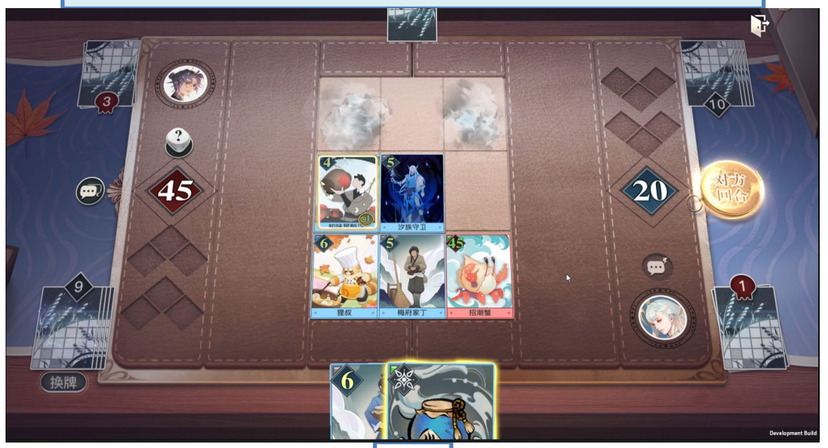
GDC

Applying AlphaZero to Develop AI in Turn-Based Card Games

Zihan Lei
Fuxi AI Lab, NetEase Inc

Outline

Revelation Mobile: Yuxi



AlphaZero

AlphaZero and MCTS

Yuxi: a turn-based card game in Revelation Mobile

Applying AlphaZero

Result and future

01

AlphaZero and MCTS

What is AlphaZero

The game of Go

- The game of Go has long been viewed as the most challenging of classic games for artificial intelligence[1]

AlphaGo[2]

- the first computer program to defeat a professional human Go player
- the first to defeat a Go world champion
- arguably the strongest Go player in history

The game of Go[3]

AlphaGo

⇩

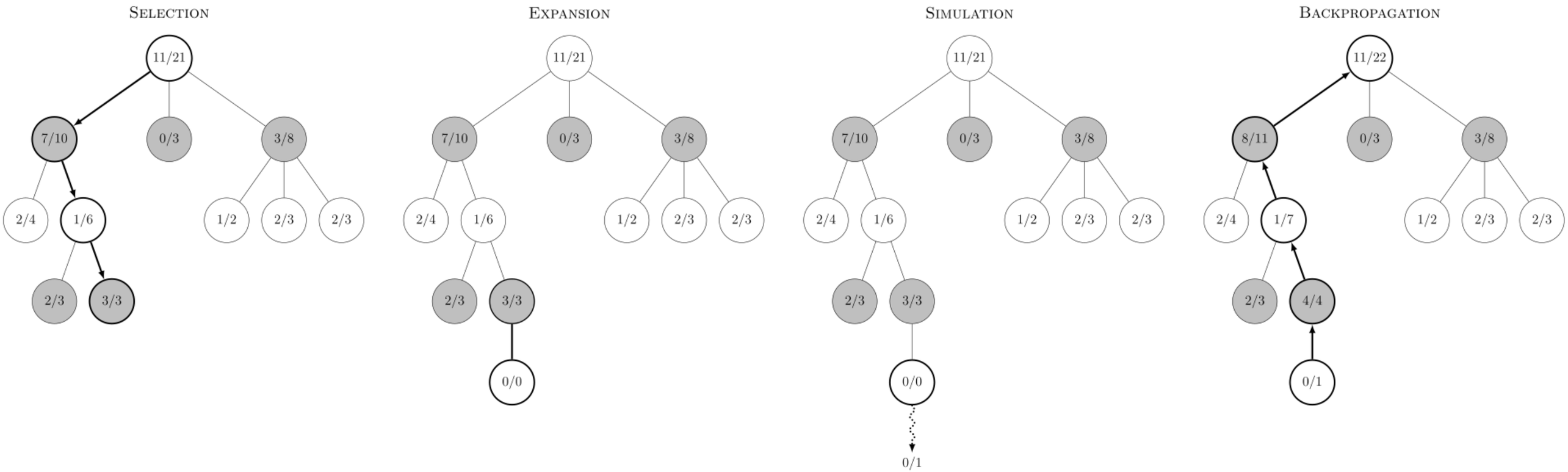
AlphaGo Zero

⇩

AlphaZero

1. https://www.nature.com/articles/nature16961
2. https://deepmind.com/research/case-studies/alphago-the-story-so-far
3. https://en.wikipedia.org/wiki/Go_(game)

