Bluff Bot

04/03/2025

Alec Yu
Jason You
Arshan Ghanbari

The Reinforced Learners

1. **Defining the Problem**
2. Exploring the Problem through a Game
3. Obtaining a Baseline
4. Comments on Baseline
5. Approach + Results + Conclusion

Deep reinforcement learning (DRL) has achieved remarkable success in perfect information games (i.e., AlphaZero, Chess, 2017), yet applying these techniques to environments with hidden/incomplete information and strategic deception remains an open challenge.

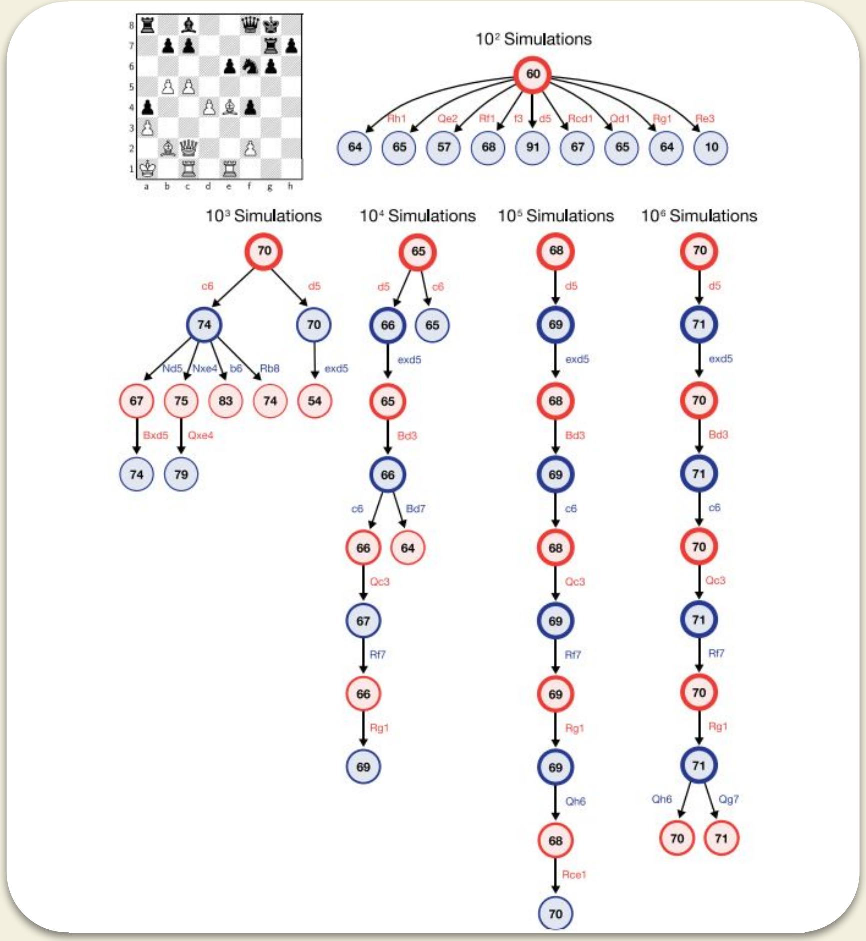
Information Types

Complete

Everything about the game is known (players, actions, sequences, etc.)

Incomplete

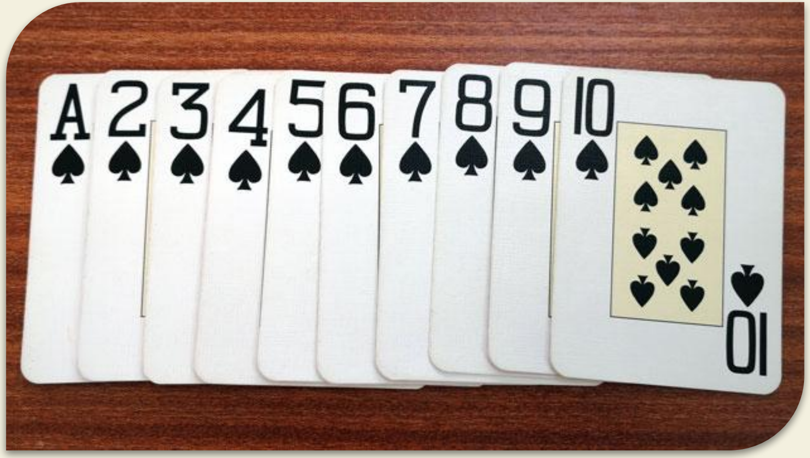
Lacking knowledge about some relevant aspect of the game (player playing habits)

