton/blogs/whatever). Assignments will have bonus sections if you want to dig deeper.

- **Practicality first.** Everything essential to solving reinforcement learning problems is worth mentioning. We won't shun away from covering tricks and heuristics. For every major idea there should be a lab that allows to “feel” it on a practical problem.
- **Git-course.** Know a way to make the course better? Noticed a typo in a formula? Found a useful link? Made the code more readable? Made a version for alternative framework? You're awesome! Pull-request it!

Coordinates and useful links

- **HSE** classes are happening on demand in Q&A mode (ping Yozhik if you want one)
- **YSDA** classes are over until next term.
- For RL reading group, ping Pasha
- Lecture slides are here.
- Online student **survival guide**
- Installing the libraries - guide and issues thread
- Magical button that creates VM: launch binder (press me. may be down time to time. If it won't load for 2-3 minutes, it's down)
- Telegram **chat room (russian)**
- Gitter **chat room (english)**
- **How to submit homeworks[HSE and YSDA only]:** anytask instructions and grading rules
- E-mail for everything else : **practicalrl17@gmail.com** (please don't submit homeworks via e-mail)
- Anonymous feedback form for everything that didn't go through e-mail.
- About the course
- A large list of RL materials - awesome rl

Announcements

- 12.06.17 - The course is over for this term. Please fill in the feedback form once you finished it. *Next term*: full tracks for tensorflow & pytorch, more ballanced assignment difficuly + whatever you vote for in the form. Meanwhile, we'll still monitor issues and pull requests at least twice a week. We're also gonna add english videolecture for week8 later this week.
- 12.06.17 - Attention @HSE students, please make sure you submit your homeworks at least 3 days prior to global term deadline for your department (even if it's coming next september).

- Previous announcements
  * 17.05.17 - !ATTENTION ysda and hse students! - there's a suspicion that anytask sometimes fails to send homework assignments. Please check that all your assignments are sent (sometimes we receive empty submissions). We will binge-check all newly sent assignments so don't worry about timing. Also this is most likely us being over suspicious, we post this warning just in case. * 1.05.17 - UPD - week8 deadlines have been prolonged till the end of holidays * 22.04.17 - YSDA deadlines for week8 set to 30th of __april__ (previously 30 may, which was a typo). * 25.03.17 - __HSE important__ next monday lecture is postponed by 1 week due to HSE mid-term exams. Deadlines have been postponed accordingly. * 25.03.17 - __week5__ you can submit any atari game you want. * 16.03.17 - __week4 homework__ max score threshold for LunarLander reduced to -100 * 16.03.17 - (hse) shifted deadline for week5 * 15.03.17 - (hse) added week6 assignment and deadline * 10.03.17 - (ysda/hse students) __important__ please consider [Course Projects](https://github.com/yandexdataschool/Practical_RL/wiki/Course-projects) as an alternative way of completing the course. * 8.03.17 - YSDA deadlines announced for weeks 3 and 3.5, sry for only doing this now. * 01.03.17 - YSDA deadline on week2 homework moved to 08.03.17 * 28.02.17 - (HSE) homework 4 published * 24.02.17 - Dependencies updated ([same url](https://github.com/yandexdataschool/Practical_RL/issues/1)). Please install theano/lasagne/agentnet until week4 or make sure you're familiar enough with your deep learning framework of choice. * 23.02.17 - YSDA homework 2 can be found [here](https://github.com/yandexdataschool/Practical_RL/tree/master/week2). If you're from HSE you can opt to submit either old or new whichever you prefer. * 17.02.17 - warning! we force-pushed into the repository. Please back-up your github files before you pull! * 16.02.17 - Lecture slides are now available through urls in README files for each week like [this](https://github.com/yandexdataschool/Practical_RL/tree/master/week1#materialshttps://github.com/yandexdataschool/Practical_RL/tree/master/week1#materials). You can also find full archive [here](https://yadi.sk/d/loPpY45J3EAYfU). * 30.03.17 - YSDA deadlines announced for HW 4 * 16.02.17 - HSE homework 3 added * 14.02.17 - HSE deadlines for weeks 1-2 extended! * 14.02.17 - anytask invites moved [here](https://github.com/yandexdataschool/Practical_RL/wiki/Homeworks-and-grading-(HSE-and-YSDA)) * 14.02.17 - if you're from HSE track and we didn't reply to your week0 homework submission, raise panic! * 11.02.17 - week2 success thresholds are now easier: get >+50 for LunarLander or >-180 for MountainCar. Solving env will yield bonus points. * 13.02.17 - Added invites for anytask.org * 10.02.17 - from now on, we'll formally describe homework and add useful links via ./week*/README.md files. [Example.](https://github.com