Monte Carlo Tree Search

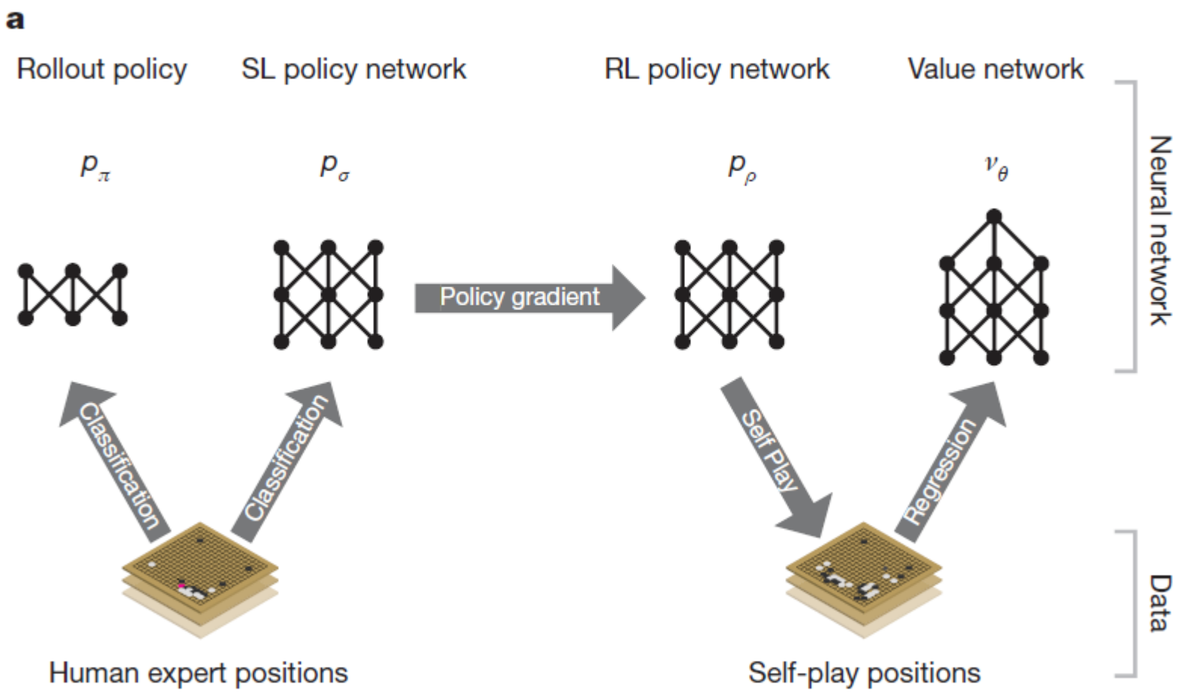


Step of Monte Carlo tree search[1]

1. https://en.wikipedia.org/wiki/Monte_Carlo_tree_search

MCTS to AlphaGo[1]

- Use deep neural networks
- Supervised learning
- Reinforcement learning
- Self-play
- Asynchronous policy and value MCTS
- Human expert data of 30 million steps
- 1202 CPUs ,176 GPUs and 48 TPUs



1. Silver, D. et al. Mastering the game of Go with deep neural networks and tree search. Nature 529, 484–489 (2016)

AlphaGo to AlphaZero

AlphaGo Zero[1]

- No supervised learning
- Smaller input scale and more complex network structure
- No rollout and new loss function
- 4 TPUs

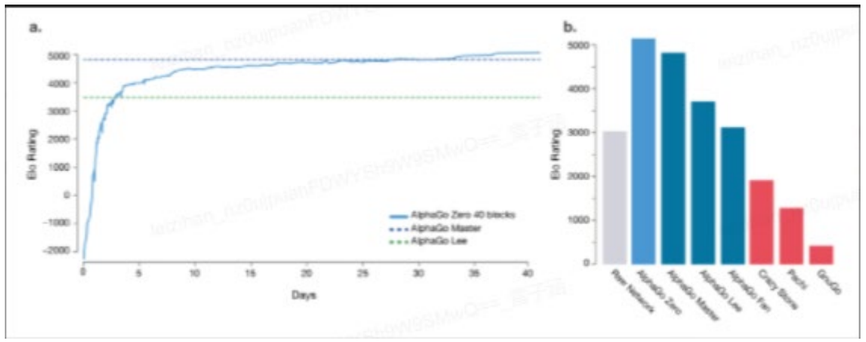
19*19*48

19*19*17

$$l = (z - v)^2 - \boldsymbol{\pi}^\top \log \mathbf{p} + c||\theta||^2$$

AlphaZero[2]