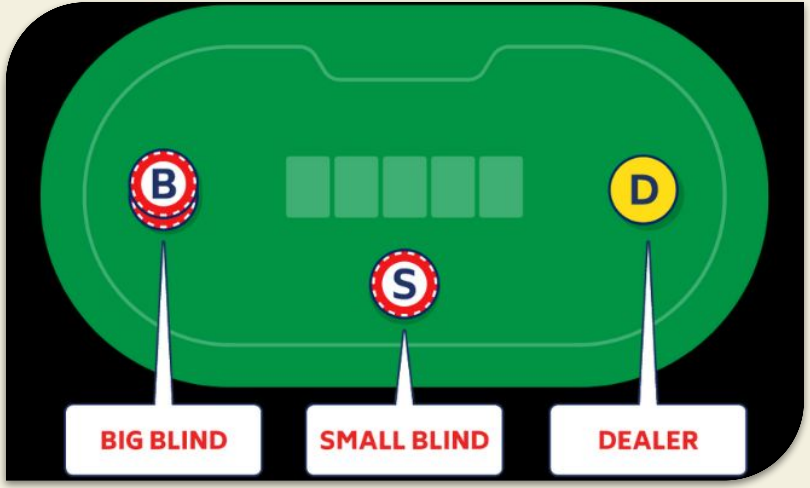


1. Defining the Problem
2. **Exploring the Problem through a Game**
3. Obtaining a Baseline
4. Comments on Baseline
5. Approach + Results + Conclusion

Bluff Game

Simple adaptations to traditional Texas Hold'em to emphasize the importance of learning incomplete information