/yandexdataschool/Practical_RL/blob/master/week0/README.md) * 9.02.17 - YSDA track started * 7.02.17 - HWs checked up * 6.02.17 - week2 uploaded * 27.01.17 - merged fix by _omtcyfz_, thanks! * 27.01.17 - added course mail for homework submission: __practicalrl17@gmail.com__ * 23.01.17 - first class happened * 23.01.17 - created repo

Syllabus

- **week0** Welcome to the MDP
  - Lecture: RL problems around us. Markov decision process. Simple solutions through combinatoric optimization.
  - Seminar: Frozenlake with genetic algorithms
  - Homework description - week0/README.md
    - HSE Homework deadline: *23.59 1.02.17*
    - YSDA Homework deadline: *23.59 19.02.17*
- **week1** Crossentropy method and monte-carlo algorithms
  - Lecture: Crossentropy method in general and for RL. Extension to continuous state & action space. Limitations.
  - Seminar: Tabular CEM for Taxi-v0, deep CEM for box2d environments.
  - Homework description - week1/README.md
    - HSE homework deadline: *23.59 15.02.17*
    - YSDA homework deadline: *23.59 26.02.17*
- **week2** Temporal Difference
  - Lecture: Discounted reward MDP. Value iteration. Q-learning. Temporal difference Vs Monte-Carlo.
  - Seminar: Tabular q-learning
  - Homework description - week2/README.md
    - HSE homework deadline: *23.59 15.02.17*
    - YSDA homework deadline: *23.59 8.03.17*
- **week3** Value-based algorithms
  - Lecture: SARSA. Off-policy Vs on-policy algorithms. N-step algorithms. Eligibility traces.
  - Seminar: Qlearning Vs SARSA Vs expected value sarsa in the wild
  - Homework description - week3/README.md
    - HSE homework deadline *23.59 22.02.17*
    - YSDA homework deadline: *23.59 14.03.17*
- **week3.5** Deep learning recap
  - Lecture: deep learning, convolutional nets, batchnorm, dropout, data augmentation and all that stuff.
  - Seminar: Theano/Lasagne on mnist, simple deep q-learning with CartPole (TF version contrib is welcome)
  - Homework - convnets on MNIST or simple deep q-learning - week3.5/README.md
    - HSE homework deadline *23.59 1.03.17*
    - YSDA homework deadline: *23.59 14.03.17* (5 pts)
- **week4** Approximate reinforcement learning
  - Lecture: Infinite/continuous state space. Value function approximation. Convergence conditions. Multiple agents trick.
  - Seminar: Approximate Q-learning with experience replay. (CartPole, Acrobot, Doom)
  - Homework - q-learning manually, experience replay - week4/README.md
    - HSE homework deadline *23.59 8.03.17*
    - YSDA homework deadline *23.59 19.03.17*
- **week5** Deep reinforcement learning
  - Lecture: Deep Q-learning/sarsa/whatever. Heuristics & motivation behind them: experience replay, target networks, double/dueling/bootstrap DQN, etc.
  - Seminar: DQN on atari
  - Homework - Breakout with DQN and advanced tricks - week5/README.md
    - HSE homework deadline *23.59 22.03.17*
    - YSDA homework deadline *23.59 26.03.17*