- Apply AlphaGo Zero to board games other than Go



1. Silver, D., Schrittwieser, J., Simonyan, K. et al. Mastering the game of Go without human knowledge. Nature 550, 354–359 (2017).
2. Julian Schrittwieser,. et al: Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model. (2019)

02

Yuxi

Revelation Mobile

Revelation Mobile is a MMORPG developed by Netease games. It is a very successful game in China since it was launched on January 8th, 2021.

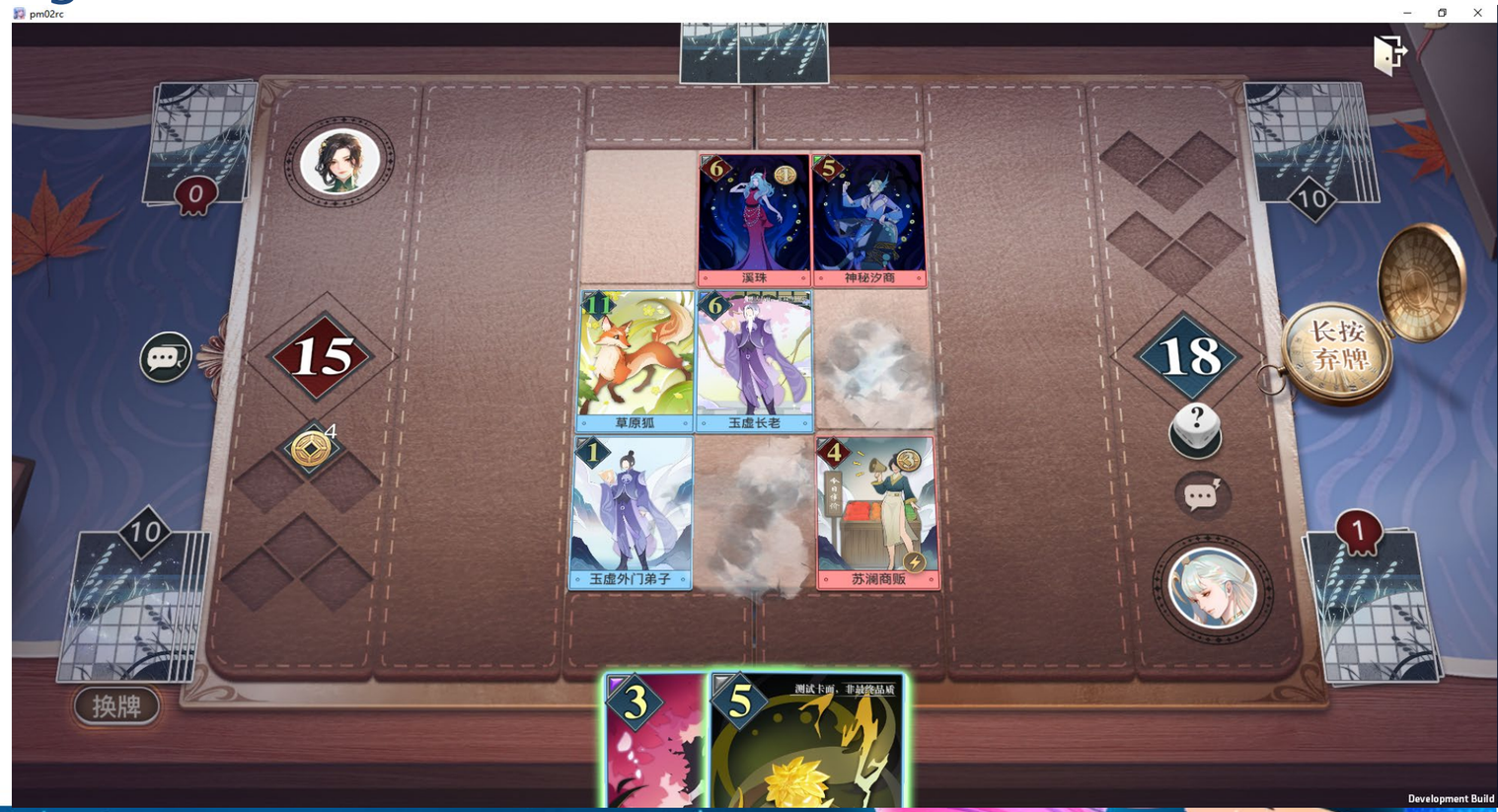
Revelation Mobile: Yuxi

Gameplay

- A turn-based card game in the Revelation Mobile, the purpose is to familiarize the player with the world view background and relax their mood.

Scene

- 1v1
- Three different board sizes: 3*3, 4*4, 6*4
- More than 200 cards, including role cards and spell cards
- No matches and leaderboards




pm02rc
enemy hand
enemy grave
3*3 board
my point and money
enemy deck
enemy point and money
wea ther
end turn
card
my deck
my grave
card exchange
my hand
溪珠
神秘汐商
草原狐
玉虚长老
玉虚外门弟子
苏澜商贩
测试卡面，非最终品质
Development Build
GDC GAME DEVELOPERS CONFERENCE | July 19-23, 2021 | #GDC21

