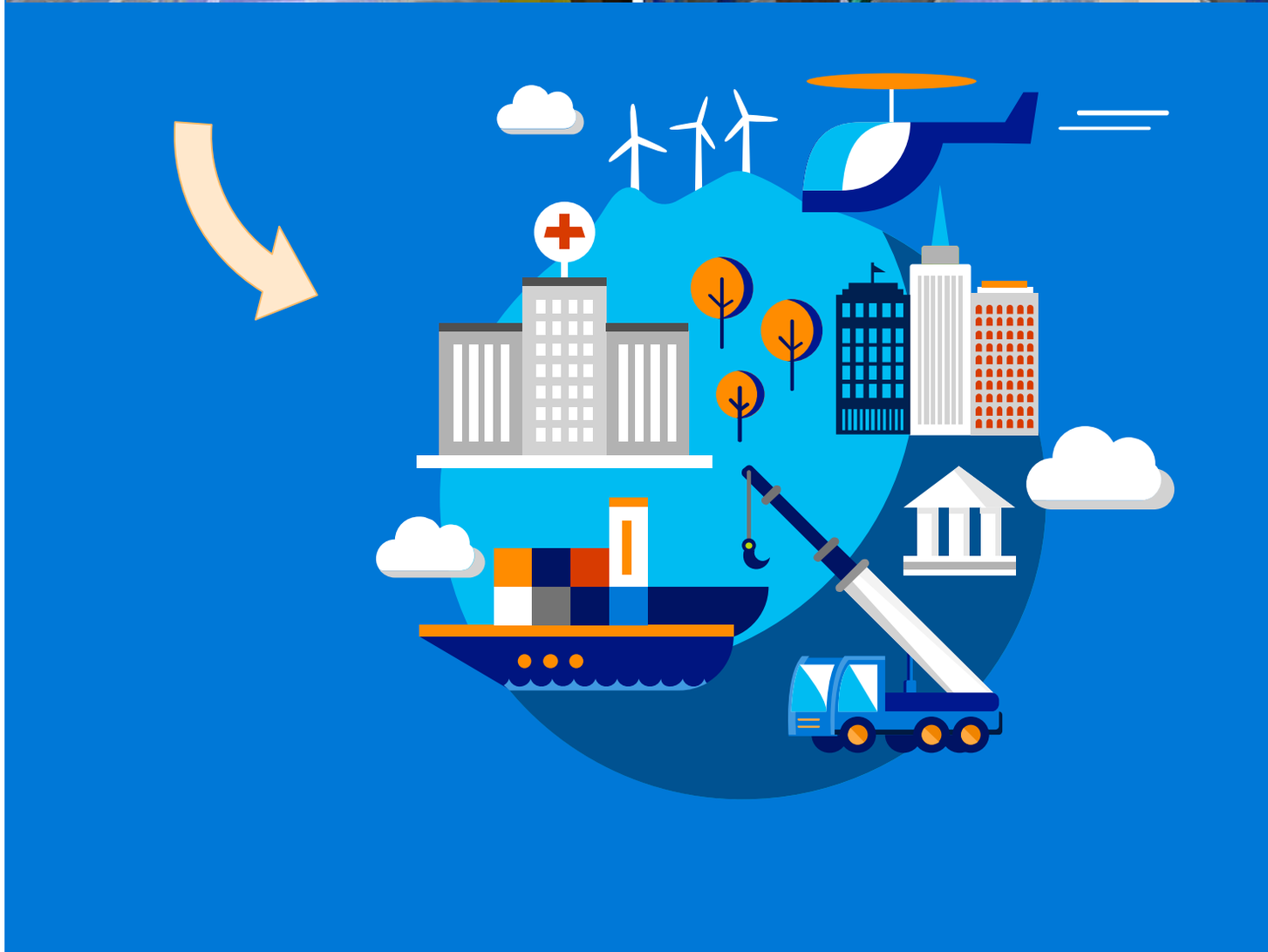
Microsoft

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 planed, starting summer 2018

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

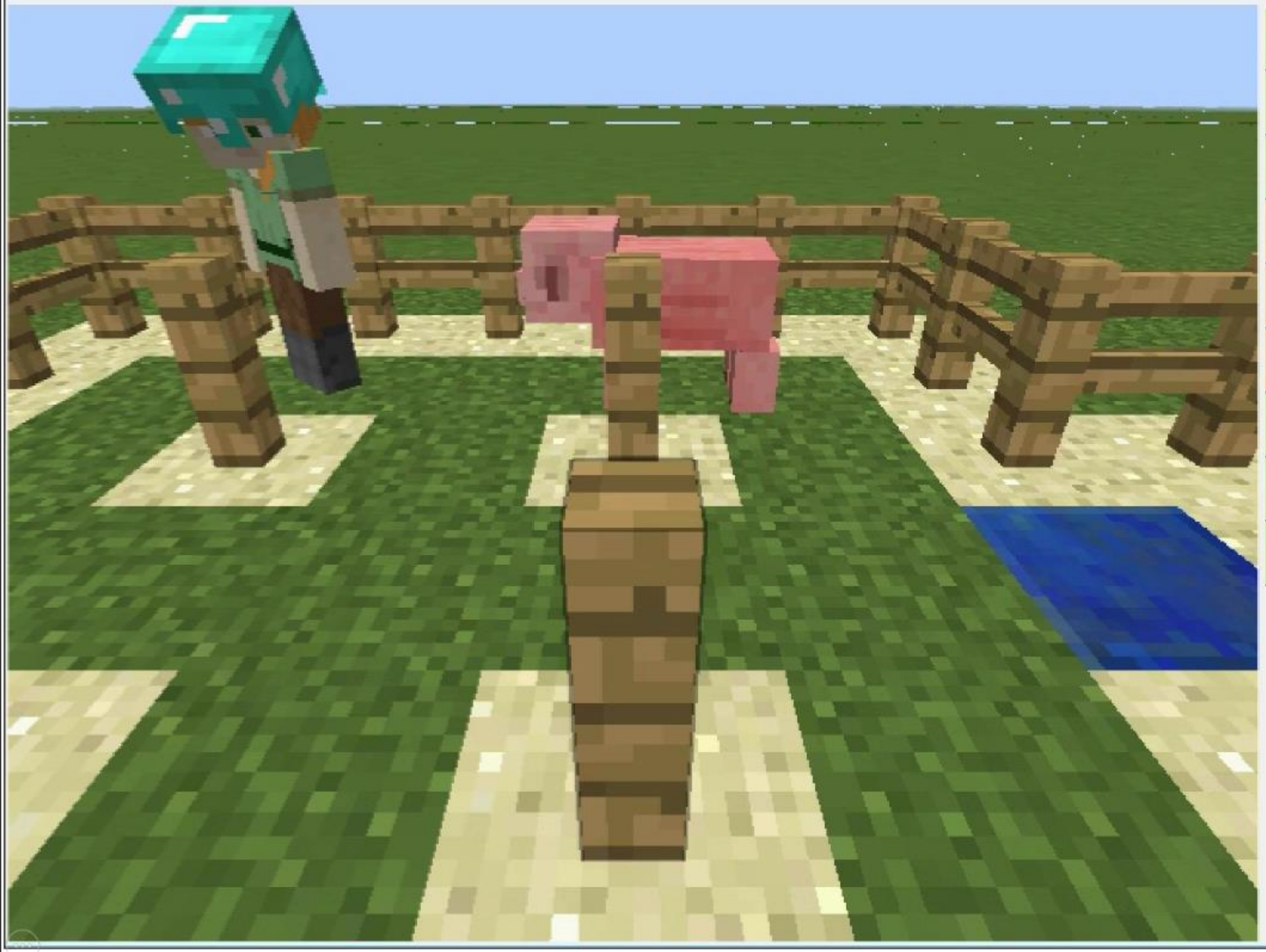
Prizes:

Microsoft Research AI Summer School 2017

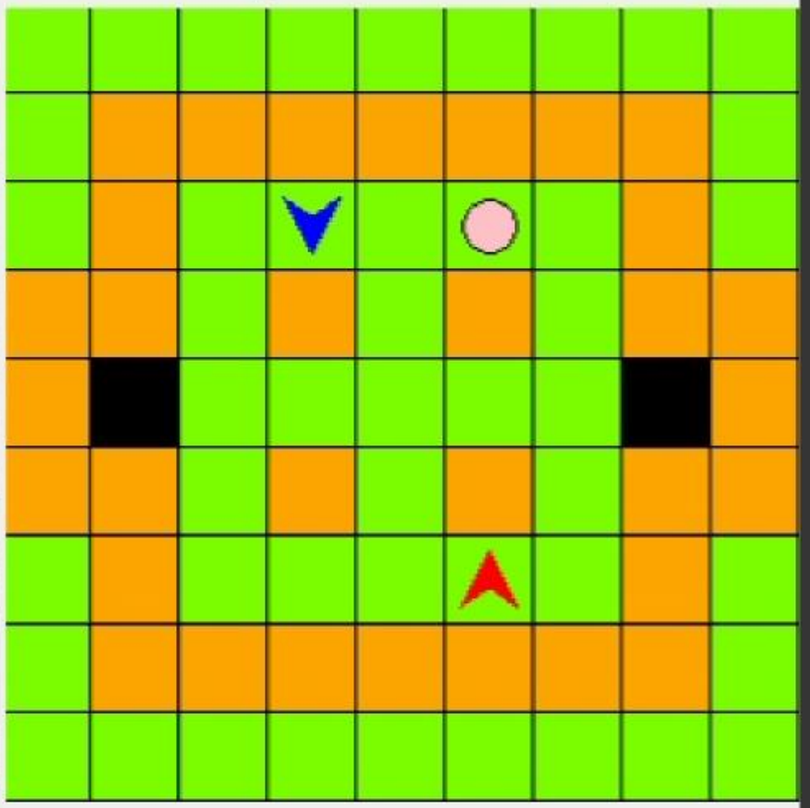
Microsoft Azure for Research

tk

First Person View



Symbolic View



Game stats

Episode: 1

Score: 0

Previous action: None

Actions taken: 0

Actions remaining: 25

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

Microsoft / malmo

Unwatch 233 | Unstar 1,998 | Fork 263

Code | Issues 49 | Pull requests 3 | Wiki | Pulse | Graphs | Settings

Project Malmo is a platform for Artificial Intelligence experimentation and research built on top of Minecraft. We aim to inspire a new generation of research into challenging new problems presented by this unique environment. --- For installation instructions, scroll down to *Getting Started* below, or visit the project page for more information: https://www.microsoft.com/en-us/research/project/project-malmo/ — Edit

695 commits | 4 branches | 10 releases | 11 contributors

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

timhutton committed on GitHub Merge pull request #300 from Microsoft/xerces_init ... Latest commit efcd5b4 3 days ago

.travis	Minor: removed comments.	20 days ago
ALE_ROMS	Applied MIT license.	2 months ago
Malmo	Fix: having two agent_host's in the same script causes a crash becaus...	4 days ago
Minecraft	Fix: use and attack in discrete movement were being sent to first pla...	4 days ago
Schemas	Fix: time 0 was invalid yet suggested in the documentation.	4 days ago
cmake	Fix: changes to make Lua work on Fedora 23.	2 months ago
doc	Minor: fixed item numbering.	5 days ago
sample_missions	Making cliff_walking_1.xml use discrete actions.	a month ago

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

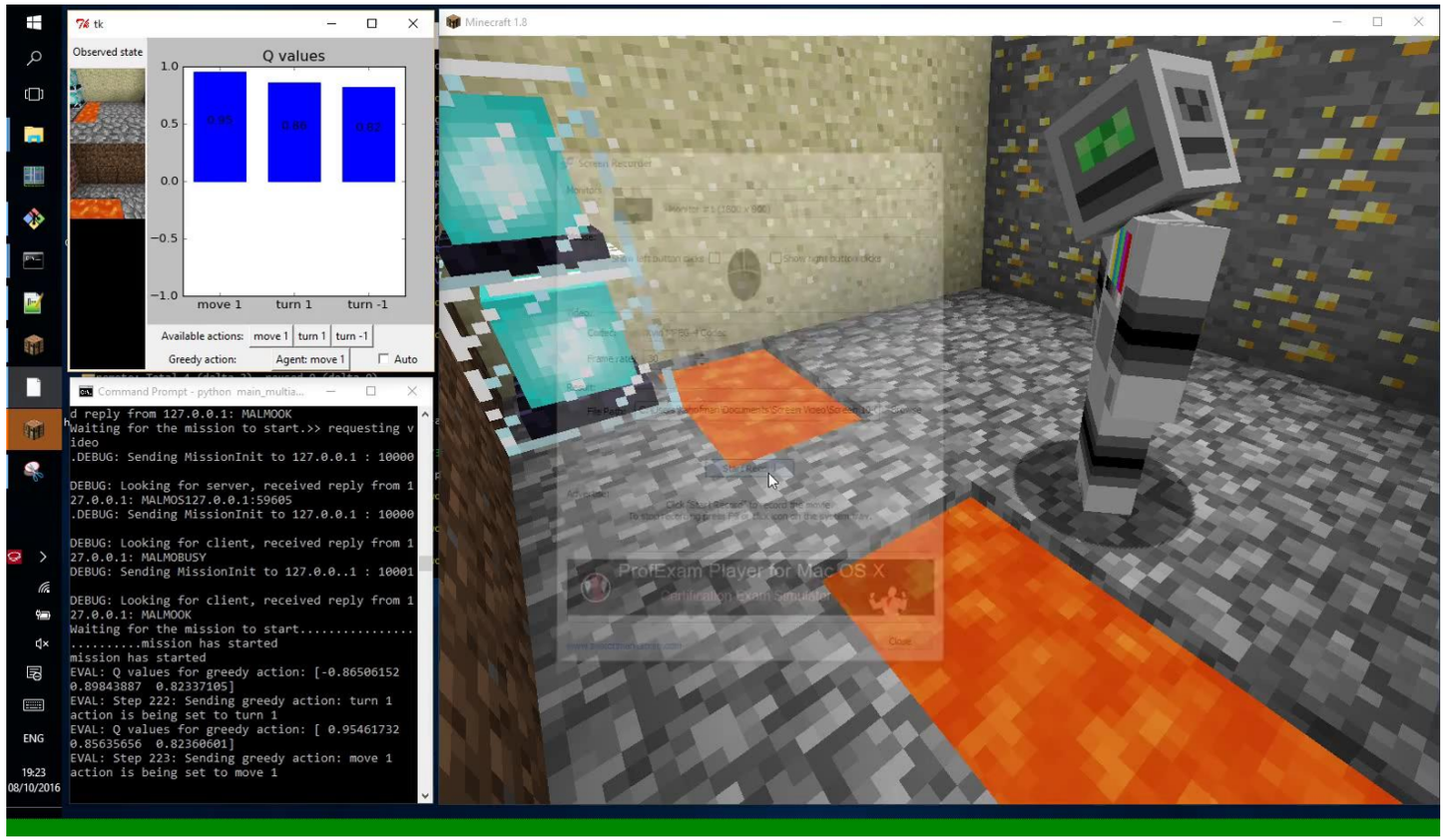
Low entry barrier: provide cross-language (currently: Java, .NET, C/C++, Python, Lua) & cross-platform (Windows, Linux, MacOS) API

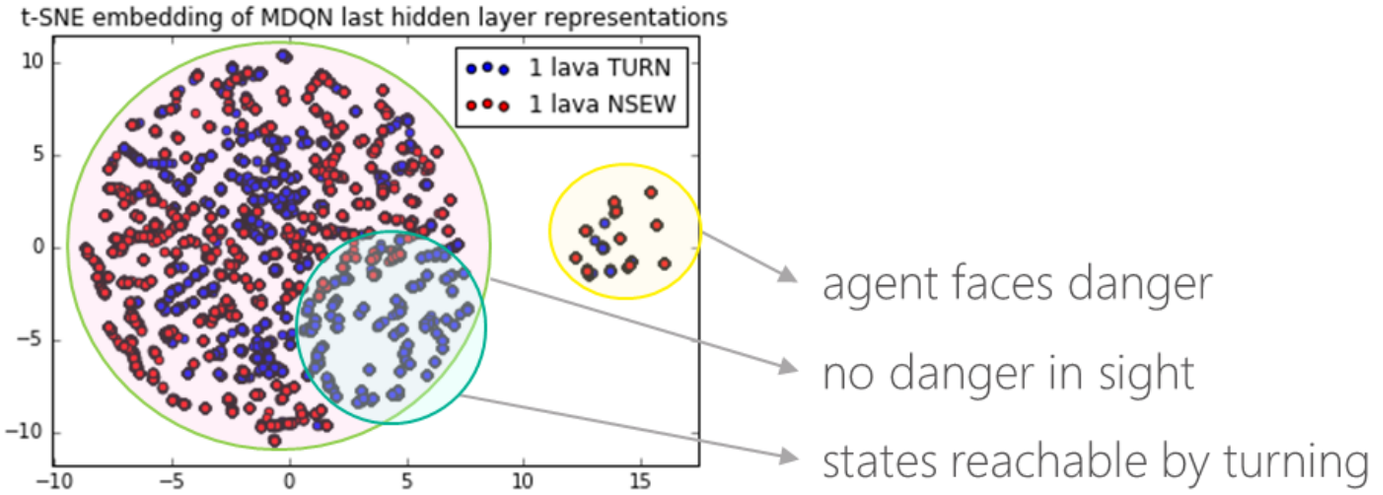
A natural environment for multi-agent learning

