Ryan Perrizo
Artificial Intelligence Nanodegree: Research Review
https://people.csail.mit.edu/rivest/pubs/Riv87c.pdf

Research Review Summary: Game Tree Searching by Min/Max Approximation

The paper addresses two apparent weaknesses when considering a tree search using minimax with alpha-beta pruning. One problem is the use of depth-first search, which decides which child node to explore next in an arbitrary fashion. The second problem is using a minimum or maximum function on a list of child scores, which only considers the top-choice and disregards strong alternative choices. The main topic discussed is penalty-based approaches using min/max approximation functions and their derivatives.

Tree searching by min/max approximation is an example of a penalty-based iterative search method, which assign a penalty to the edge between each parent and child node. The lowest sum of penalties between the root of a tree and the tip nodes determines which node is to be expanded. Min/max approximation substitutes the traditional minimum/maximum calculation with the generalized p-mean function. The derivative of the p-mean function can be used to calculate the 'sensitivity' between two nodes, while the derivative of the exact min/max function provides no additional information. The sensitivity score is then used to assign weights to edges.

Implementing generalized p-means is computationally expensive. The implementation described in the paper does not use generalized mean values and instead follows a "reverse approximation" approach. The traditional min and max calculations are used. The main point of using generalized means was for using the derivatives to assign sensitivity weights between each parent node and child node. They achieved this effect with a weight equation (Eq. 15), which provides a weight to every configuration based on the configuration value (calculated by an evaluation function for the configuration state) and the most favorable descendent value. The function provides a variable for penalizing depth and another for adjusting based on confidence in evaluation function accuracy. Using this implementation, they were able to test the performance of a penalty-based scheme against minimax with alpha-beta pruning.

Connect Four was chosen as the game to compare the penalty-based approach to minimax with alpha-beta pruning. The evaluation function scored each set of four spaces that would end the game if filled by one player. The score was zero if no tokens were in the segment or if both players had a token in the segment. The score also took into account which player was up next.

For evaluating performance, the two tree search approaches were constricted on either time or number of calls to 'move'. Because the penalty based approach requires more computation time, alpha-beta was able to explore more configurations and outperform when the restriction was on time. When the restriction was on calls to 'move', the penalty based approach outperformed alpha-beta since each move takes more information into account.

Min/max approximation and other penalty-based approaches can outperform minimax with alpha-beta pruning if limited to the same number of moves. If the constraint is on time and not on moves, alpha-beta can explore more configurations and outperform the penalty based approach.