A Short Essay on Go as a Markov Process

Christof Cebulla

Abstract

In the present essay the Go game is studied in terms of a homogeneous Markov chain.

1 Introduction

The world-famous game of Go has simple rules and a complexity exceeding that of chess. In the context of Go the question arises whether it is possible to define its rules in mathematical language which could be useful for studying it. There are several known approaches, mainly focusing on finding an optimal playing strategy based on Monte Carlo tree search (cf. e.g. [Brown]). The present approach considers the time evolution of the game as a discrete stochastic process.

1.1 The state space of board configurations

Each configuration of stones on the board together with the number of captured stones, corresponds to a state in a state space denoted as E. Regarding captured stones it is sufficient to consider the difference between black and white captured stones. Since there may be - at least in principle - any number of captured stones, the state space E of possible configurations is countable infinite.

1.2 The graph of the Go game

Legal Go moves (moves allowed by the Go rules) define relations between the states in E. The moves, together with the states define a digraph (directed graph) with nodes being states of E and edges being legal moves between states. This digraph will in the following be denoted by $\mathbb{G}$. There is an infinite transition matrix $P_{\mathbb{G}}$ of the graph $\mathbb{G}$, where $(P_{\mathbb{G}})_{i,j} > 0$ if and only if there is a (single) legal move from a state $i \in E$ to another state $j \in E$. $P_{\mathbb{G}}$ corresponds to the adjacency matrix of the graph $\mathbb{G}$.

1.3 End states

Let the initial state in every game (no stones on the board) be denoted as the zero state s_0.

Definition 1.1 *A state $e \in E$ is called an end state if for at least one of the players (called player A) holds that any legal sequence of legal moves of player A while the other player (player B) always passes cannot lead to a win of player A. The set of all end states will be denoted by F, $\quad F \subset E$. The win of a player is hereby defined by the usual Go rules.*

Any state $i \in E$ can be written as a tuple $i \equiv (i_1, i_2)$, where i_1 is a board configuration from the finite set I_1 of board configurations, $i_1 \in I_1$, $|I_1| < \infty$. In fact, $|I_1| \leq 3^{N^2}$, where N^2 is the size of the board. $i_2 \in \mathbb{Z}$ is the difference between the number of black and white captured stones.

1.4 Description of a playing strategy

It directly follows from the Go game rules that every game started at the zero state ends either in an endless loop (under the assumption that there is no rule forbidding that) or in an end state. It also clear that the set F is composed of two distinct subsets which contain winning states of the one or the other player.

Definition 1.2 *A strategy of a player is defined by a matrix S with $0 \leq (S)_{i,j} \leq (P_{\mathbb{G}})_{i,j}$, $\forall i, j \in E$, and $\sum_i (S)_{i,j} = 1$, $\forall j$.*

The strategy defines the probability that the player takes a move from state i leading to next state j for every possible state j which may occur in a game.

The strategy S can be formulated by properly chosing the entries $(S)_{i,j}$ to form a transition matrix.

2 Go as a Markov Process

Let us introduce a slightly amended transition graph $\tilde{\mathbb{G}}$ which takes into consideration that a usual game just ends whenever an end state is reached. Instead we define an endless game, where after reaching an end state the game is reset and starts again:

$$\begin{cases} (\tilde{\mathbb{G}})_{e,s_0} = 1 \quad \forall \quad \text{end states} \quad e \\ \\ (\tilde{\mathbb{G}})_{i,j} = (\mathbb{G})_{i,j} \quad \text{else} \end{cases} \tag{1}$$

According to the definition of $\tilde{\mathbb{G}}$ every time an end state is reached, the system automatically returns to the zero state s_0.

Correspondingly we define a strategy for the amended model by $\tilde{S}$, where $(\tilde{S})_{e,s_0} = 1$.

In the real game, each player has his own strategy, so that $\tilde{S} = \tilde{W}\tilde{B}$, where the $\tilde{W}$ is the strategy of the white and $\tilde{B}$ the strategy of the black player, respectively. $\tilde{S}$ represents then two moves at once.

2.1 Discrete time stochastic process

Any given amended strategy $\tilde{S}$ is a transition matrix of a discrete time stochastic process $\{X_n\}_{n\geq 0}$ with $X_0 = s_0$ and $X_n \in E$, $\forall n \geq 1$. Moreover, it follows from the Go game rules that $\tilde{S}$ defines a homogeneous Markov chain, because the probability $p(x_{n+1} = j)$ that at time (or move) $n+1$ the state (or board configuration) j occurs, depends on the board configuration at time n only. As a standard result in graph theory (cf. e.g. [Har], the probability $p(x_{n+2} = i|X_n = j)$ is given by $(\tilde{S}^2)_{i,j}$.

2.2 Finitness of $E \setminus F$

Lemma 2.1 *For any $i_1 \in I_1$ there is an $i_2^+ \in \mathbb{Z}_+^0$ and an $i_2^- \in \mathbb{Z}_-^0$ such that for all $a^+ \geq i_2^+$ it holds that $i_a^+ \equiv (i_1, a^+)$ is an end state, and for all $a^- \leq i_2^-$, $i_a^- \equiv (i_1, a^-)$ is also an end state.*

Proof 2.2 *Let i_1 be a given board configuration with N_B black stones and N_W white stones on the board, $N_B, N_W \in \mathbb{N}$. The maximum number of points achievable for the black player by putting black stones on board while white only passes is $p_B^{max} = N^2 - (N_B + N_W) + 2N_W = N^2 + N_W - N_B$ (ingoring for the moment the captured stones). Accordingly, $p_W^{max} = N^2 + N_B - N_W$. Any black stone added to the board would reduce black's points by one, while all white stones are considered dead already.*

Now, set $i_2^+ = p_B^{max} + 1$ and $i_2^- = -(p_W^{max} + 1)$. Then it is apparent that (i_1, i_2^+) and (i_1, i_2^-) are end states and so are all (i_1, a^+) and (i_1, a^-) with $a^+ \geq i_2^+$ and $a^- \leq i_2^-$.

From 2.1 it follows immediately that $|E \setminus F| < \infty$. Moreover, the graph corresponding to $\tilde{S}$ has a finite component containing s_0.

2.3 Invariant Measure $\pi_{\tilde{S}}$

If $\tilde{S}$ is such that the zero state s_0 (as a node of the corresponding graph) communicates with any other state reachable from s_0 (there is a path from s_0 to that state and a path back and no paths in the graph which end in loops not containing s_0), than $\tilde{S}$ is a recurrent homogeneous Markov chain. But, for recurrent Markov Chains it is known (cf. e.g. [Bre]) that there exists an invariant measure Π (depending on $\tilde{S}$) with

$$\tilde{S}\pi_{\tilde{S}} = \pi_{\tilde{S}}. \tag{2}$$

By the ergodic theorem, $\pi_{\tilde{S}}$ can be computed by counting the occurance rate of the states during the time evolution of the infinite game defined above. Moreover, $\pi_{\tilde{S}}$ contains all the information about how good is a strategy $\tilde{W}$ for white, given any strategy $\tilde{B}$ for black (and vice versa).

References

[Brown] Brown, C. et al. *A Survey of Monte Carlo Tree Search Methods*, IEEE Transactions on Computational Intelligence, Vol. 4, No. 1, March 2012.

[Har] Harary, F. *Graph Theory*, Adison-Wesley, 1972.

[Bre] Bremaud, P. *Markov Chains, Gibbs Fields, Monte Carlo Simulation and Queues*, Springer, 1999.