Game Tree Searching by Min/Max Approximation

Ronald L. Rivest MIT Cambridge MA 02139
Reviewer: Bhinesh Patel 27 April 2017

This paper introduces a new technique for searching in game trees, based on the idea of approximating the min and max operators with generalized mean-value operators. The main reason for using generalized means is that they are more suitable to sensitivity analysis than the min or max functions.

For a two person zero-sum perfect information game, a game tree can be constructed to map out all the possible move combinations. Minimax search can be then used to expand all the nodes and come up with an optimal move. Alpha-beta search with pruning will reduce the number of nodes that need to be expanded. In reality the number of nodes to be expanded is so large that only nodes up to a certain depth can be explored. A heuristic approximation is needed so that the search can be terminated within a reasonable time. Given a limit on the computing time available, one can successively compute the backed-up values for depth d=1, 2, … until one runs out of time and return the last search completed. This technique is known as iterative deepening.

A different class of heuristics are iterative heuristics which grow the search tree one step at a time. The tree grown need not be of uniform depth: some branches may be searched to a much greater depth than other branches. The heuristic proposed in this paper is an iterative heuristic technique. The key to such an idea is to only expand the nodes that are likely to lead to an optimal move. It is proposed that those nodes that have the greatest rate of change of the value at the root node are likely to contain 'interesting moves' and thus should be expanded. In order to implement this idea, it is necessary to compute the derivatives at each node. The minimax function is not differentiable so the author proposes using a generalized mean value function, p-mean, that approximates the minimax function but more importantly is differentiable and allows the derivatives or rates-of-change to be calculated at the nodes. A penalty based iterative search method is proposed where a penalty is assigned to a node based on penalizing bad moves over good moves. The minimax values are replaced with p-mean values calculated using a large value for p (p=10 for example). Similarly derivatives are calculated at the nodes using the differential for the p-mean function. These can now be used to find the tip nodes which will have the greatest influence on the value at the root node. These nodes are expanded further and the search repeated.

Experiments were carried out by the author to test the effectiveness of the new search technique. 980 games of Connect-Four were played where time was limited in one set of games and moves were limited in another set. It was found that when moves were limited, the author's technique was superior to minimax with alpha-beta pruning whereas when CPU time was limited, the additional overhead of computing led to it being not as effective.