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

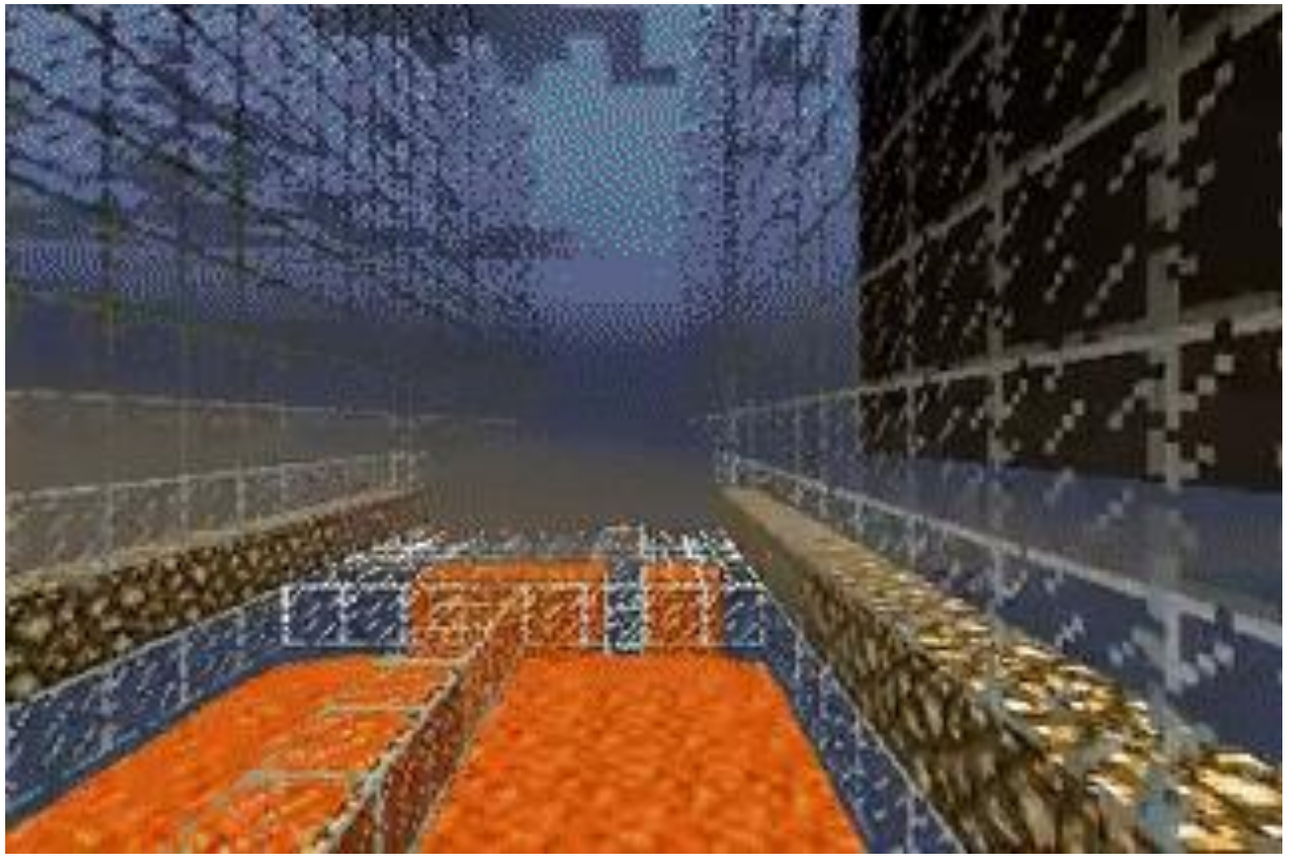


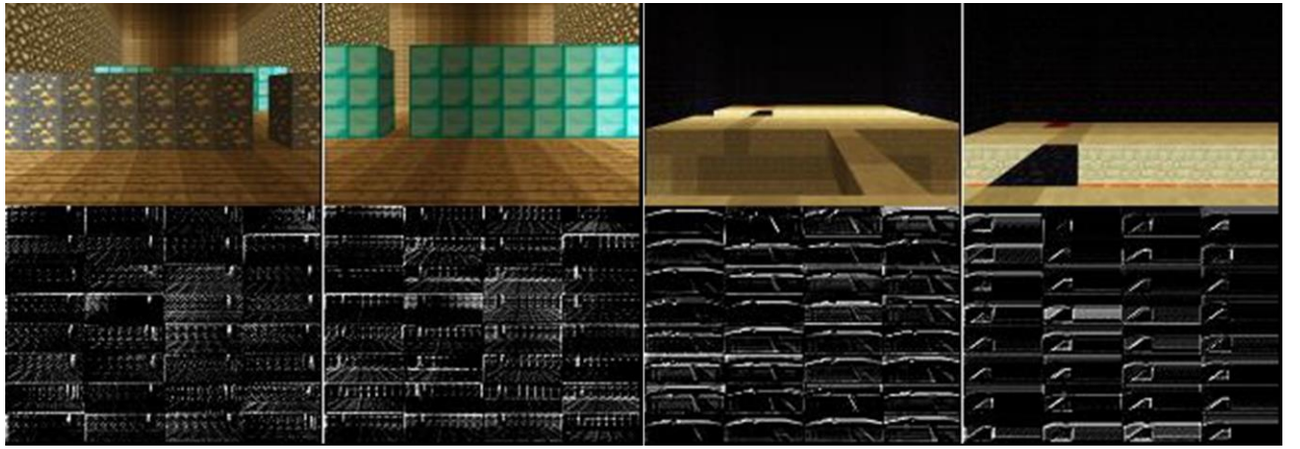


Asynchronous Data Aggregation for End to End Visual Navigation in Minecraft

Mathew Monfort, Matthew Johnson, Aude Oliva, Katja Hofmann

AAMAS 2017
ifaamas.org/Proceedings/aamas2017/pdfs/p530.pdf





The MARLÖ Competition – **M**ulti-**A**gent **R**einforcement **L**earning in Malm**Ö**