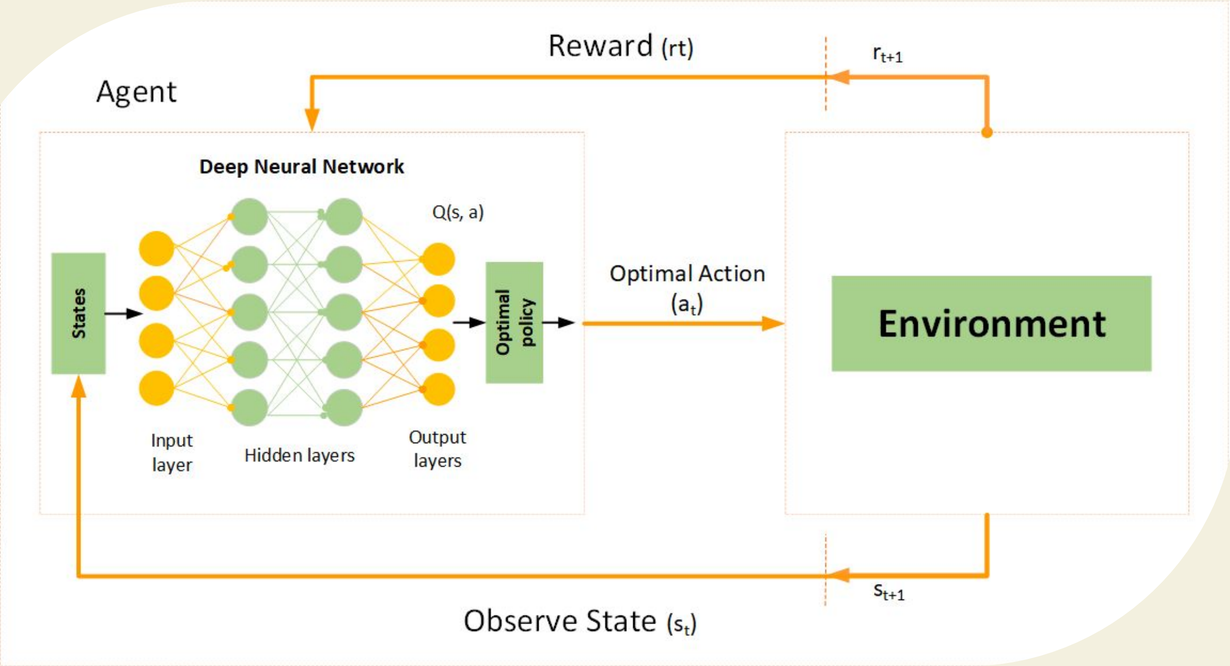




Information Type	Details
Complete Information	Player Card/Value Each Player's Chip Bet for the Round
Incomplete Information	Adversarial Card/Value Adversarial Betting Patterns Adversarial Bluffing Patterns

1. Defining the Problem
2. Exploring the Problem through a Game
3. **Obtaining a Baseline**
4. Comments on Baseline
5. Approach + Results + Conclusion

