

コンピュータ大貧民クライアント「wisteria」の製作

大渡 勝己

東京大学大学院総合文化研究科

ohto @ tanaka.ecc.u-tokyo.ac.jp

概要

著者がUECda2015向けに作成したコンピュータ大貧民大会クライアント「wisteria」について、昨年の著者のクライアント「FujiGokoro」と比較しつつ、本年度用いた手法について紹介する。

1 行動決定方法の概略

カード交換と役提出の選択は、いずれも必勝探索モンテカルロ法の順で行った。

モンテカルロ法のシミュレーション中は常に必勝探索を行い、末端では残り二人時の全読みや、三人以上で枚数が少ない際のMaxN探索を行うことで終盤評価の正確化を図った。

昨年度の「FujiGokoro」ではシミュレーション中の着手の選択にさらにシミュレーションを用いていたが、本年度は須藤ら [1] の「snowl」と同じく、softmax方策関数にて選択した着手そのままにシミュレーションを進行させた。

2 方策関数パラメータの学習

方策関数は、須藤らの用いた評価項目をベースに、著者が自身の大富豪の知識を用いて作成した評価項目を加え、合計のパラメータ数を500程度とした。

値の学習には、UECdaの過去のプログラムの試合の棋譜80,000試合程度を用いた。その内、シミュレーション中に実際に方策関数を用いて着手を選ぶ局面のみを学習対象とし、softmax方策関数に棋譜の着手が選ばれやすくなるように、方策勾配法にてパラメータを調整した。

結果、棋譜の着手との一致率は、softmax法にて温度パラメータ $T=1$ の場合には約50%、温度パラメータ $T=0$ 、つまり方策点の値最大のものを選ばせる場合には約63%となった。

試合では、方策パラメータはプレー人数分用意し、事前に学習したパラメータを初期値として、毎試合ごとに自分以外のプレイヤーの実際の着手に対して学習を行う。それによって、対戦相手に合わせたシミュレーションが行われることを期待している。

3 方策関数を用いた相手手札推定

相手手札を推定する要素として、カード交換と着手の情報を用いることが出来る。

FujiGokoroでは、ランダムに初期手札を生成してカード交換をヒューリスティックに進め、現在得ている情報と矛盾しない手札組のみ候補として残すことでカード交換の効果を反映させた。

その後、作成された手札組と、それによって定まる現在の手札組との手札得点の差分が大きい、つまり出し方に無駄が少なそうな組を、自然な着手が行われていると推定して採用するという方針を取っていた。

今年度は相手着手の効果の反映の方法を変更し、一定個数生成した手札組に対して、その手札組から当該試合の進行が生まれる尤度を方策関数によって計算し、尤度最大の組を採用するという手法を用いた。それによって手札の推定率に上昇が見られた。

また、シミュレーションと手札分配を同一の枠組みで行うことで、手札推定器を別に開発する手間が省け、さらには対戦相手の着手傾向によって手札推定の挙動を変化させることが可能になった。

参考文献

- [1] 須藤 郁弥, 成澤 和志, 篠原 歩. UEC コンピュータ大貧民大会向けクライアント「snowl」の開発. 第2回 UEC コンピュータ大貧民シンポジウム