Learning to Play: The Multi-Agent Reinforcement Learning in Malmo (MARLO) Competition

Speaker: Katja Hofmann
Machine Intelligence and Perception
Microsoft Research

@katjahofmann
The Malmo Collaborative AI Challenge

Goal: foster research in collaborative AI

First round: April / May 2017 (83 registered teams)

Second round planned, starting summer 2018

Details: https://www.microsoft.com/en-us/research/academic-program/collaborative-ai-challenge
Key questions

Can agents generalize?
To new (instances of) games and new opponents?

How can we lower the barrier to entry?
Consider: engineering, compute
Project Malmo: Minecraft as platform for AI research
Project Malmo

A platform for AI experimentation, built on Minecraft

microsoft.com/en-us/research/project/project-malmo/

Open source on github

github.com/Microsoft/malmo

The Malmo Platform for Artificial Intelligence Experimentation
Matthew Johnson, Katja Hofmann, Tim Hutton, & David Bignell 2016
Use Cases and Design Principles

Connect AI agents into the game through an intuitive yet powerful API

Provide researchers with tools for task creation – building on existing Minecraft capabilities

Build for extensions and novel uses – open source; “plug-and-play” design of observation, command, reward handlers

Low entry barrier: provide cross-language (currently: Java, .NET, C/C++, Python, Lua) & cross-platform (Windows, Linux, MacOS) API
A natural environment for multi-agent learning
Decoding multitask DQN in the world of Minecraft

Lydia Liu, Urun Dogan, Katja Hofmann

EWRL & Deep Learning Workshop @ NIPS 2016
ewrl.files.wordpress.com/2016/11/ewrl13-2016-submission-29.pdf
Asynchronous Data Aggregation for End to End Visual Navigation in Minecraft

Mathew Monfort, Matthew Johnson, Aude Oliva, Katja Hofmann

AAMAS 2017
ifaamas.org/Proceedings/aamas2017/pdfs/p530.pdf
The MARLÖ Competition – Multi-Agent Reinforcement Learning in Malmö

Competition Framework
MARLO 2018
Multi-Agent Reinforcement Learning in Minecraft
By Microsoft Research
Starting soon

✓ Streamline
✓ Standardize
✓ Provide baselines
Standardizing the Reinforcement Learning Loop
import gym
import marlo

env = gym.make('MinecraftBasic-v0')
env.init(
    allowContinuousMovement=['move', 'turn'],
    videoResolution=[800, 600]
)
env.reset()

done = False
while not done:
    env.render()
    action = env.action_space.sample()
    obs, reward, done, info = env.step(action)
    print(action)

env.close()
Baselines

ChainerRL is a deep reinforcement learning library built on top of Chainer.

ChainerRL
reinforcement-learning
deep-learning
machine-learning
python
dqn
actor-critic

1,745 commits
7 branches
3 releases
15 contributors
MIT

Branch: master
New pull request

tosunar Merge pull request #279 from muupan/add-dqn-loss-test
assets
add logo
11 months ago

chainer
Merge pull request #271 from udidr/master
docs
Fix wrong directives: autoclass -> autofunction
docs
5 months ago

examples
Update README.md
tests
Parameterize tests
tests
a month ago

tools
Remove the ale install script
tools
update .gitignore
tools
17 days ago

.gitignore
update .gitignore

.travis.yml
Add opencv-python as an optional dependency

CONTRIBUTING.md
Mention autopep8 in CONTRIBUTING.md
LICENSE
Create LICENSE
LICENSE
a year ago

README.md
Add CategoricalDQN to README
readthedocs.yml
Add readthedocs.yml to install chainerrl for docs
requirements-dev.txt
Add opencv-python as an optional dependency
requirements-dev.txt
2 months ago

https://github.com/chainer/chainerrl
## Baselines

ChainerRL is a deep reinforcement learning library built on top of Chainer.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Discrete Action</th>
<th>Continuous Action</th>
<th>Recurrent Model</th>
<th>CPU Async Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>DQN (including DoubleDQN etc.)</td>
<td>✓</td>
<td>✓ (NAF)</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>Categorical DQN</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>DDPG</td>
<td>x</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
</tr>
<tr>
<td>A3C</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ACER</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>NSQ (N-step Q-learning)</td>
<td>✓</td>
<td>✓ (NAF)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>PCL (Path Consistency Learning)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>PPO</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>TRPO</td>
<td>✓</td>
<td>✓</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

[github.com/chainer/chainerrl](https://github.com/chainer/chainerrl)
A BIG Thank You to Monday’s MARLO bootcamp participants!!
The MARLÖ Competition – Multi-Agent Reinforcement Learning in Malmö