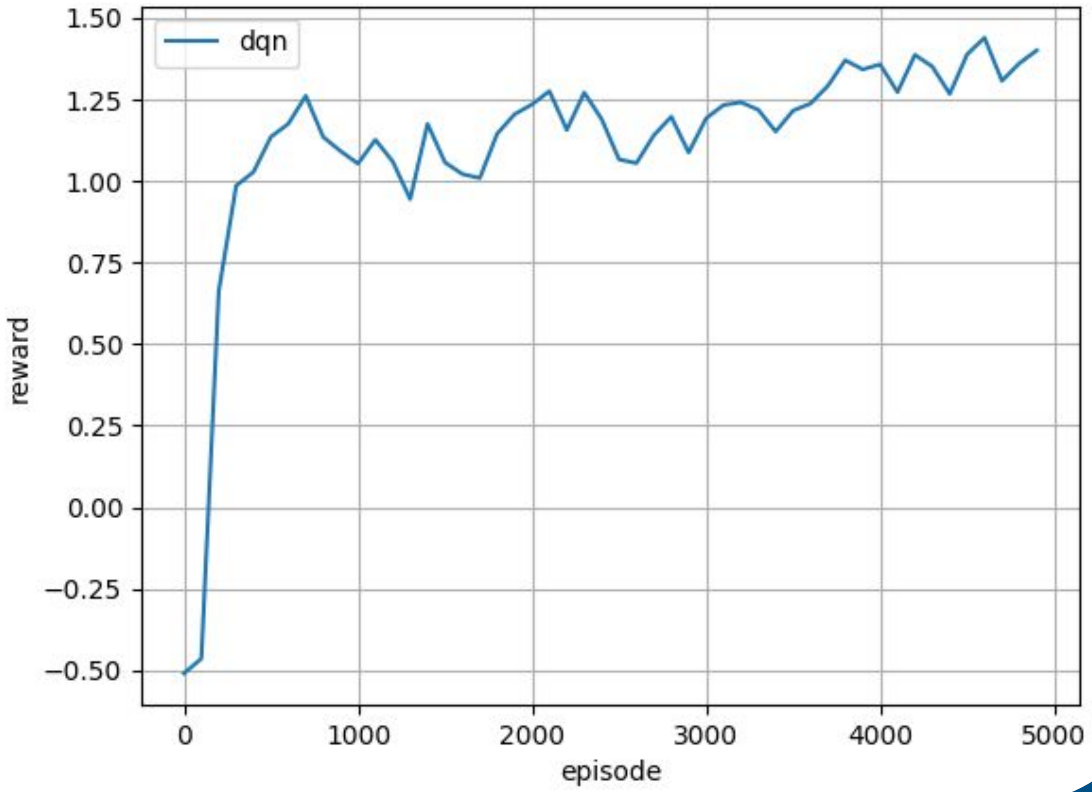
Deep Q-Learning Network



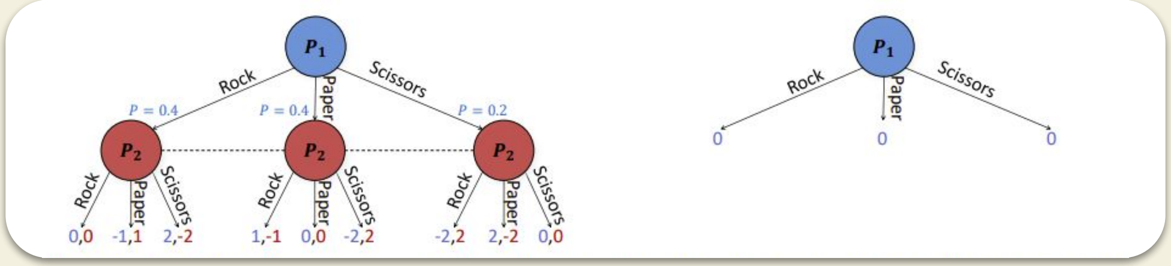
Baseline Outcome

Rewards are measured in big blinds per hand

0.5 (-0.5) indicates that the player wins (loses) 0.5 times the big blind amount



1. Defining the Problem
2. Exploring the Problem through a Game
3. Obtaining a Baseline
4. **Comments on Baseline**
5. Approach + Results + Conclusion

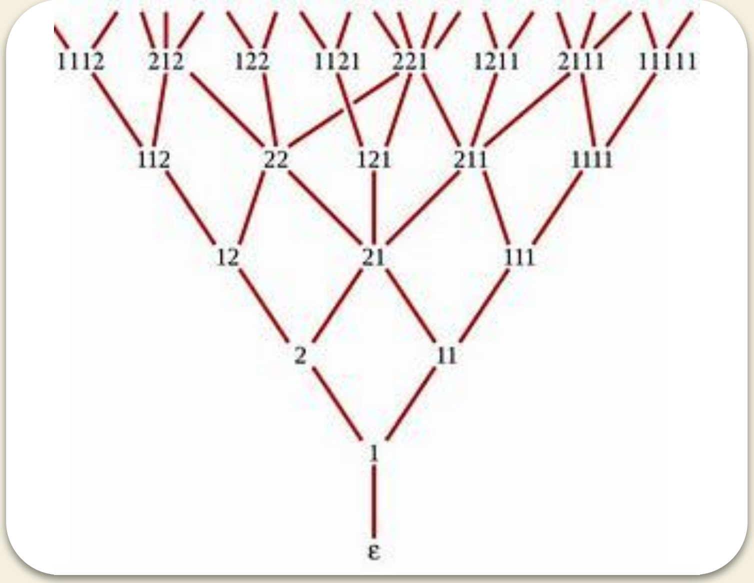


Noam Brown and Tuomas Sandholm. Depth-limited solving for imperfect-information games. Advances in Neural Information Processing Systems (NeurIPS), 2018.

