762alpha アピール文書

2022 年 3 月 31 日 Ver.0.3

天野　史斎

762alpha は学習の自動化を目的とするコンピュータ将棋プログラムです。

学習の自動化のゴールとしては、コードの手直し無しで計算資源（CPU パワーとメモリー）を投入すればするほど完全解析の結果に漸近していくソフトウェアの実現にあります。

学習というとディープラーニングに代表される関数近似の理論に基礎を置く手法が今日日メジャーですが、これはレイヤーとノードの数が学習能力を規定することが明らかな割に、それらの学習対象に応じた事前見積もりが困難であり、かつ問題の複雑さが後からわかったからと言って能力をスケールさせることが簡単ではありません。この方法では上記目的には今一つです。

これに対し、762alpha では局面の順序を片っ端から暗記していく手法をとります。
、とだけ書くと種々の疑問が渦巻く方法ですが、局面の良し悪しを表現する目的からすると順序を記憶する以上に直接的で安定的な手段はなく、学習に要する手間も比較的単純に求められる（と考えています）。すなわち、N 個の局面の順序を記憶するのに必要な順序の最悪の変更回数は、途中経過を気にしないなら O(N^2)、1 回の変更でそれまでの記憶順序を壊さずに進めるなら O(N^3)と見積もられます（※1）。

ただし探索を有限の深さ（毎回読み切るわけではない）で行う場合、学習途上の暫定的な評価関数を用いることから 1 回の変更（2 局面の比較）で局面の順序を誤るエラー率 ε が 0 より大きくなることが避けられないため、この場合は

O((正しい方向に確率(1-ε)で進む酔歩運動で目的方向に 1 だけ進むのに要する試行回数の期待値)×N^3)

の手間で（概ね）押さえられる、という話になります（※2）。

残念ながら ε が大きいと多項式時間にはなりそうもありませんが、これは組み合わせ問題の厳密な学習の手間を最悪計算量で評価する限り、どんな方法を採っても避けられないかもしれません。

※1:　1 回の変更は、2 手分の探索と 2344 個の 0/1 としてベクトル表現された局面の線形分離の手間で遂行されます。

※2:　（酔歩運動で目的方向に 1 だけ進むのに要する試行回数の期待値）を知るには n 回の試行後に 1 つ隣に居る移動の数え上げ（比較的簡単）ではなく、n 回未満の試行で 1 つ隣に到達してしまう移動を除いての経路の数え上げが必要です（※3）。

※3:　x 軸上の 1 次元の酔歩運動において、x=0 から出発し、2*k ＋ 1 手かけて最初に x≧1 に到達する経路の総数はカタラン数 C_k なので、x=0 から出発して最初に x≧1 に到達するまでの手数の期待値 S は、+方向に移動する確率を p として、
Σ[k=0..無限大]{（2 ＊ k ＋ 1）＊ C_k ＊ p^(k+1）＊（1−p)^k }、ということになると思います。
この p を 1−ε と置いたものが

(正しい方向に確率(1-ε)で進む酔歩運動で目的方向に 1 だけ進むのに要する試行回数の期待値)

です。