Task Design
General Video Game AI: a Multi-Track Framework for Evaluating Agents, Games and Content Generation Algorithms


https://arxiv.org/pdf/1802.10363
http://www.gvgai.net
Games and Tasks for MARLO

- **Game 1**
  - Training: Task 1:1, Task 1:2
  - Validation: Task 1:3
  - Test: Task 1:4

- **Game 2**
  - Training: Task 2:1, Task 1:2
  - Validation: Task 2:3
  - Test: Task 2:4

- **Game N**
  - Training: Task N:1, Task N:2
  - Validation: Task N:3
  - Test: Task N:4

Features:
- **Public**
  - Preliminary rankings: ✔
  - Results: Preliminary
- **Private**
  - Preliminary rankings: ✔
  - Results: Preliminary
  - Preliminary rankings: ✗
  - Results: Final
MARLO Tournament

Each group has $P$ players and the same $N$ games.

The best $P/2$ players of each group progress to the next round.

Quarterfinals:
- Group Q1
- Group Q2

Semifinals:
- Group S1
- Group S2

Quarterfinals:
- Group Q3
- Group Q4

Round of 16:
- Group A
- Group B
- Group C
- Group D

Final Group:
- Grand Final

Round of 16:
- Group E
- Group F
- Group G
- Group H
MARLO Competition games: Mob Chase
Mob Chase – Level Design
Parameters:

- Time & Weather
- Number & Type of mobs
- Number & Block type of exits
- Number of obstacles
- Edge block type (fences)
- Ground block type
- Size of play area
- Number of maximum steps allowed

= Game space size: 6.05E+6 (* level configurations)

Mob Chase - Variants
## Challenge 1: General Sum Games

<table>
<thead>
<tr>
<th></th>
<th>Catch the Pig</th>
<th>Run for the Exit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catch the Pig</td>
<td>5, 5</td>
<td>0,1</td>
</tr>
<tr>
<td>Run for the Exit</td>
<td>1,0</td>
<td>1,1</td>
</tr>
</tbody>
</table>
### Challenge 1: General Sum Games

<table>
<thead>
<tr>
<th></th>
<th>Catch the Pig</th>
<th>Run for the Exit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catch the Pig</td>
<td>5, 5</td>
<td>0,1</td>
</tr>
<tr>
<td>Run for the Exit</td>
<td>1,0</td>
<td>1,1</td>
</tr>
</tbody>
</table>

**Aim:** encourage approaches for general sum games – most realistic but hard!

**Consider uncertainty over reward structure – encourage generalization**
Challenge 2: Extensive Form
Challenge 2: Extensive Form

Values depend on trajectories – combinatorial blow-up

But provides key information, e.g., for opponent modelling
Challenge 3: Incomplete (Partial) Information
Challenge 3: Incomplete (Partial) Information

First-person view provides natural direction for learning to generalize

But provides only a partial view of the game state (and opponent actions)
MARLO Competition games: Build Battle
MARLO Competition games: Treasure Hunt
What’s next?
Summary

Can agents generalize?
To new (instances of) games and new opponents?

How can we lower the barrier to entry?
Consider: engineering, compute

- Streamline
- Standardize
- Provide baselines
Schedule (draft)

Open
July 10, 2018
• Competition Open at Crowd.ai
• Final version of the framework, sample controllers are available
• Training tasks – run locally as warm-up

Qualifying Round
July 10-Oct 7, 2018
• Multi-agent validation tasks
• Top 32 evaluated teams are invited to the final round

Final Round
AIIDE, Nov, 2018
• Multi-agent games in remote server for final tournament
• Live competition!
• Academic program, including papers and invited keynotes!

Tutorials
Aug 14-18, 2018
• Tutorial at IEEE CIG 2018
• 2 hours session of talk and hands-on
Multi-agent validation tasks

Top 32 evaluated teams are invited to the final round

Multi-agent games in remote server for final tournament

Live competition!

Academic program, including papers and invited keynotes!
Prizes!

• Award
  • 1<sup>st</sup> place: 10,000 USD-equivalent Azure plus a travel grant to join a relevant academic conference or workshop.
  • 2<sup>nd</sup> place: 5,000 USD-equivalent Azure.
  • 3<sup>rd</sup> place: 3,000 USD-equivalent Azure.

• Publication
  • The top three entries will be invited as co-authors in a paper summarizing the competition structure, rules, approaches, results and main take-aways.
Follow @Project_Malmo

Project Malmo website aka.ms/malmo

Competition website aka.ms/marlo