Areas of Improvement

1. Lacks probabilistic action selection
   → no mixed strategies

2. RL generally inefficient
   → amplified by combinatorial explosion



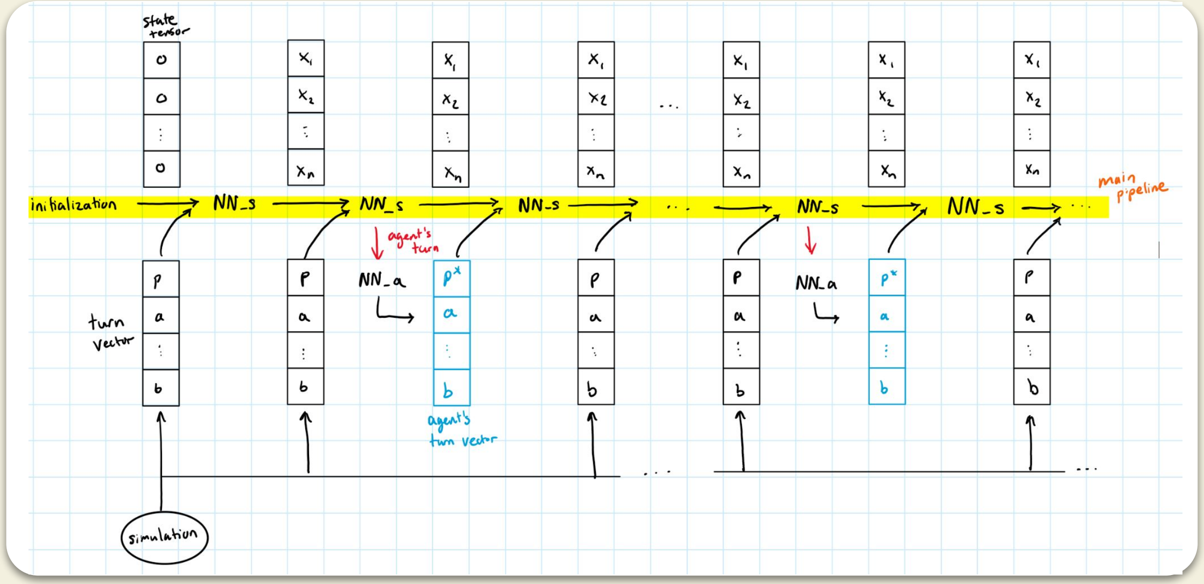
1. Defining the Problem
2. Exploring the Problem through a Game
3. Obtaining a Baseline
4. Comments on Baseline
5. **Approach + Results + Conclusion**

Our Approach

Inductive Biases

Use VAE+LSTM

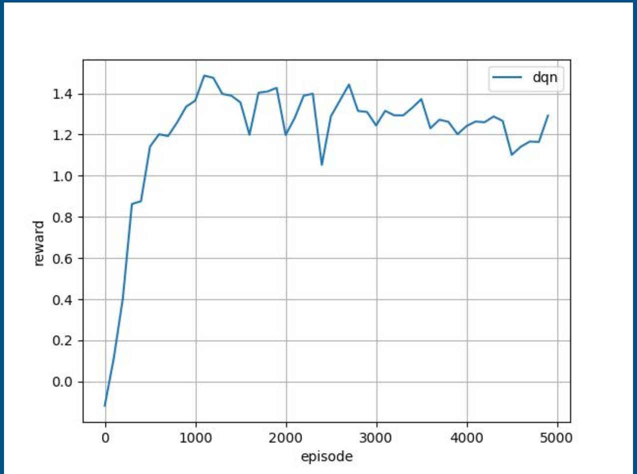
Intermediate Reward Signals (akin to custom loss function)

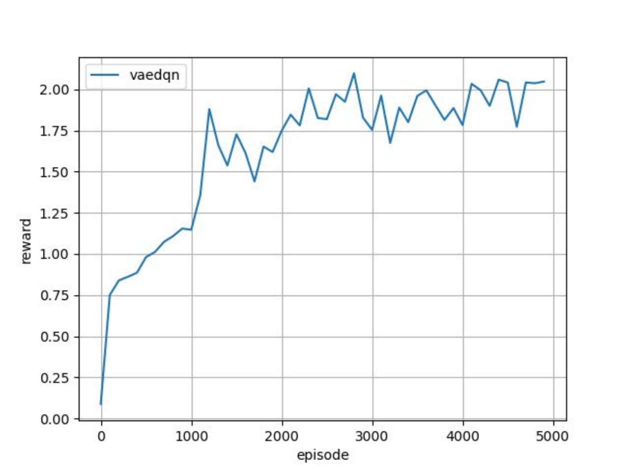


Model Outcomes

Currently, when we increase game length, model converges a bit slower but gets better performance.

Still haven’t fully implemented the model freezing and intermediate reward signal





Reflection

Focus on the problem, learn the problem. That's the work of an ML engineer and data science.