Competition Framework

Organizers

crowdAI

Challenges Knowledge Base Job Board

MARLÖ 2018

Multi-Agent Reinforcement Learning in Minecraft

By Microsoft Research

Starting soon

3091 Views

34 Participants

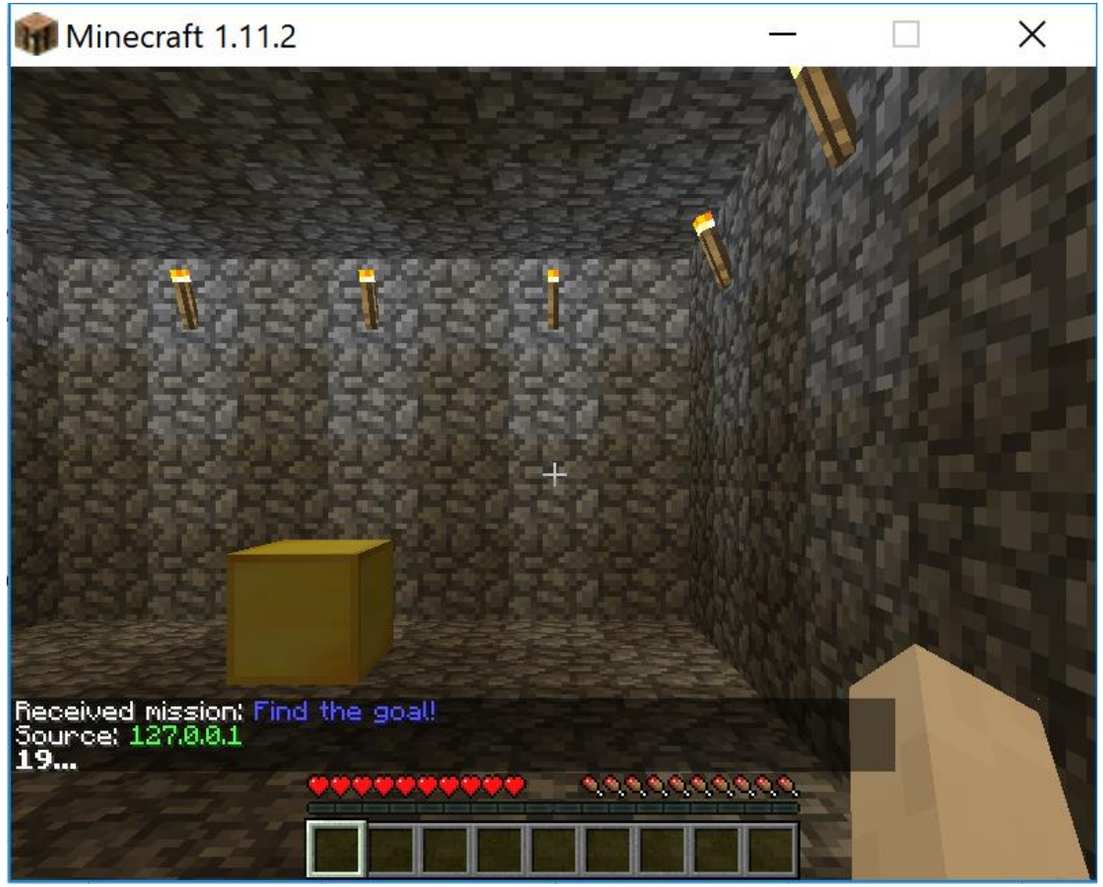
34

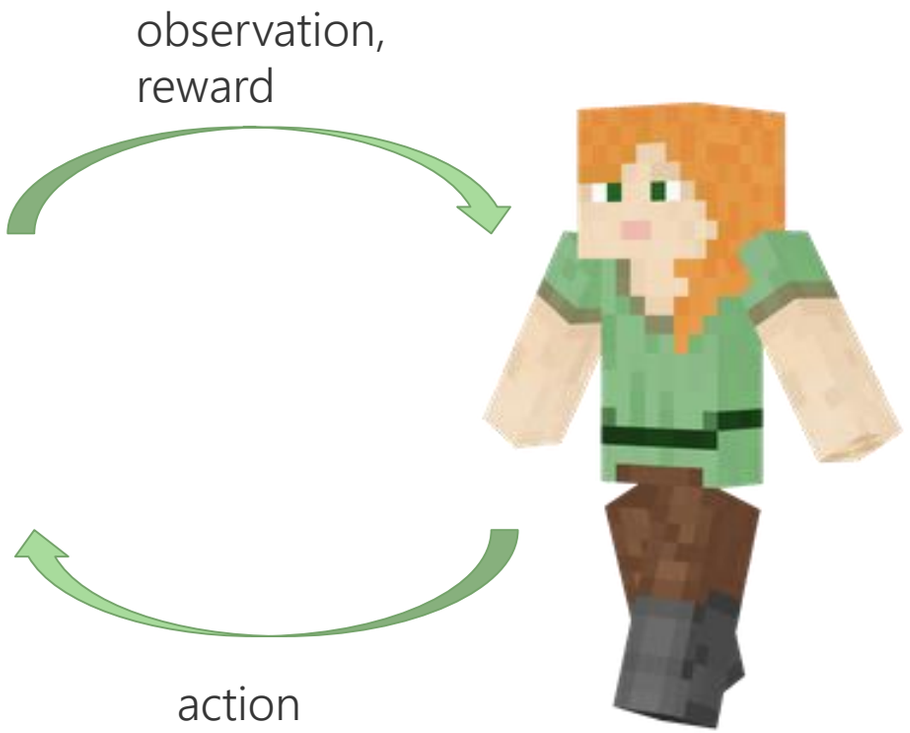
UNFOLLOW

Overview Leaderboard Discussion Dataset Submissions Participants

- ✓ Streamline
- ✓ Standardize
- ✓ Provide baselines

Standardizing the Reinforcement Learning Loop





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

chainer / chainerrl

Watch 76 | Star 485 | Fork 112

Code | Issues 52 | Pull requests 21 | Projects 0 | Wiki | Insights

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

chainer 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 | Create new file | Upload files | Find file | Clone or download

toslunar Merge pull request #279 from muupan/add-dqn-loss-test ... Latest commit e424a75 14 days ago

assets	add logo	11 months ago
chainerrl	Merge pull request #271 from uidilr/master	a month ago
docs	Fix wrong directives: autoclass -> autofunction	5 months ago
examples	Update README.md	a month ago
tests	Parameterize tests	17 days ago
tools	Remove the ale install script	11 months ago
.gitignore	update .gitignore	2 months ago
.travis.yml	Add opencv-python as an optional dependency	2 months ago
CONTRIBUTING.md	Mention autopep8 in CONTRIBUTING.md	4 months ago
LICENSE	Create LICENSE	a year ago
README.md	Add CategoricalDQN to README	3 months ago
readthedocs.yml	Add readthedocs.yml to install chainerrl for docs	a year ago
requirements-dev.txt	Add opencv-python as an optional dependency	2 months ago

https://github.com/chainer/chainerrl

Baselines

chainer / chainerrl

Watch 76 | Star 485 | Fork 112

Code | Issues 52 | Pull requests 21 | Projects 0 | Wiki | Insights

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

chainer | reinforcement-learning | deep-learning | machine-learning | python | dqn | actor-critic

MIT

Clone or download

ommit e424a75 14 days ago

Algorithm	Discrete Action	Continous Action	Recurrent Model	CPU Async Training
DQN (including DoubleDQN etc.)	✓	✓ (NAF)	✓	x
Categorical DQN	✓	x	✓	x
DDPG	x	✓	✓	x
A3C	✓	✓	✓	✓
ACER	✓	✓	✓	✓
NSQ (N-step Q-learning)	✓	✓ (NAF)	✓	✓
PCL (Path Consistency Learning)	✓	✓	✓	✓
PPO	✓	✓	x	x
TRPO	✓	✓	x	x

11 months ago
a month ago
5 months ago
a month ago
17 days ago
11 months ago
2 months ago
2 months ago
4 months ago
a year ago
3 months ago
a year ago

requirements-dev.txt Add opencv-python as an optional dependency 2 months ago

https://github.com/chainer/chainerrl

A BIG Thank You to Monday's
MARLO bootcamp participants!!

The MARLÖ Competition – **M**ulti-**A**gent **R**einforcement **L**earning in Malm**Ö**

