CS230 - Learning Minichess Without Human Knowledge

Karthik selvakumar Bhuvaneswaran

karthik0@stanford.edu

Abstract

Using WebRTC we are trying to develop a product called 'Ping' used for audio, mult-video, file and screen sharing.

- Ping uses WebRTC for the source of data exchange and XMPP Server for signalling and transporting.
- Ping works on Browser to Browser connections instead of naive client server approach.
- Ping guarentees high scalability and upto 60% more efficiency than existing systems.

Logic and MCTS

- WebRTC is an upcoming standard that aims to improve real time communication among web browsers in peer to peer fashion.
- Ping uses WebRTC which allows browsers to natively support interactive peer to peer communication and realtime collaboration.

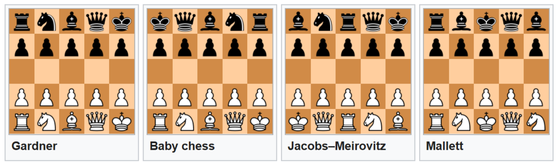

Figure 1:Popular Minichess Board Layouts

$$l = \Sigma(v_\theta(s_t) - z_t)^2 + \vec{\pi}_t log(\vec{p}_\theta(s_t))$$

Distributed Architecture

We designed an efficient architecture for Ping to withstand high traffic.

- At start browsers do not know each other.
- WebRTC mediates the setup process through the XMPP server of Ping.

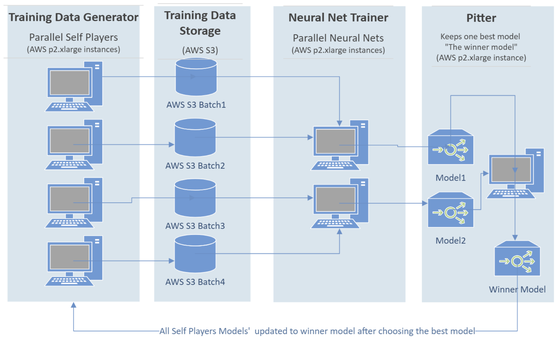

Figure 2:Distributed Cloud Architecture for Reinforcement Learning

Neural Network and hyperparams

- RTC Peer connection is the WebRTC API that handles stable and efficient communication of streaming data between peers.

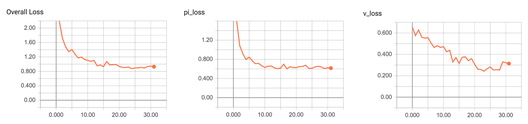

Figure 4:Loss values after each epoch

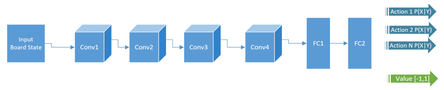

Figure 5:Baseline Random vs Greedy vs Neural Net

Performance Gain Distributed vs Single Instance

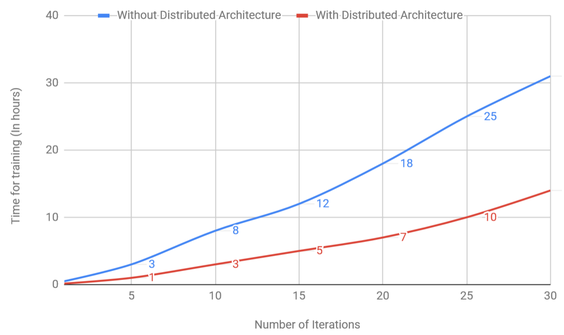

Figure 3:Distributed Cloud Architecture for Reinforcement Learning

Baseline Random vs Greedy vs Neural Net

Baseline	Color	Won	Lost	Draw
Random Player	White	10	0	0
	Black	10	0	0
Greedy Player	White	10	0	0
	Black	3	0	7
Model Version 1	White	10	0	0
	Black	7	0	3
Model Version 5	White	10	0	0
	Black	0	0	10
Model Version 15	White	0	0	10
	Black	0	0	10

Figure 6:Baseline Random vs Greedy vs Neural Net

Chess Layout	Baseline	Color	Won	Lost	Draw
BabyChess	Random	White	10	0	0
		Black	9	1	0
	Greedy	White	10	0	0
		Black	0	10	0
Mallot	Random	White	10	0	0
		Black	9	1	0
	Greedy	White	10	0	0
		Black	10	0	0

Figure 7:Best Model on directly used on different layouts

Iteration to Reach the performance

- Though there are existing systems like Skype and Hangouts, Ping guarentees around 60% more efficiency.
- Ping also guarentees high scalability. Around 3000 users per second.
- As WebRTC is leveraging communication, Ping uses WebRTC to improved end user experience.

Interesting Moves and Neural Net values

- Though there are existing systems like Skype and Hangouts, Ping guarentees around 60% more efficiency.
- Ping also guarentees high scalability. Around 3000 users per second.

Conclusion

- Though there are existing systems like Skype and Hangouts, Ping guarentees around 60% more efficiency.
- Ping also guarentees high scalability. Around 3000 users per second.

References

- Gardner's Minichess Variant is solved. Mehdi Mhalla et al. *arXiv e-print (arXiv:1307.7118)*
- Learning to Play Othello Without Human Knowledge Surag Nair et al *github.com/suragnair/alpha-zero-general*
- Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm. Silver et al. 2017a

Contact Information

- Web: https://github.com/karthikselva/alpha-zero-general/tree/minichess
- Email: karthik0@stanford.edu