- **week6** Policy gradient methods
  - Lecture: Motivation for policy-based, policy gradient, logderivative trick, REINFORCE/crossentropy method, variance theorem(advantage), advantage actor-critic (incl.n-step advantage)
  - Seminar: REINFORCE manually, advantage actor-critic for MountainCar - week6/README.md
    - HSE homework deadline *23.59 2.04.17*
    - YSDA deadline *23.59 6.04.2017*
- **week6.5** RNN recap
  - Lecture: recurrent neura networks for sequences. GRU/LSTM. Gradient clipping. Seq2seq
  - Seminar: char-rnn and simple seq2seq
    - HSE homework deadline **23.59 5.04.17**
    - YSDA deadline *23.59 9.04.2017*
- **week7** Partially observable MDPs
  - Lecture: POMDP intro. Model-based solvers. RNN solvers. RNN tricks: attention, problems with normalization methods, pre-training.
  - Seminar: Deep kung-fu & doom with recurrent A3C and DRQN
    - HSE homework deadline *23.59 16.04.17* (first submission; kung fu assignment is worth 6pts isntead of 3)
    - YSDA homework deadline *23.59 19.04.17* (first submission)
- **week 8** Case studies 1
  - Lecture: Reinforcement Learning as a general way to optimize non-differentiable loss. Seq2seq tasks: g2p, machine translation, conversation models, image captioning.
  - Seminar: Simple neural machine translation with self-critical policy gradient
    - HSE deadline *23.59 10.05.17* (first submission)
    - YSDA deadline *23.59 10.05.17* (first submission)
- **week 9** Advanced exploration methods
  - Lecture1: Improved exploration methods for bandits. UCB, Thompson Sampling, bayesian approach.
  - Lecture2: Augmented rewards. Density-based models, UNREAL, variational information maximizing exploration, bayesian optimization with BNNs.
  - Seminar: bayesian exploration for contextual bandits
- **week 10** Trust Region Policy Optimization.
  - Lecture: Trust region policy optimization in detail. NPO/TRPO.
  - Seminar: approximate TRPO vs approximate Q-learning for gym box2d envs (robotics-themed).
    - HSE deadline *23.59 18.05.17* (first & last submission)
    - YSDA deadline *23.59 18.05.17* (first & last submission)
- **week 11** Model-based RL: Planning
  - Seminar: MCTS
    - HSE deadline *23.59 18.05.17* (first & last submission)
    - YSDA deadline *23.59 18.05.17* (first & last submission)

Future lectures:

- **week 11** RL in Large/Continuous action spaces.
  - Lecture: Continuous action space MDPs. Value-based approach (NAF). Special case algorithms (dpg, svg). Case study:finance. Large discrete action space problem. Action embedding.
  - Seminar: Classic Control and BipedalWalker with ddpg Vs qNAF. https://gym.openai.com/envs/BipedalWalker-v2 . Financial bot as bonus track.
- **week 12** Advanced RL topics
  - Lecture 1: Hierarchical MDP. MDP Vs real world. Sparse and delayed rewards. When Q-learning fails. Hierarchical MDP. Hierarchy as temporal abstraction. MDP with symbolic reasoning.
  - Lecture 2: Knowledge Transfer in RL & Inverse Reinforcement Learning: basics; personalized medical treatment;

robotics.

Course staff

Course materials and teaching by

- Fedor Ratnikov - lectures, seminars, hw checkups
- Alexander Fritsler - lectures, seminars, hw checkups
- Oleg Vasilev - seminars, hw checkups, technical support
- Pavel Shvechikov - lectures, seminars, HW checkups

Contributors

- Using pictures from http://ai.berkeley.edu/home.html
- Massively refering CS294
- Tensorflow assignments by Scitator
- Other awesome people: see contributors