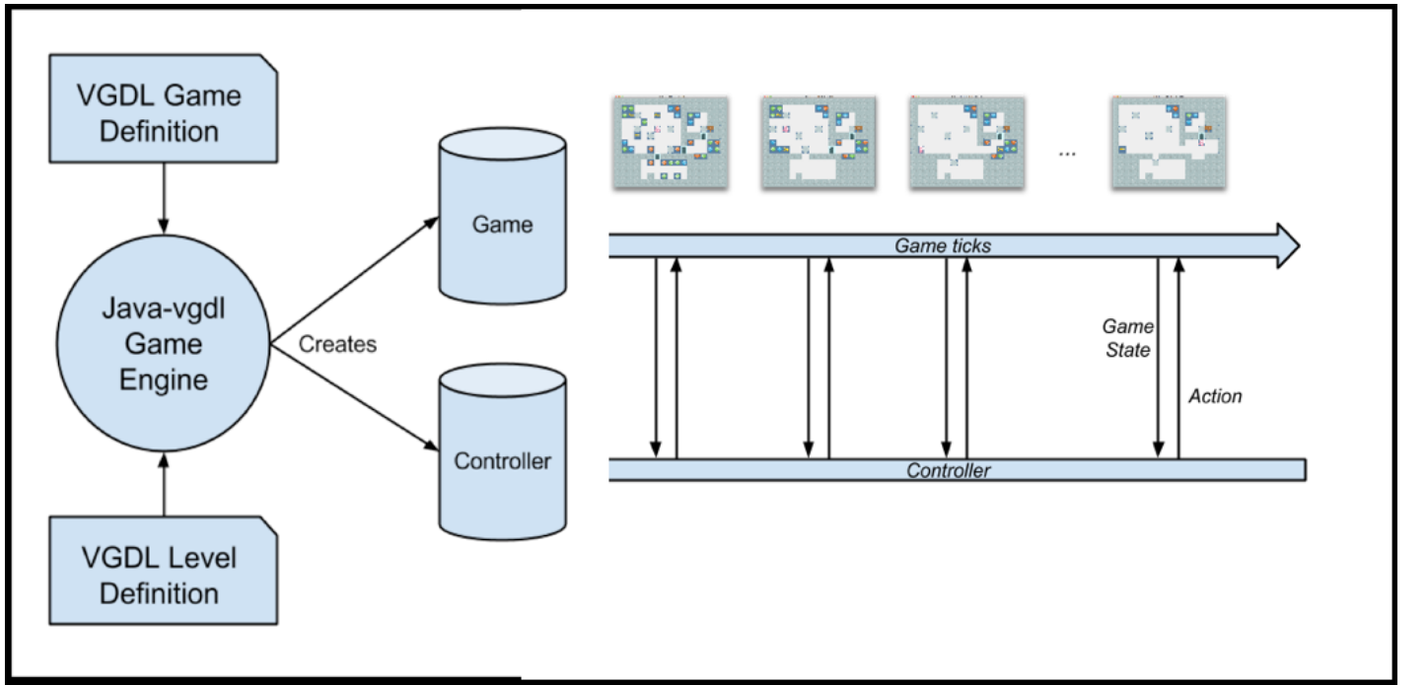
Task Design

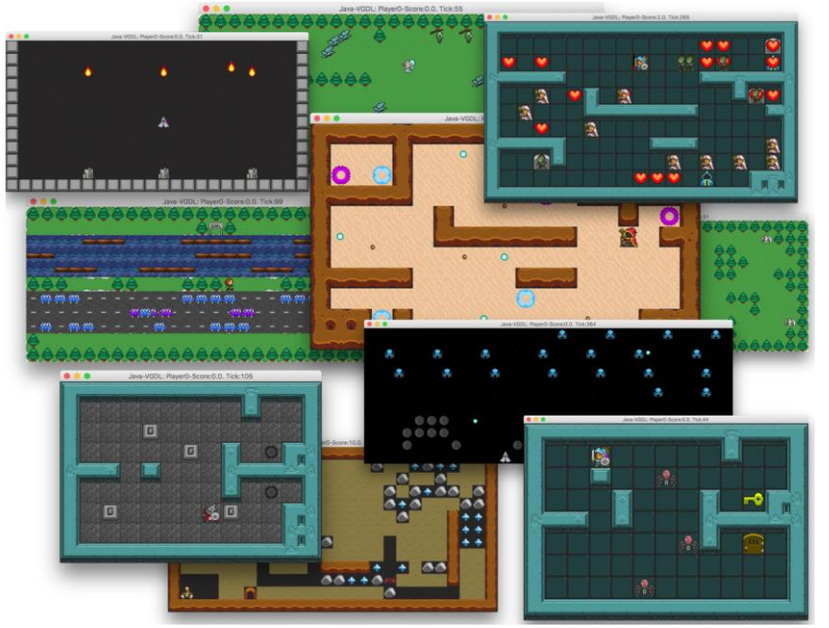
General Video Game AI: a Multi-Track Framework for Evaluating Agents, Games and Content Generation Algorithms

Diego Perez-Liebana, Jialin Liu, Ahmed Khalifa, Raluca D. Gaina, Julian Togelius, Simon M. Lucas

