Early retirement

van den Herik, H.J.

Published in:
ICCA Journal

Publication date:
1997

TABLE OF CONTENTS

EARLY RETIREMENT

Human geniuses are unpredictable. In the early 1960s Robert Fischer complained that Mikhail Botvinnik only played occasionally in tournaments, whereas a World Champion was supposed to show his supremacy as often as possible to the community. As soon as Fischer himself became World Champion he quit the scene, only to return some twenty years later to play the same opponent again. Maybe there was a good reason for Fischer to stop playing, maybe not. Many top sportsmen receive the proper advice of stopping with their activities as soon as they have reached their peak. A few do stop, but most of them continue nevertheless. They enjoy their sport, they feel at ease in the community, or they would like to raise their performances, showing that the last results were by no means their best.

Many fans believed that the ultimate IBM challenge was to achieve near perfection in chess, viz. playing a match against Kasparov for the title of World Champion. So far, only in the world of Checkers we have seen such a match: CHINOOK versus Tinsley. Indeed, that match had its own complications, and its success was tainted by Dr. Tinsley's forfeit.

The world of chess and computer chess have a common interest: both would like to see a third match between Kasparov and DEEP BLUE. However, IBM has decided otherwise: they followed the proper advice, and stopped. They have won a good reputation and can now sleep at ease. Although it is not our choice, we have to respect their decision and are still grateful for what they have offered to computer chess.

Meanwhile DEEP BLUE's victory over Kasparov has invoked many thoughts, opinions and proposals for future research. In three workshops (Princeton New Jersey, Providence Rhode Island, and Nagoya Japan) the question arose which game should take the prime research position over from chess? There were pros and cons for Go, Shogi and Chinese Chess. But still, Chess itself was also in the running. My opinion is that fundamental computer-chess research has not finished with the 1997 Kasparov-DEEP BLUE match. DEEP BLUE's early retirement makes with respect of W ($\nabla_W P_k = \left(\frac{\partial}{\partial w_1} P_k, \frac{\partial}{\partial w_2} P_k, \ldots, \frac{\partial}{\partial w_n} P_k\right)$, also called the gradient of $_W P_k$). The recency parameter, λ, allows for an exponential weighting with recency of predictions occurring k steps in the past. In the experiments we used a value for λ of 0.95, and a variable value of α that decreased over time, in the range 0.1 to 0.

2.1 Comparison with other Prediction-Outcome Learning Methods

Prediction-outcome learning methods are driven by the difference between prediction-outcome pairs. For example, one might make a prediction after every move and compare this prediction with the actual outcome of the game. The resulting error term can then be used to make adjustments to the prediction. Obviously, this method can only be applied once the result of the game is known.

In contrast, the TD learning method is driven by an error term generated by the comparison of successive predictions, and need not wait for the actual outcome of the game. Sutton (1988) shows that TD learning methods make more efficient use of their experience than conventional prediction-learning methods. They converge faster and produce more accurate predictions. In addition he shows that they are easier to compute, since TD learning methods are incremental instead of waiting for a final outcome.

2.2 Converting Evaluation Scores into Prediction Probabilities

In order to use temporal difference methods, the evaluation score for the position after a move is regarded as a prediction of the final outcome of the game. [To be precise, it is the evaluation score from the position at the end of the principal variation (the *principal position*) which is backed up to the root of the search tree, and used as a prediction to be compared with future values.] The values returned by the search are converted from position value (we use a relative material balance) into an estimation of the probability of winning. For the conversion we use a standard sigmoid squashing function. Thus the probability of winning for a given position is taken to be $P_{win(v)}$, denoted by

$$P(v) = \frac{1}{1+e^{v}} \tag{2}$$

where v is the evaluation-function value of the position. For v we use the following formula:

$$v = \sum_{piecetypes} w_{type} c_{type} \tag{3}$$

where *piecetypes* range over five types, viz. Pawn, Knight, Bishop, Rook, and Queen; w_{type} is the weight of a piecetype; c_{type} is the count of the pieces of a type of the player-to-move minus the opponent's pieces of that type.

The squashing function used has the advantage that it has a simple derivative:

$$\frac{dP}{dv} = P(1-P) \tag{4}$$

Hence for each type we have for each step k in the past, the following formula (cf. (1)):

$$\nabla_w P_k = c_{type} P_k (1 - P_k) \tag{5}$$

Figure 1 shows the relation between position value and prediction probability. The example score of two Pawns (using the Trial-A value (cf. Section 5) of Pawn = 0.471) is converted into a probability of winning of 0.72. Of course, the resulting probabilities for any given material advantage will vary according to the piece values that have been learnt. For instance, when using the value from Trial E, the two-pawn advantage converts into a probability of winning of 0.74.