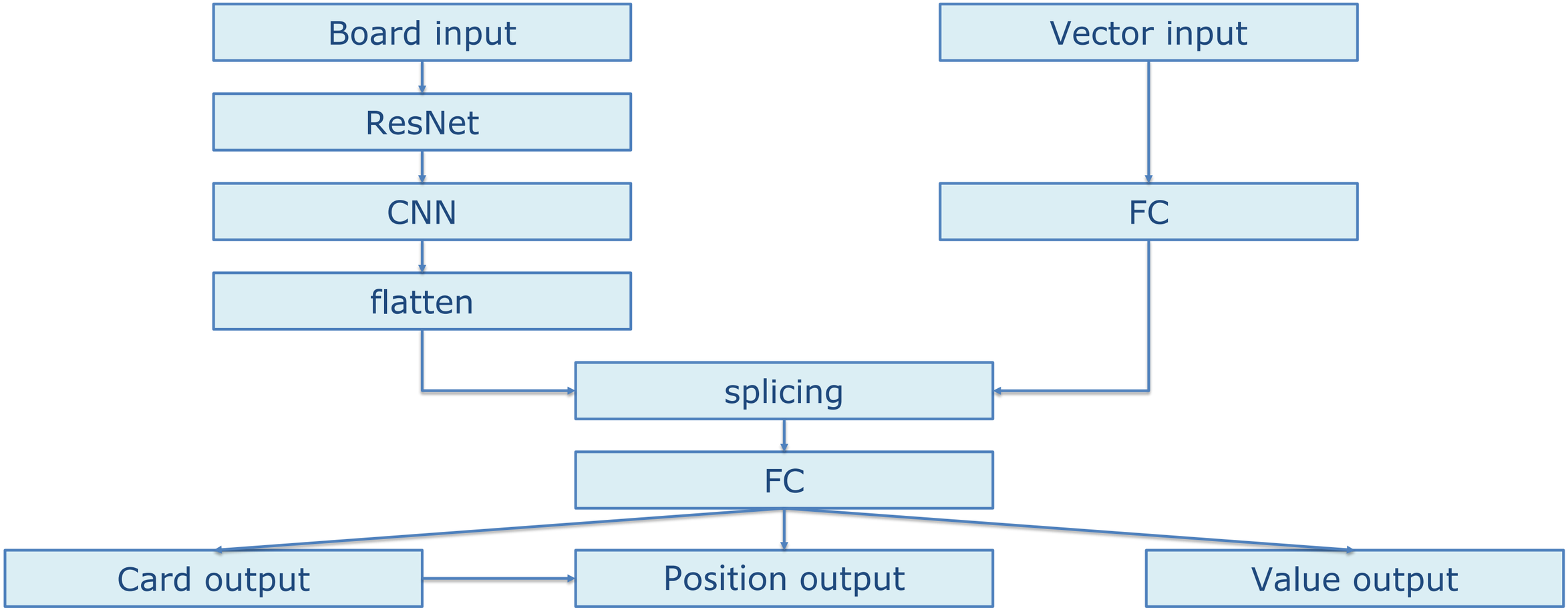
role card

spell card

03

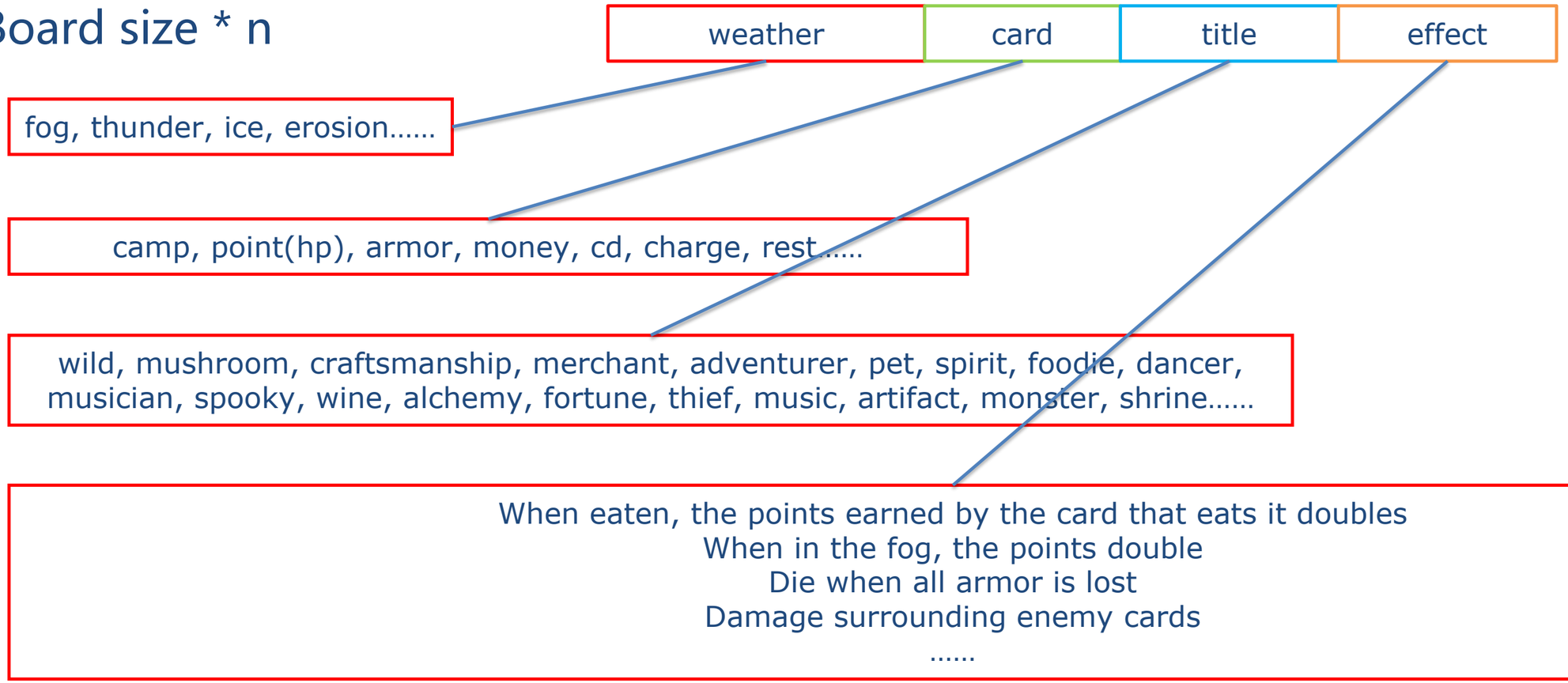
Applying AlphaZero

Network structure



Board input

Board size * n



Vector input

Use the card ID as the index and the number of cards as the value

4 parts

- my hand → [1, 2, 1, 0, 0, 1, 0, 0, 1, 0]
- my deck
- my grave
- enemy grave

There are 2 cards with ID 1 and 1 card with ID 0, 2, 5, and 8 in my hand

Because we can’t see the enemy’s hand and deck, we can’t add them to the input

Output

We have a legal mask to ensure that the actions output by the model are legal

- Step1: Choose a hand card or end turn
- Step2: Choose a position

Loop until the end of the round

Some cards need to choose 2 or more positions, we use rules to achieve these subsequent choices

The model will also output a value to estimate the quality of the scene

Train

Method

- more than 50 fixed decks
- each model controls a deck
- let these decks play against each other for training
- no card exchange

Resources

- The first time: 100 CPUs and 8 GPUs, 1 month
- Iteration: 50 CPUs and 4 GPUs, 1 week
- Online: 4 CPUs

How to search?