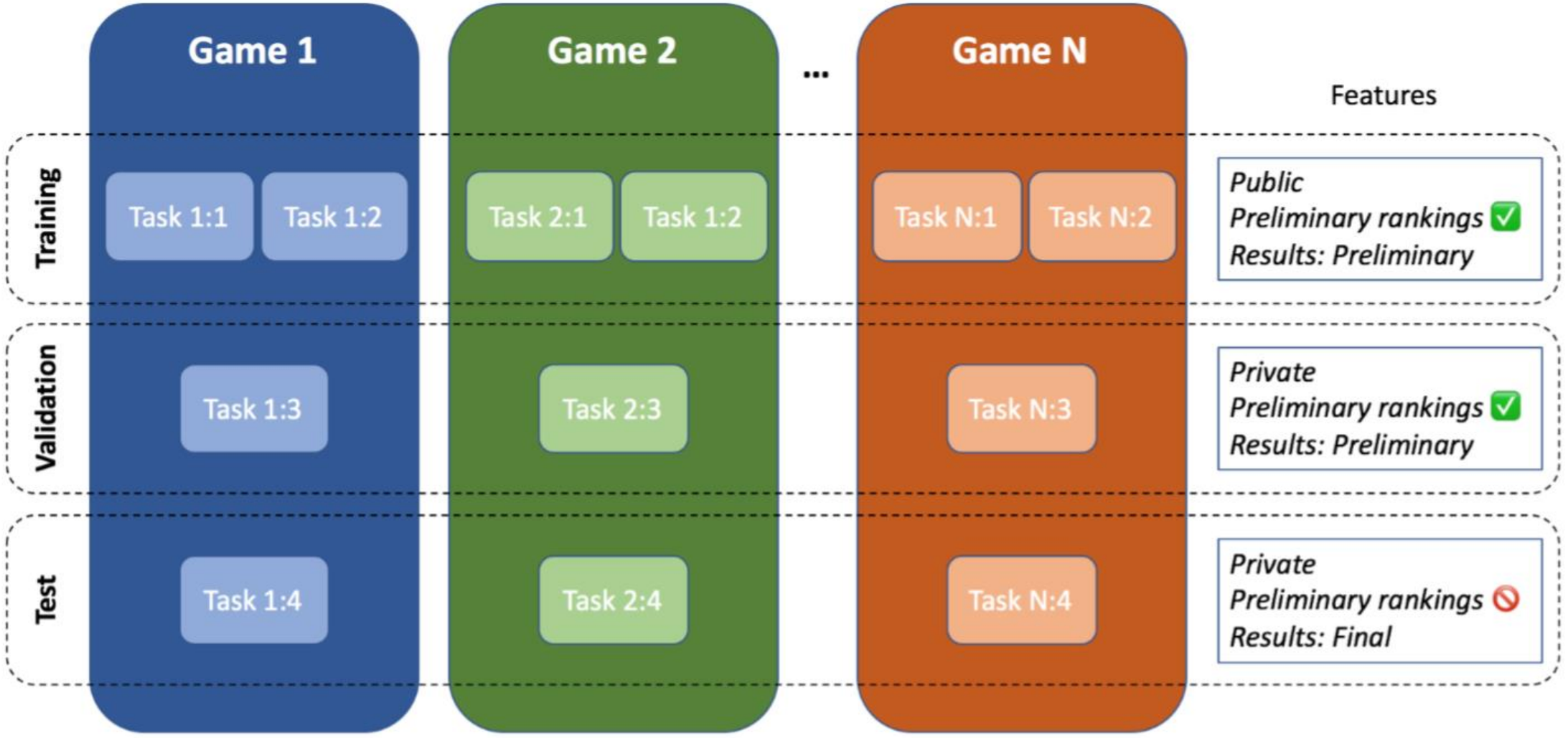
https://arxiv.org/pdf/1802.10363

http://www.gvgai.net

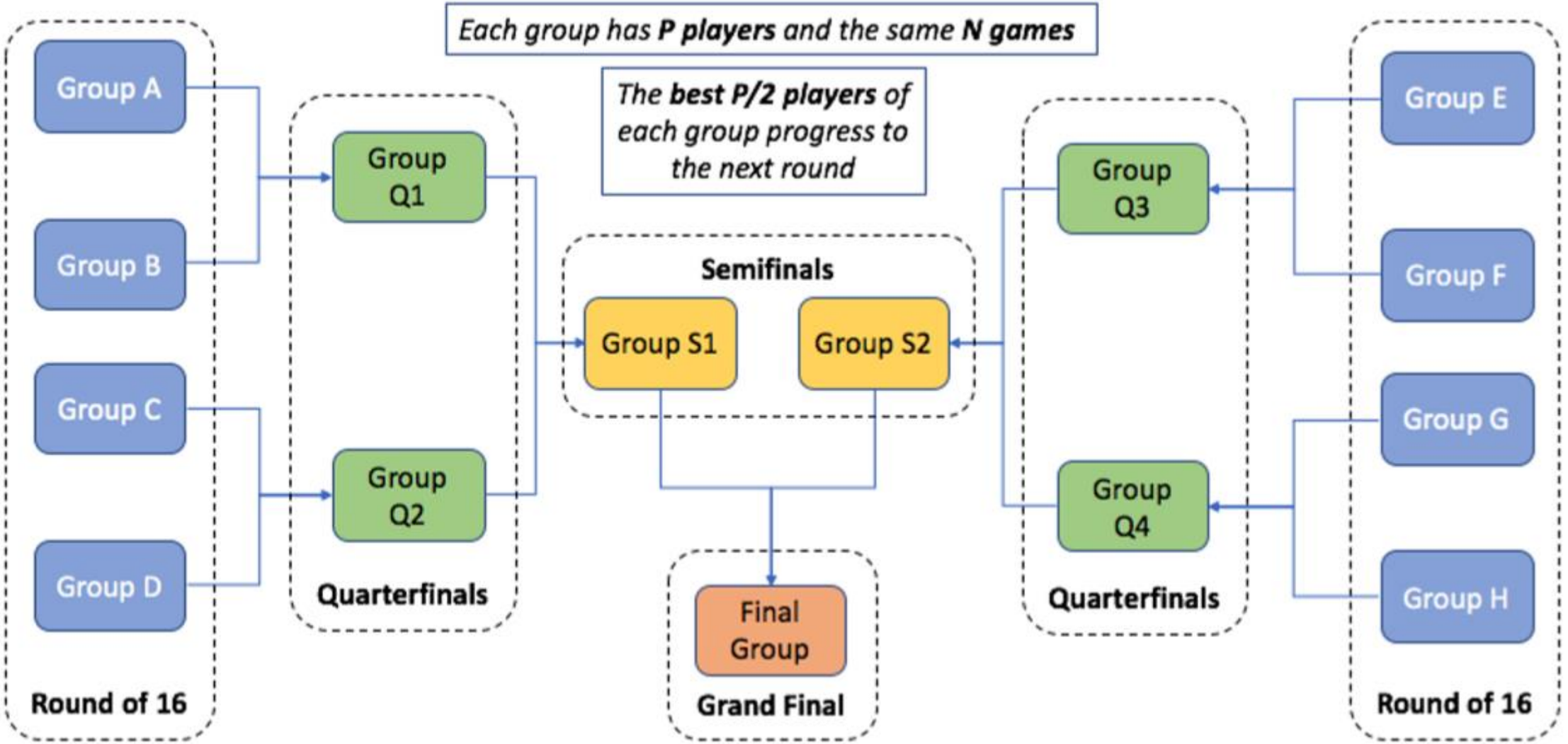




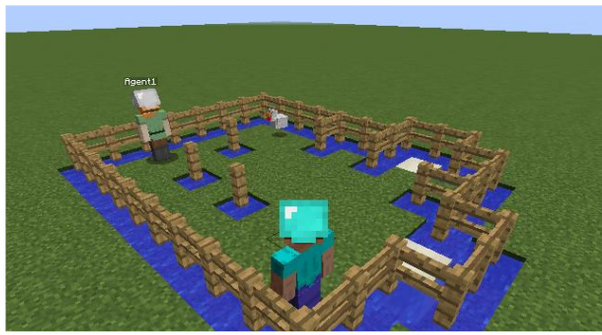
Games and Tasks for MARLO

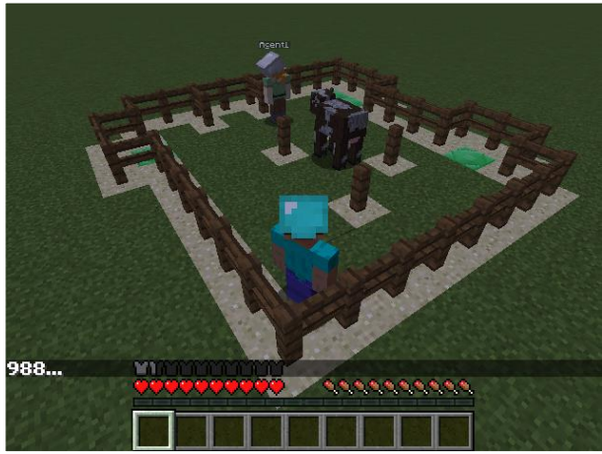


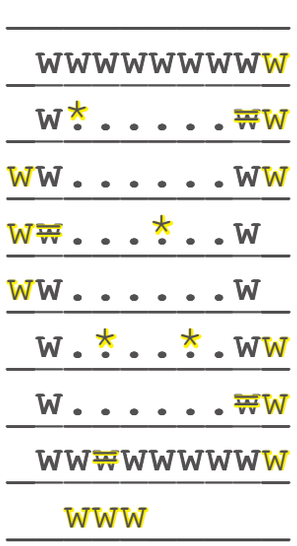
MARLO Tournament

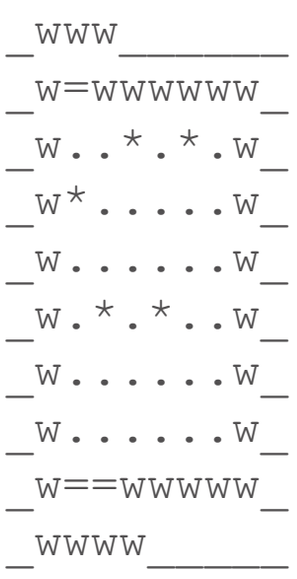


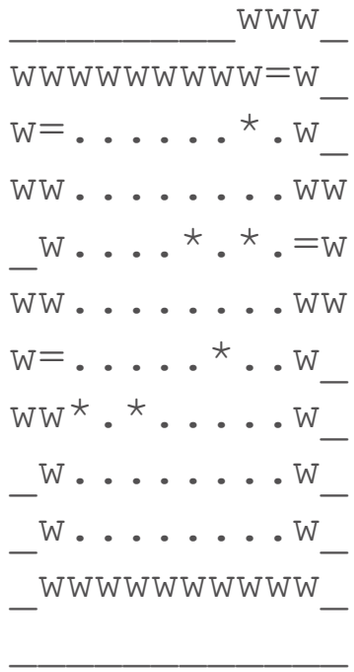
MARLO Competition games: Mob Chase











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

	Catch the Pig	Run for the Exit
Catch the Pig	5, 5	0,1
Run for the Exit	1,0	1,1

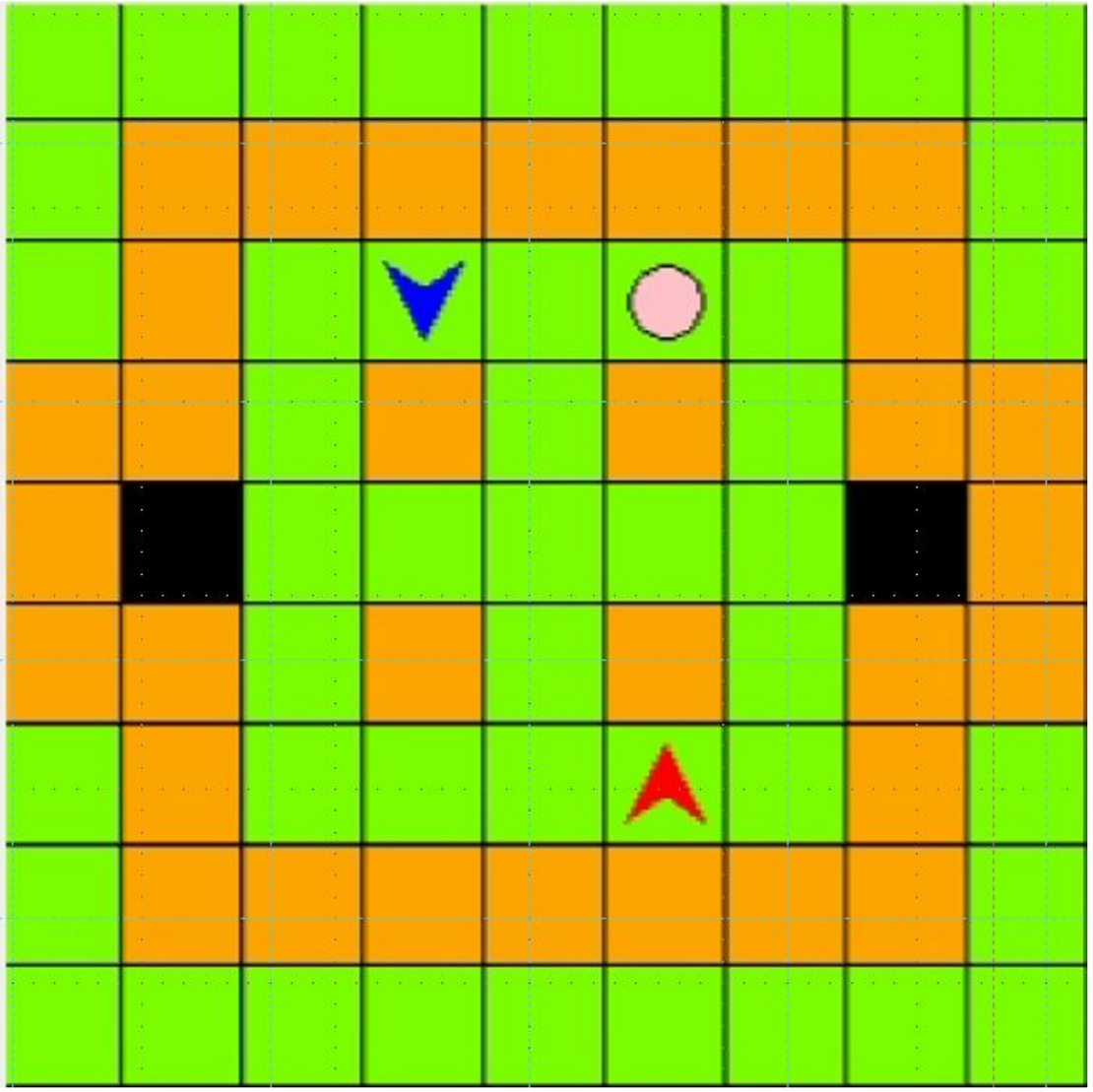
Challenge 1: General Sum Games

	Catch the Pig	Run for the Exit
Catch the Pig	5, 5	0,1
Run for the Exit	1,0	1,1

