

Board Game Al Final Presentation

European Master Team Project @ InES



Introduction



The Team



Tobias Sesterhenn



Raluca Chis



Roman Hess



Julie Naegelen



Tudor Esan



Timur Carstensen



The Timeline

	Feb	Mar	Apr	Mai	June	July	Aug
Kick-off							
Phase 1: Mannheim							
Onboarding Cluj							
Phase 2: Mannheim and Cluj							
Intermediate Presentation							
Phase 3: Mannheim							
Final Presentation							



The Challenge







Deliverables

See GitHub Repo



Connect Four Game AI: RL Training Wheels

Connect Four: a competitive two-player game of limited dimensions.

Train a RL agent against:

- 1) random player
- 2) heuristic player

Used Ray [RLlib & Tune]: entire pipeline in python.







Game Simulator

Image of Board State

TT Environment



BugPlus: Turing Tumble RL Environment





RL Pipeline



(1) Supervised Pretraining: Expert Play

(2) RL: Curriculum Learning



Pretraining (1/5)

 \rightarrow Network trained on Expert Play

- \rightarrow Expert = Simple Algorithm that decides on:
 - which bits to connect next
 - which edge to remove next



\rightarrow Supervised training:

Network Output		Label	
[0.2 0.6 0.2]	<->	[0 1 0]	ightarrow next edge prediction
[0.4 0.4 0.2]	<->	[0.5 0.5 0]	\rightarrow edge removing

→ Expert Play can be improved >>> Reinforcement Learning



Pretraining (2/5)





Pretraining (3/5)





Pretraining (4/5) - Rollout

 \rightarrow Start with random board initialization and input-output pairs

 \rightarrow Let the network take actions until the target board is reached (all input-output pairs are included)

 \rightarrow Solved Programs = 88/100





Pretraining (5/5) - Rollout

 \rightarrow Start with zero board initialization and input-output pairs

 \rightarrow Let the network take actions until the target board is reached (all input-output pairs are included)

 \rightarrow Found Solutions = 83/100





RL Pipeline



(1) Supervised Pretraining: Expert Play

(2) RL: Curriculum Learning



Curriculum Learning (1/3)

Gradually increase the difficulty of the problems during training

 \rightarrow start with problems that are one step away from the solution at first





Curriculum Learning (2/3)

Initialise learner with weights from pretrained model:

 \rightarrow start training closer to the (global) optimum of the loss surface

 \rightarrow allow the learner to find alternative (more sophisticated) solutions to that of the (pretrained) expert





Curriculum Learning (3/3)

Takeaways:

→ **curriculum learning works**: learner is able to adapt to increasing challenge difficulty

- \rightarrow pretraining accelerates the learning progress, but:
 - learning progress stagnates and falls below non-pretrained models

 \rightarrow so far: fail to initialise value branch of PPO trainer

Going forward:

- \rightarrow pretrain the value branch of the network
- \rightarrow modify exploration behaviour



Recap and Outlook



Lessons learned

Remote Online Working in an International Team

Project Management Platforms can be useful, however:

- Effort needed to keep it up to date
- Esp. when tasks change dynamically
- Dailies were often sufficient to keep track of progress

Theoretical Knowledge: Reinforcement Learning

Practical Knowledge:

- Scrum
- Libraries: Ray [Tune & RLLib]; JPype; Weights & Biases
- Version Control with Git & Github
- Technicalities: properly debugging code
- Technicalities: use custom FFN instead of RLlib out-of-the box networks













Possibilities for improvement

- Fully integrated GUI:
 - not only generate visual board as output, but use board as input.

😉 😑 🧸

- HPO
- Better communication with project supervision on theoretical foundations:
 - \circ had to change plans late in the project pursuing unrealistic goals.
- Different Platforms were an issue
 - Spend more time in the beginning on the technical setup: Docker Solution early on.



Credits

- Graphical Turing Tumble Simulator: Jesse Crossen
 - <u>https://github.com/jessecrossen/ttsim</u>
- BugBit Java environment: Christian Bartelt



Further information

- You are interested in learning more about this project?
- After August 7th, our website launches
- More Information about:
 - \circ The project
 - The study trips to Cluj and Mannheim
 - The team members





Thank you for your Attention!

Now to Q&A...