Reinforcement Learning in Collectible Card Games: Preliminary Results on Legends of Code and Magic

Ronaldo Vieira[1], Luiz Chaimowicz[1], Anderson Rocha Tavares[2]

[1] Universidade Federal de Minas Gerais, Departmento de Ciência da Computação, Belo Horizonte, Brazil

[2] Universidade Federal do Rio Grande do Sul, Instituto de Informática, Porto Alegre, Brazil

{ronaldo.vieira, chaimo}@dcc.ufmg.br, artavares@inf.ufrgs.br

1 Motivation

In the recent years, games such as Go and poker have been played at superhuman level, but performance in collectible card games remain limited.

Can we play well a collectible card game with a pure reinforcement learning approach?

2 Legends of Code and Magic

Legends of Code and Magic (LOCM) is a two-player digital collectible card game in the likes of the popular Hearthstone, except with simpler rules and designed specifically for research.

A match of LOCM has two phases:

- In the **draft** phase, the players build their decks by choosing a card between three random cards presented by the game, 30 times.

- In the **battle** phase, the players take turns playing cards from their hands and attacking their opponent until one of them has zero or fewer health points.

3 Methodology

In our initial plan, each phase will be played by a dedicated **(deep) neural network** that learns by self-play with a **state-of-the-art reinforcement learning algorithm**, using the match result (win or loss) as reward signal.

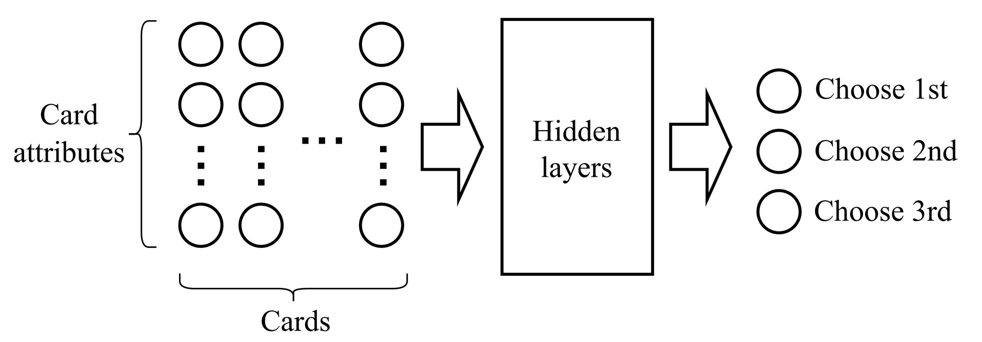

Figure 1: Expected input/output for the draft neural network. It takes the attributes of the cards and outputs the chosen card.

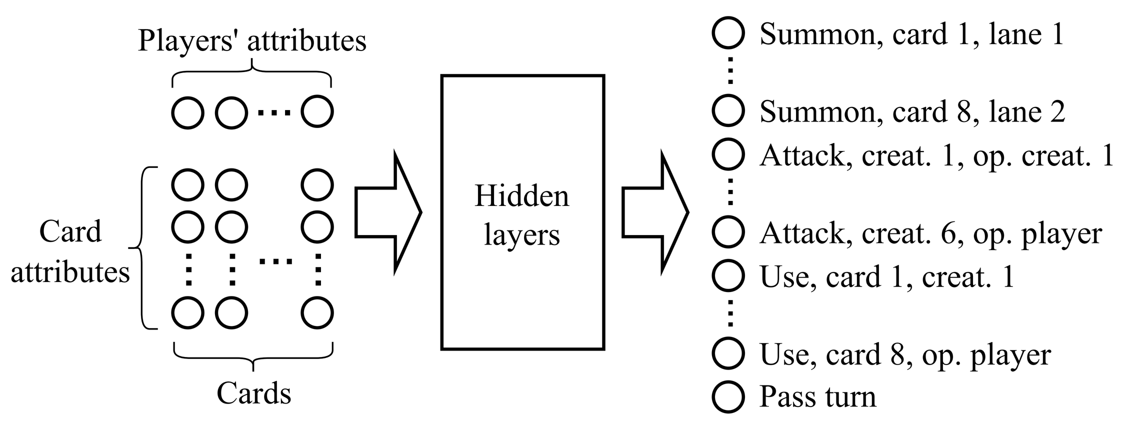

Figure 2: Expected input/output for the battle neural network. It takes the players' stats and attributes of the cards in hand and in the board and outputs the chosen action.

4 Preliminary & Expected Results

- To facilitate further research on LOCM, we **reimplemented the game engine** and developed draft-only, battle-only and full game **OpenAI Gym environments**.
- We studied the influence of each game phase on the outcome of a match, and confirmed that **the strategies used on the draft and battle phases are interdependent and both contribute significantly to a win**.

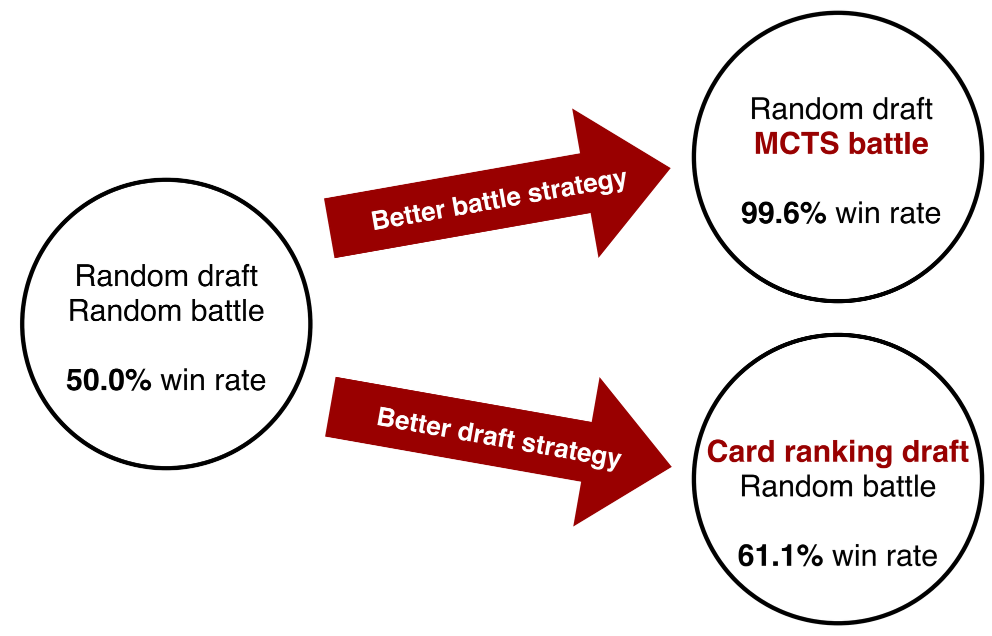

Figure 3: Win rate of strategies against a baseline (random draft and random battle).

- We plan to evaluate our approaches for both phases individually and jointly against state-of-the-art strategies, with win rate as main metric.
- By training both draft and battle models simultaneously, we expect them to optimize themselves for one another and **perform better than the best LOCM agents currently**.

5 Conclusion

With this work, we expect to show that pure reinforcement learning approaches are viable in the domain of collectible card games.

6 Future Work

Some future work ideas include:

- Addressing the uncertainty in the game by, for example, modeling it as an partially observable Markov decision process (POMDP).
- Porting this methodology to a more complex collectible card game.
- Studying the synergy between different draft and battle strategies.
- Using the battle phase network as a tree policy for a Monte Carlo tree search method.

This work is supported by: