Review of paper "Game Tree Searching by Min/Max approximation" by Ronald L. Rivest

The author proposes a technique for searching game trees. His idea is based on minimax algorithm with iterative heuristic.

Minimax algorithm with alpha-beta pruning expands all nodes at each depth level pruning those that won't be selected. When using iterative heuristic we expand one node at a time which can be at any depth level in the tree - thus forming a tree of non-uniform depth. The idea of iterative heuristic is to focus on expanding the node which has the most impact on the root value and thus the most promising path. The way we choose the node to expand is determined by the penalty function - "good" nodes has lower penalty than "bad" nodes. The node with the lowest penalty is called the tip. The author proposes a new way to define the penalty function based on min/max approximation.

Generalized mean values theory states that the maximum of a set of values is a p-mean when p approches the infinity and the minimum of a set of values is a p-mean when p approches the minus infinity. A large value of p (e.g. 10) will give a good approximation of min/max values. The difference between the usual min/max function and the one proposed in the paper is that the latter has continuous derivative and that fact the author uses to define the penalty function.

In order to cope with computational difficulty of computing the generalized p-means the author proposes to skip the computation altogether and to use appropriate min/max values instead which must be very close to the generalized p-means. This is called "reverse approximation". In the end we are interested in the derivative of that function hence an approximation of its value should suffice.

The author concludes with experimental results demonstrating that his technique can give better results than minimax algorithm with alpha-beta pruning. The proposed algorithm (MM) was compared to alpha-beta pruning algorithm (AB) to play the game "Connect-Four". Both time (1 to 5 seconds) and move (1000 to 5000 moves) bounds were used to compare the algorithms. The results show that AB is superior to MM when time bounds are used. However, when resource limits are move-based the MM algorithm is definitely superior.

In the end the author discusses some important properties of penalty-based algorithms like the one proposed in the paper. Here is the summary:

- space requirements are higher than for depth-first search schemes
- orientation towards improving the heuristic value at the root and not choosing the best move
- the necessity to expand all successors of the tip
- in some cases the penalty-based schemes might be inefficient
- penalty-based schemes do spend some time evaluating non-optimal paths