Game Tree Searching by Min / Max Approximation by Ronald Rivest

Goals and techniques

Game playing by computer entails searching of futuristic winning position in game tree. It is impossible to search in game tree of high branching factor like chess due to problem of combinatorial explosion. To tackle this problem, we can use minimax with alpha-beta pruning and iterative deepening. While alpha-beta pruning optimizes minimax by pruning useless nodes, iterative deepening limits number of nodes to visit so that computer can respond within specified time. Despite these methods, we still need some heuristic to find next node to explore which have high probability to be part of solution.
In this paper, author introduces mini-max approximation operator instead of max and min operator in minmax algorithm to find heuristic evaluation value of different leaf nodes in explored part of game tree. With high value of power factor, generalized mean value tends to be maximum of its arguments. Similarly with low value of power factor, generalized mean value tends to be minimum of its arguments. This generalized mean operator has advantage that it is differentiable every where unlike max and min operators which are differentiable only at max and min value.

Paper explains policy based iterative procedure in game trees and explains a novel instance of this method as follows. While exploring a game tree, we need to decide a leaf node to expand. Here we choose a leaf node which has highest effect on the value of root i.e. differentiation of value at root with respect to a leaf node. We start with differentiation of leaf node with respect to its parent and back pass to root annotating each edge with a penalty (negative logarithm of derivatives) corresponding to differentiation value by chain rule. Now we would explore a node with minimum penalty as root has highest differentiation value with respect to this leaf. We iterate through this process selecting a new leaf node and updating heuristic value of nodes with back passed derivatives until we are left with terminal nodes or we have reached a time limit.

Experiment and Results

Author chooses commercially available and well known game of Connect Four as basis of experiments. Author tries 2 different metrics, time usage and number of moves.
While on time usage alone, alpha-beta seems to be superior of the min/max approximation approach. However, if we base our comparison on move-based resource limits, the min/max approximation is superior. Author note that the number of distinct positions considered by alpha-beta was approximately three times larger than the number of distinct positions considered by min/max when a time bound was in effect.

After explaining success of its methods, paper ends with open problems like best generalized mean value functions, parallelization and pathological games.