Using Reinforcement Learning to Develop a Better Strategy for Ramsey Games

Kandasamy Chokkalingam

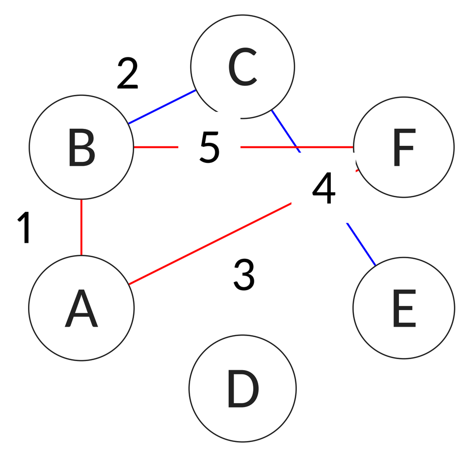
Poolesville High School

Ramsey Theory

Ramsey's Theorem states that for all m there exists an edge coloring of a graph of size n where there is a fully connected group of m nodes. [1] Ramsey Numbers are the minimum graph size, n, needed to ensure a monochromatic K_m exists. Currently, proofs are hard to find as bounds from current methods are not tight. [2]

Ramsey games involve two players coloring the edges of a graph to get a monochromatic K_m in their color, at which point they win.

A completed Ramsey Game



With Ramsey's Theorem, we know that only on graphs where n is greater than $R(m)$ will there always be a winner.

Reinforcement Learning

Reinforcement learning models learn from playing. An ideal policy will always produce the best move.

- MCTS
  - MCTS uses a random sampling of possible playouts.
  - It chooses the move that most often leads to a win.
- Tabular Q-Learning
  - Stores each value in a table
  - Precise, but takes large amounts of memory
- Deep Q-Learning
  - Trains a neural network to predict values
  - Not as precise, but is memory efficient
- Graph Q-Learning [3]
  - Uses GNNs to extract information about groups
  - Similar to DQN in performance

Methods

Models trained against another model of the same type. Models were tested on graphs of size $R(3)$ and $R(4)$ and trained for 30,000 games, or at a maximum of 6 hours.

After every game, the number of moves, winner, mean time taken to make a move, and memory used were collected.

Each model played against a random agent 3 times for 500 games each.

The policy of the top models for both $R(3)$ and $R(4)$ was studied.

Results

GQN models performed the best on $R(3)$ and DQN did so on $R(4)$. MCTS showed high performance on $R(3)$ only.

Metrics

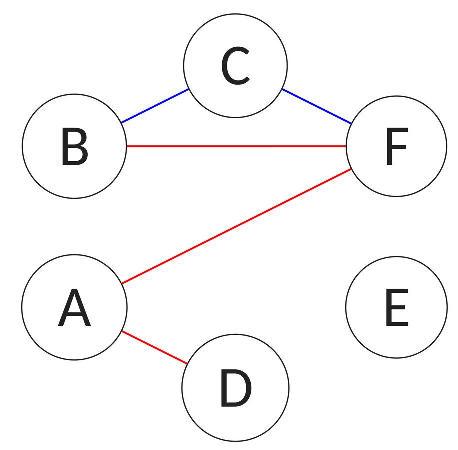
Metric	$R(3)$	$R(4)$
Win Rate		
Fastest Growing	GQN-3	DQN
Highest Value	TQL	DQN
Average Move Time		
Lowest Value	TQL	DQN
Highest Value	GQN-1	MCTS
Number of Moves		
Lowest Value	TQL	MCTS
Highest Value	DQN	GQN-1
Memory Usage		
Lowest Value	GQN-1	DQN
Highest Value	TQL	TQL
Most Consistent	GQN-3	GQN-1

For the random player win rate metric on $R(3)$, MCTS had the highest for the player and opponent. The TQL and GQNs are equal and better than DQN.

On $R(4)$, DQN had the highest for the player and opponent. Both Player and Opponent DQN is better than all the other models. The other models cannot be proved to be not equal.

Results

An Example of Forking



The MCTS Player on $R(3)$ took 8 moves to win. The MCTS Player prioritized forking from the same node, connecting multiple nodes to the same node for multiple moves.

The DQN Player on $R(4)$ took 65 moves to win. Both DQN agents did not prioritize forking as much as MCTS agents did. More clusters of connections were present, with multiple separate forks from different nodes.

Conclusions

GQN models performed the best on $R(3)$ and DQN on $R(4)$.

MCTS showed high performance on $R(3)$ but not on $R(4)$.

Forking was a strong strategy and done differently on $R(3)$ vs. $R(4)$.

This method is viable for learning how to play Ramsey games.

Future Work

Applying models to higher Ramsey number sized graphs.

Evaluating the performance of newer GNNs that may be more suited for this task.

References

[1] W. Gasarch. *Ramsey's Theorem on Graphs*. en. URL: `https://www.cs.umd.edu/users/gasarch/COURSES/858/S13/ramsey.pdf`.

[2] W. Gasarch. *Small Ramsey Numbers, Lower Bounds On Ramsey Numbers: An Exposition by William Gasarch [Lecture*. en.

[3] Jiechuan Jiang et al. "Graph Convolutional Reinforcement Learning". en. In: *arXiv:1810.09202 [cs, stat]* (Feb. 2020). arXiv: 1810.09202. URL: `http://arxiv.org/abs/1810.09202` (visited on 07/06/2020).