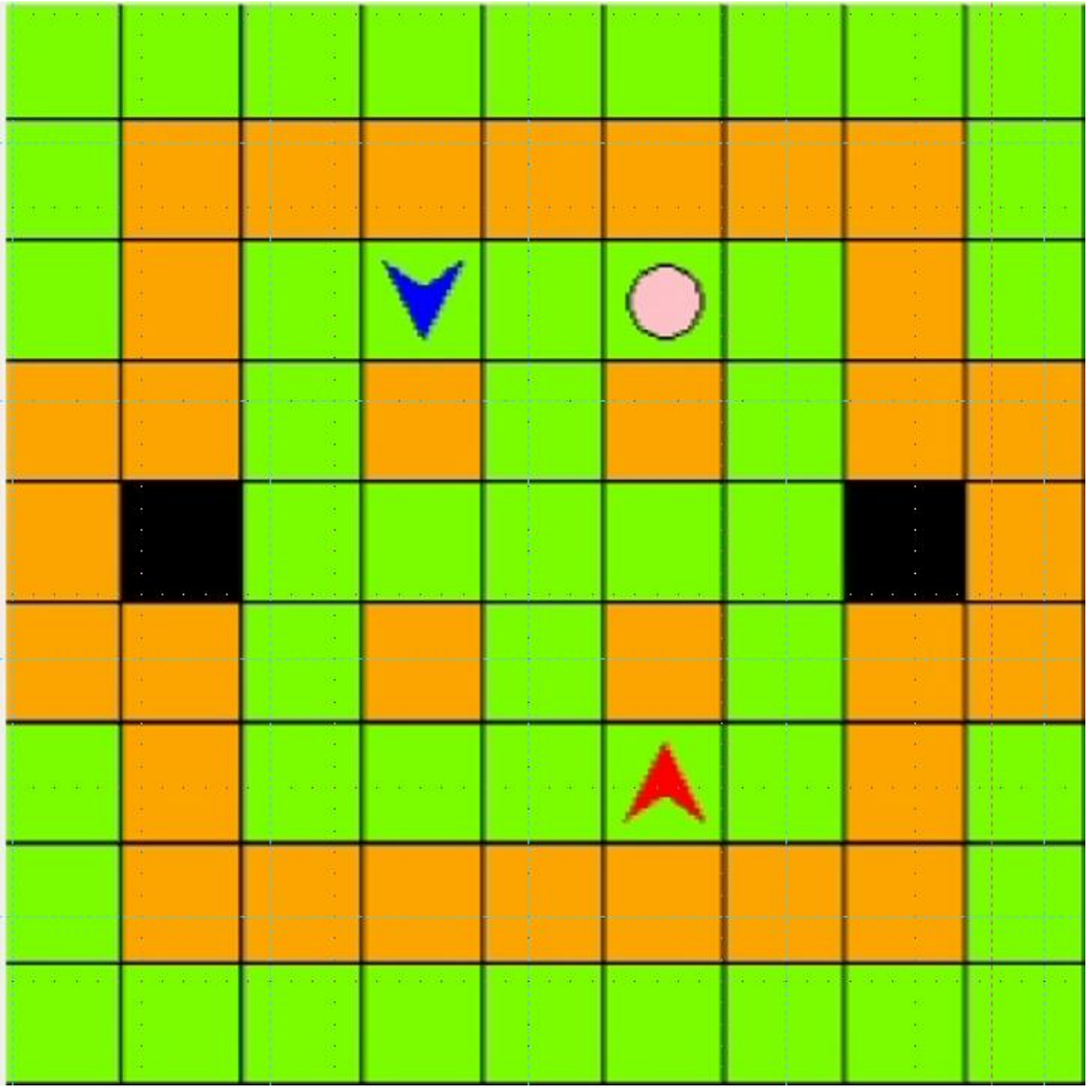
Aim: encourage approaches for general sum games – most realistic but hard!

Consider uncertainty over reward structure – encourage generalization

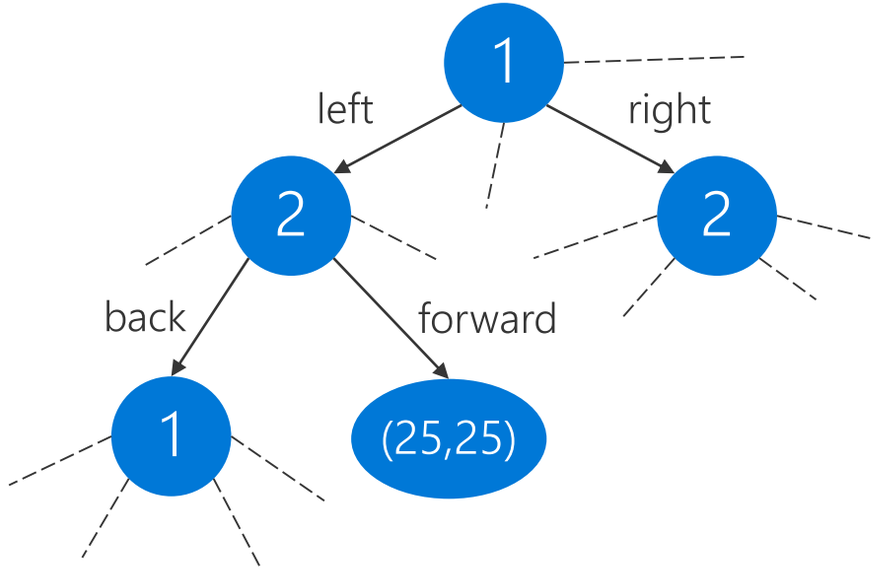
Challenge 2: Extensive Form



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

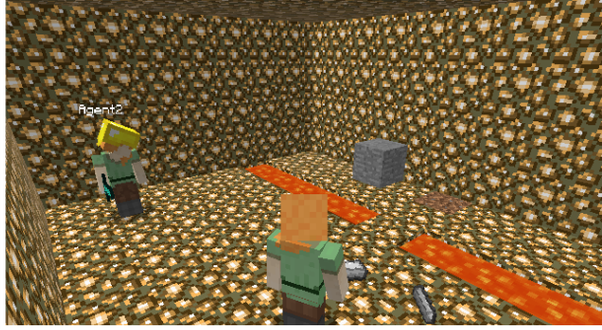
Challenge 3: Incomplete (Partial) Information

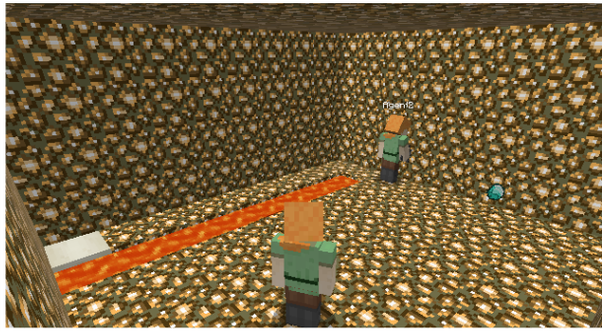
First-person view provides natural direction for learning to generalize

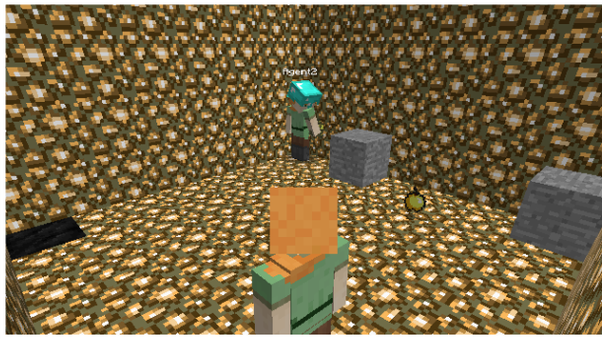
But provides only a partial view of the game state (and opponent actions)

MARLO Competition games: Build Battle

MARLO Competition games: Treasure Hunt







What’s next?