Unlike complete information games like Go and Chess, how to search without knowing the enemy's hand?

- Training: Use the enemy's model
- Online: Only search card exchange

A game environment that can generate and load snapshots.

Once in each round, we can exchange a hand card with one of top two cards in the deck

Online

Card exchange

- Traverse all possible exchanges
- Input these states to the model
- Find max value output
- Do card exchange!

Robustness

- Real players may use decks or cards that the AI has not seen in training?
- The modeling method can identify these cards in board input, but cannot guess the deck of real players.

04

Result and future

退出房间

房间号：104921295

对战

53 韩子惠

55 lzh006

3*3快速局\苏澜郡的雾霭

随机套牌

点击这里输入信息

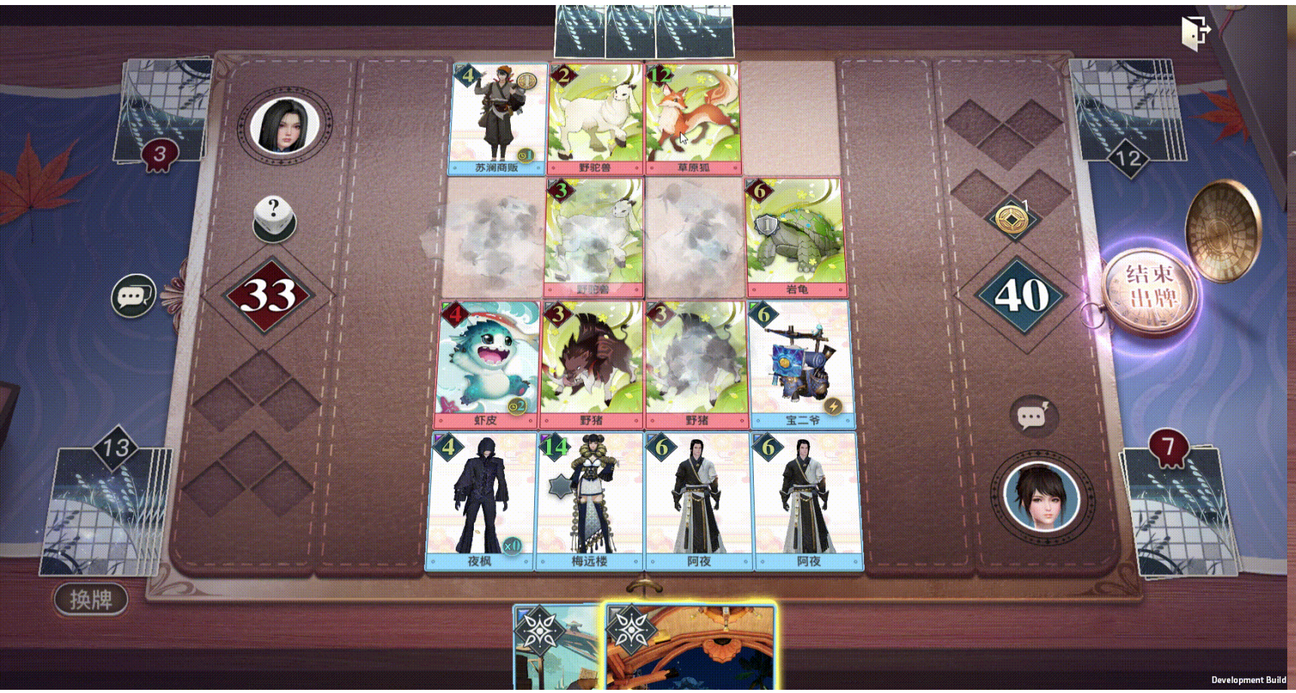
发送

准备

Development Build

BUG

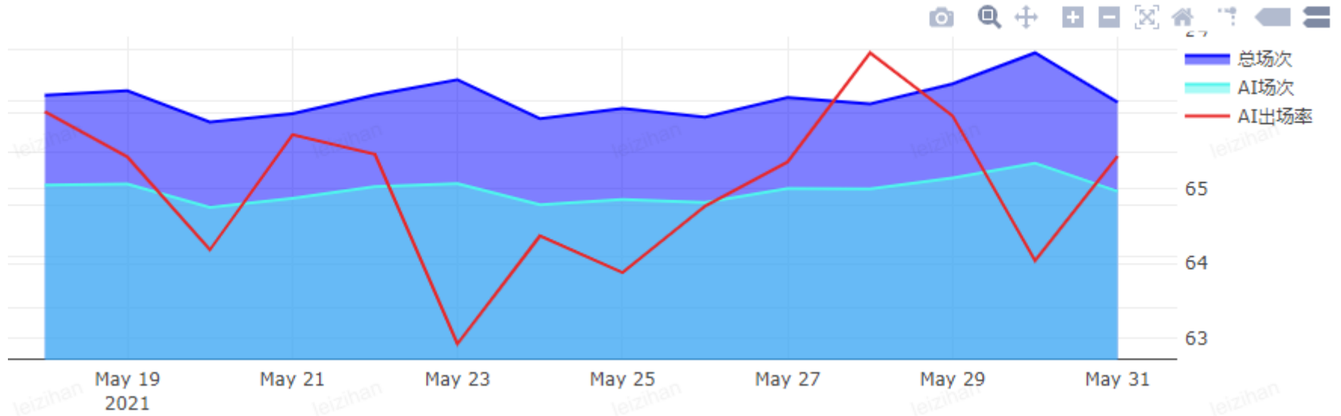
