Modelling Go Positions with Planar CRFs

Dmitry Kamenetsky, Nicol N. Schraudolph, Simon Günter, S.V.N. Vishwanathan

NICTA, Australian National University, Australia

Machine Learning and Games Workshop, December 2007
What is Go?

- Two players alternate in placing stones on the intersections of a grid
- Neighbouring stones of the same colour form a contiguous block
- A block can be *captured* if all its empty neighbours are occupied by opponent stones
What is Go?

- Two players alternate in placing stones on the intersections of a grid
- Neighbouring stones of the same colour form a contiguous block
- A block can be captured if all its empty neighbours are occupied by opponent stones
What is Go?

- Two players alternate in placing stones on the intersections of a grid
- Neighbouring stones of the same colour form a contiguous block
- A block can be captured if all its empty neighbours are occupied by opponent stones
What is Go?

- Two players alternate in placing stones on the intersections of a grid
- Neighbouring stones of the same colour form a contiguous block
- A block can be captured if all its empty neighbours are occupied by opponent stones
What is Go?

- Two players alternate in placing stones on the intersections of a grid
- Neighbouring stones of the same colour form a contiguous block
- A block can be captured if all its empty neighbours are occupied by opponent stones
Go Modelling Go

What is Go?

- Two players alternate in placing stones on the intersections of a grid
- Neighbouring stones of the same colour form a contiguous block
- A block can be captured if all its empty neighbours are occupied by opponent stones
What is Go?

- The game terminates once players agree on the life status of blocks
- The blocks and their surrounding area count towards territory
- **Territory prediction:** Given a board position predict the owner of each intersection
- Challenging problem for ML!

D. Kamenetsky, N. Schraudolph, S. Günter, S.V.N. Vishwanathan

Modelling Go Positions with Planar CRFs
Go is played on a grid graph $G$, so it is natural to model it with a graphical model such as CRF.

If we want to perform exact inference we can use the Junction Tree Algorithm ($G$ is loopy).

But... Junction Tree’s complexity is exponential in the tree-width, which is $N$ for $N \times N$ grid.
Globerson & Jaakkola 2006 presented an exact polynomial-time algorithm for computing the partition function

Restrictions:
- Graph is planar: can be drawn without crossing edges
- Binary-valued labels
- Only edge potentials, no node potentials
Our work - overview

- Faster and simpler version of Globerson and Jaakkola algorithm
- No need to compute the expanded dual of the original graph
- Showed how to compute gradients and thus perform parameter estimation
- Applied to territory prediction in Go
Graph abstraction: common fate graph

- Blocks always live or die as a unit; Grid graph $G$ does not capture this

- Common fate graph $G_f$ (Graepel et al., 2001) merges all stones of a block into a single node
Graph abstraction: block graph

- Use Manhattan distance to classify empty regions into 3 types: *black surround* (■), *neutral* (◇) and *white surround* (□)

- Collapse empty regions of $G_f$ to form the *block graph* $G_b$
Graph abstraction: block graph

- Surrounds encode the possibility for obtaining territory

- $G_b$ is more concise than $G_f$, but preserves the kind of information required for predicting territory
Graph abstraction: group graph

- **Group**: set of blocks of the same colour that share at least one surround

- Construct the *group graph* $G_g$ by collapsing groups of $G_b$
Feature engineering: nodes

- Given a node \( v \in G_b \), for each point \( i \in v \) compute the number of adjacent points \( A_i \) that are also in \( v \).

- Node’s feature is a vector \( F \), where \( F_k = |\{i : A_i = k\}| \).

- Provides a powerful summary of the region’s shape.

\[
F = \{2, 4, 2, 1\}
\]
Feature engineering: edges

- For two nodes $v_1, v_2 \in G_b$, $A^1_i$ is the number of points in $v_2$ that are adjacent to $i \in v_1$ and vice-versa for $A^2_i$.

- Edge’s features are two vectors $F^1$ and $F^2$ that are constructed using $A^1$ and $A^2$ respectively.

- Provide extra information such as the boundary shape.

$$F^1 = \{3, 3, 1\}, F^2 = \{6, 3, 0\}$$
Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.

Parameter sharing takes into account all relevant symmetries.
Experiments

- $9 \times 9$ endgame positions of van der Werf et al., 2005

**Learning**
- Use the block graph $G_b$
- Optimization with LBFGS

**Prediction**
- Use the group graph $G_g$
- Use CRFs MAP assignment as predicted label
### Results

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Vertex</td>
</tr>
<tr>
<td>Naive</td>
<td>6.79</td>
</tr>
<tr>
<td>Stern et al., 2004</td>
<td>4.77</td>
</tr>
<tr>
<td>Block graph</td>
<td>2.36</td>
</tr>
<tr>
<td>Block graph + neighbour features</td>
<td>1.87</td>
</tr>
<tr>
<td>Block graph + other enhancements</td>
<td>1.54</td>
</tr>
<tr>
<td>* GnuGo</td>
<td>-</td>
</tr>
<tr>
<td>* van der Werf et al., 2005</td>
<td>0.19</td>
</tr>
</tbody>
</table>

*: employs Go-specific features and was used to label data
First real application of Globerson and Jaakkola algorithm

2-stage graph reduction of the Go positions. The first used for learning, the later for prediction

Generic node and edge features. Parameter sharing between equivalent node and edge types
Future work

- Find better ways to classify empty regions
- Add more domain-specific knowledge
- Extend to $19 \times 19$ games
- Incorporate into a Monte-Carlo based program
"Some cause happiness wherever they GO, others whenever they GO" - Oscar Wilde