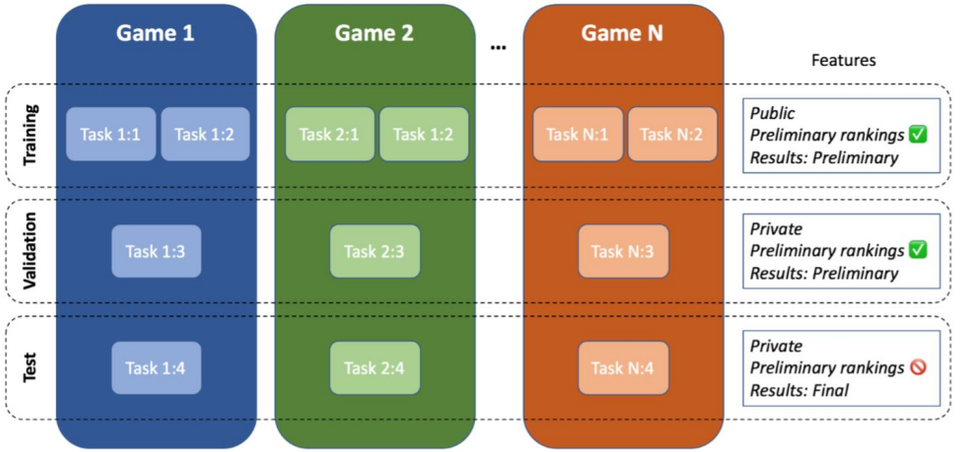
Summary

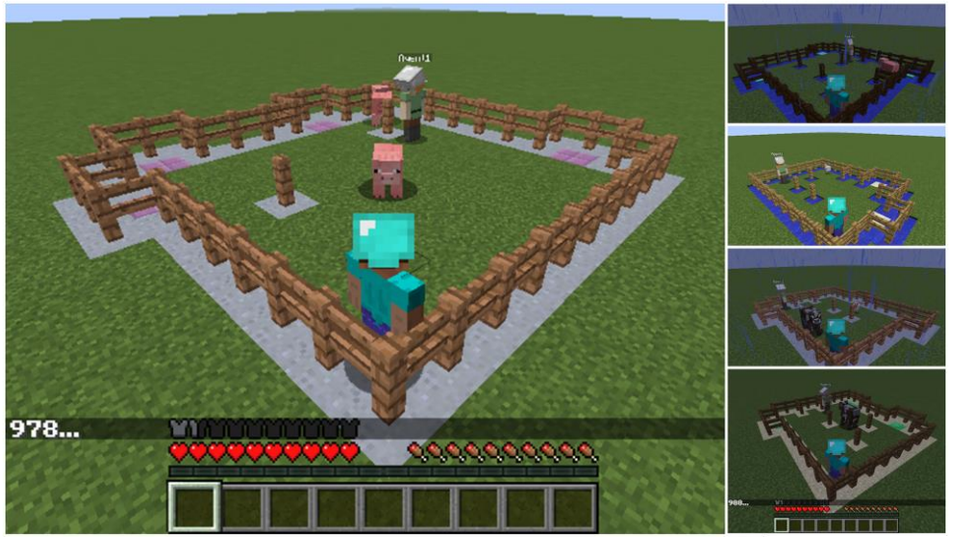
Can agents generalize?

To new (instances of) games and new opponents?

How can we lower the barrier to entry?

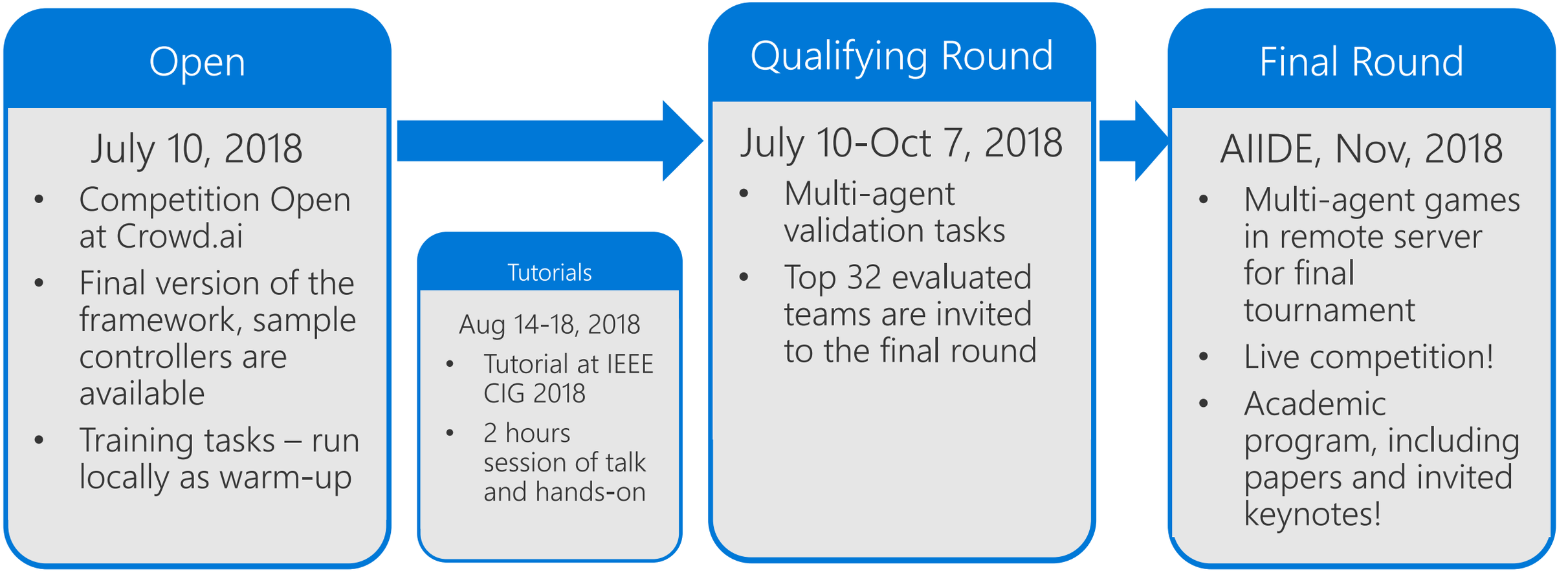
Consider: engineering, compute





- ✓ Streamline
- ✓ Standardize
- ✓ Provide baselines

Schedule (draft)



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

Tutorials

Aug 14-18, 2018

- Tutorial at IEEE CIG 2018
- 2 hours session of talk and hands-on

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!

Submissions for contributed talks + extended abstracts open now until July 27!

Prizes!

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

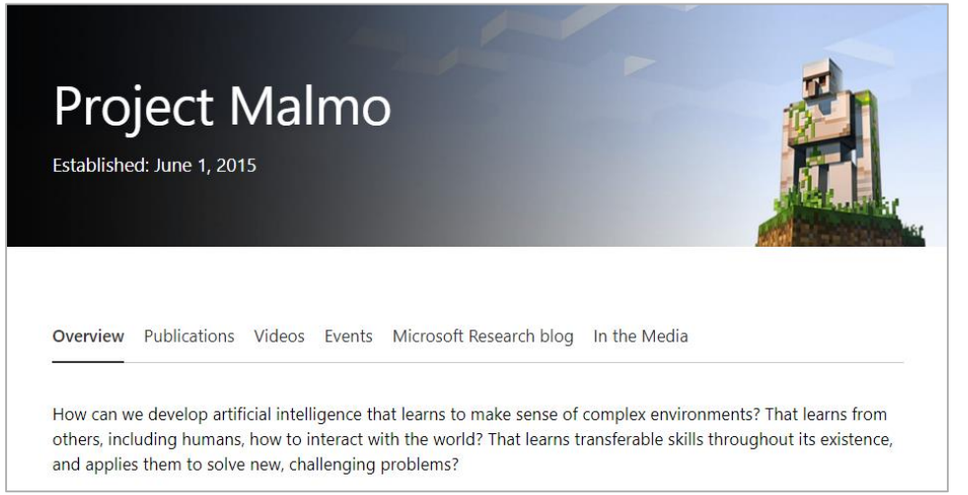
@Project_Malmo

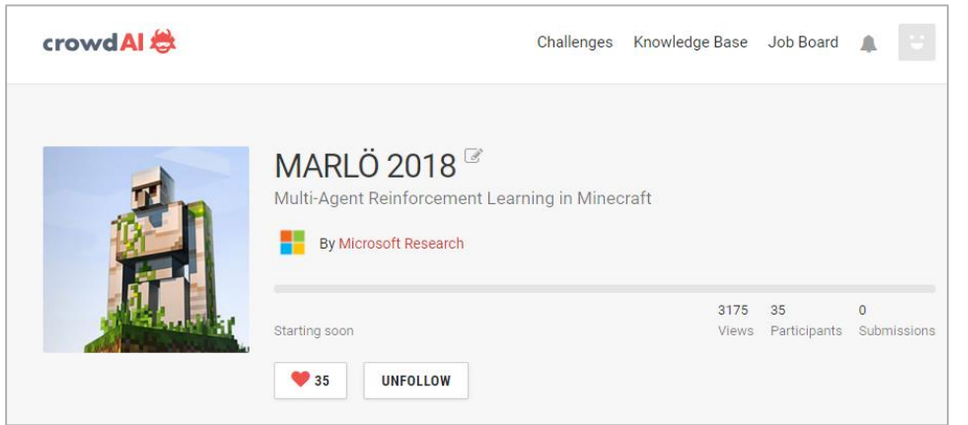
Project Malmo website

aka.ms/malmo

Competition website

aka.ms/marlo





Microsoft

© Copyright Microsoft Corporation. All rights reserved.