During the training process, AI can not only find bugs that will report errors, but also find bugs in design



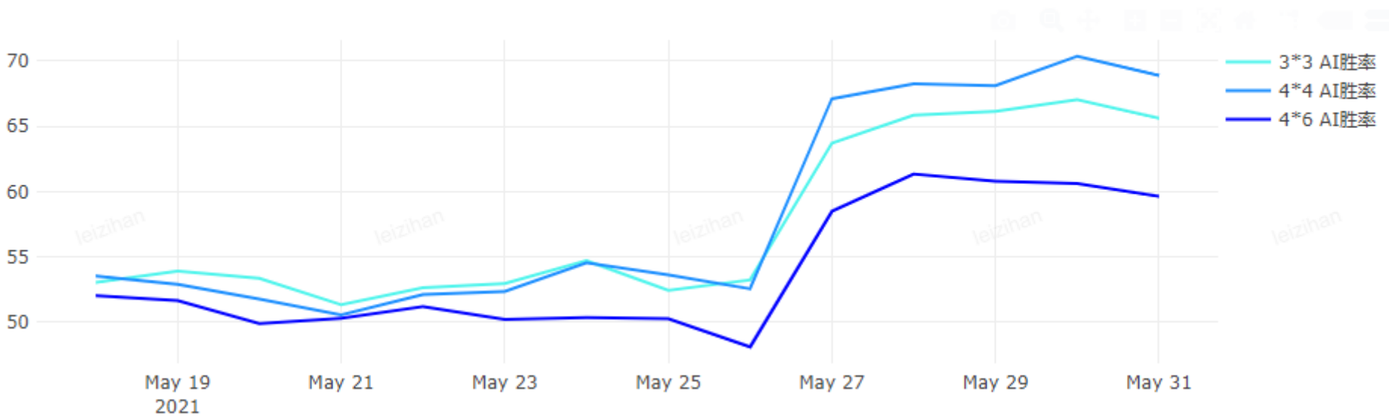
Balance

At the end of the game, the first player will randomly get a bonus point between 1 and 6.
After AI training, We found that the first players have an advantage in the 3*3 aboard and do not need the final bonus. Instead, the second players need.
Later, when we create a Yuxi room in the game, we can set a rule to remove this bonus.

Online data



RL-PM02-谕戏-胜率-分场景



Future Work

- Enable AI to predict the player's hand and deck
- As the cards are updated, the model also needs to be updated manually, otherwise it cannot adapt to the new cards. We will make it automatic in the future.
- Apply AlphaZero to other turn-based games such as Mahjong and Gobang.
- ……