DISCRETE PATTERN RECOGNITION IN THE PATTERN EXPERT MODULE OF MOYO GO STUDIO

Frank de Groot

An article for the 4th International Conference on Baduk
at the Baduk department of the Myong-Ji University in Yong-In, Korea.

www.moyogo.com

2006

Abstract

Pattern recognition in Baduk generally follows the “fuzzy” approach or the “discrete” approach. Examples of the former include neural networks and the latter approach includes rule-based expert systems and pattern matching. This article presents a novel method of pattern recognition applied to Baduk: Using a very large database of pre-computed “statistical move likelihood” patterns.

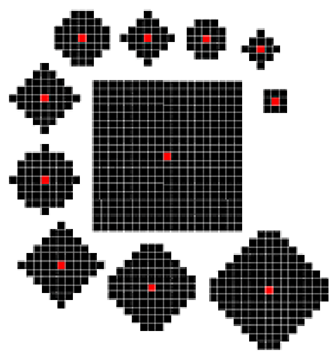
These patterns include both local and global context data, and their sole use results in an average move prediction of 40%, when all moves in contemporary pro games are guessed.

Introduction

Baduk is generally assumed to be a game where patterns play a major role. Standardized full-board and especially corner openings are an obvious example, but “good shape” moves and many obvious tactical moves are also typical “pattern moves”.

We postulated that with a large enough pattern database, and patterns with sufficient embedded information, it would be possible to play like a pro, as far as pure pattern-moves are concerned. Our postulation proved to be of merit, using a database of 16,777,216 rotational-, mirror- and color-invariant patterns. The patterns used included:

1. Twelve pattern sizes, from very small to the entire board,



2. The exact stone configuration inside the physical pattern perimeter,

3. The exact distance to the edges of the board,

4. The Ko-status of the game,

5. The number of stones in all connected strings of stones adjacent to the patterns’ center point,

6. The number of liberties of all connected strings of stones adjacent to the patterns’ center point.

Using a training set of 500,000 high-ranking amateur games and 30,000 pro games, an average “pro-prediction” of 40% is achieved on all moves played in entire games.

Our work is inspired by the research done by David Stoutamire *(Machine Learning, Game Play and Go, 1991).* In turn, our results inspired David Stern, Ralf Herbrich and Thore Græpel of Microsoft Research in Cambridge to adopt a similar approach with similar results *(Bayesian Pattern Ranking for Move Prediction in the Game of Go, 2005)*. Microsoft made their Baduk pattern system available on their online Baduk server.

Our system is available on a DVD as the commercial Baduk software “Moyo Go Studio”.

The Learning Phase

The pattern expert system learns its patterns unsupervised from a quality-controlled sample of high-ranking amateur games and professional games. To decide which patterns are included, the criteria are:

1. How “urgent” (statistically likely) the move is, e.g. the sooner a move is played on a patterns’ center point, the higher the likelihood that the pattern will be included in the database.

2. How often does a play on the pattern (on its center point) occurs. For quality reasons, exotic patterns are discarded.

Each pattern in the database has an information record on its statistical move likelihood, or “move urgency”.

Earlier experiments with Bayesian learning have proven computationally intensive and too “fuzzy” (we require exact statistical data for each pattern), therefore the algorithm for computing the statistical move likelihood is:

```
     How often the pattern occurred in all games
------------------------------------------------------
How many turns the average player waited to play there
```

Move Predictions

Plausible moves are ranked according to their statistical move likelihood. Because of the relatively large pattern database, and due to the fact that all relevant smaller patterns are included, every Baduk position always contains several to many recognized patterns.

Strengths

1. The system matches and classifies patterns near-instantaneously. On a modern PC, a Baduk position is scanned for matching patterns in less than a millisecond.
2. The system achieves an extraordinary high pro-prediction rate (40% average).
3. The system is whole-board context aware and often plays “like a pro” in positions without much tactical complexity.

Weaknesses

1. The more complex and “hot” a tactical situation, the less well the system performs.
2. The system does not take “ladders” into account.
3. Due to the unsupervised learning of many amateur games, sub-optimal moves can sometimes be suggested.

Future work

1. The inclusion of *n*-th order liberties in the patterns. “Order liberties” are liberties of liberties, and are a measure for a string of stones to escape enclosure. (Liberties are empty board points adjacent to connected stones).
2. Using a larger pro games database.
3. Using the rank of the player as a heuristic for the reliability of the statistical move likelihood of the pattern-move.