

jh170038-DAH

## 大規模な強化学習技術の実証と応用

金子知適（東京大学大学院情報学環）

### 概要

本研究では、近年注目されている強化学習技術の研究を進め、大規模なデータを利用することで人工知能システムの性能を向上させることを目指し、また将来への応用として学習した内容を自然言語で表現する可能性について模索するという計画で始めたものである。DeepMind 社の Alpha 碁の成功が示したように、各分野で人工知能システムが人間の専門家の判断力を超える時代が到来しつつあるため、高精度の AI 構築技術は判断の説明技術とともに知見を蓄積することが重要である。そのために、高性能な人工知能システムを作成する過程で行われる、1000 万から 1 億といった高次元の機械学習の過程に注目することが有力で、関連する複数のドメインでそのような大規模な学習の実証実験を行った。将棋で約 3,000 万棋譜、囲碁で 19 路と 9 路盤をあわせて、約 200 万局強の棋譜を作成し、それをもとに強化学習の実験を進め、教師データに対する効率を改善する学習方法に関する論文としてまとめた。

### 1. 共同研究に関する情報

#### (1) 共同研究を実施した拠点名

東京大学

#### (2) 共同研究分野

- 超大規模数値計算系応用分野
- 超大規模データ処理系応用分野
- 超大容量ネットワーク技術分野
- 超大規模情報システム関連研究分野

#### (3) 参加研究者の役割分担

- 代表者 金子 知適 東京大学・情報学環 強化学習の大規模化
- 副代表者 山口 和紀 東京大学・総合文化研究科 モデルの精緻化
- 共同研究者
  - 万代 悠作 東京大学・総合文化研究科 モンテカルロ木探索
  - 横山 秀 東京大学・総合文化研究科 疎結合  $\alpha\beta$  探索
  - 万 山川 東京大学・情報学環 不

### 完全情報ゲーム

- 森永 雄也 東京大学・総合文化研究科 逆辞書
- 伊部 早紀 東京大学・総合文化研究科 統計的機械翻訳
- 

### 2. 研究の目的と意義

本研究では、近年注目されている強化学習技術の研究を進め、大規模なデータを利用することで人工知能システムの性能を向上させることを目指す。また将来への応用として、学習した内容を自然言語で表現する可能性について模索する。DeepMind 社の Alpha 碁の成功が示したように、各分野で人工知能システムが人間の専門家の判断力を超える時代が到来しつつあるため、AI の判断を言葉での表現する技術への期待は高い。そのためには、高性能な人工知能システムを作成する過程で行われる、1000 万か

ら 1 億といった高次元の機械学習の過程に注目することが有力で、関連する複数のドメインでそのような超大規模な学習の実証実験を行うことで、知見を蓄積する。人工知能システムの要素技術としては深層学習が有名となっているが、本研究では強化学習の理論と未来に関する確率的な推論を行うグラフ探索の組み合わせに特に着目する。強化学習とグラフ探索の重要性を示す例としては、Alpha 碁において、深層学習に与える教師データに強化学習の観点で独自の工夫がなされていること、木探索を思考の基幹として深層ネットワークはそれに組み込まれる形で用いられたことなどがあげられる。強化学習は、人が教師データを与える必要がない点で応用への期待が高い一方、大量の試行錯誤を行いそれらのデータを処理するための計算機の負荷が大きいという側面を持つ。また理論を実際のドメインに適応させる際の工夫など、未知の部分も多い。そこで、本研究では、互いに関連する複数のテーマにおいて大規模な強化学習を行い、実際のデータを総合的に分析することで汎用性や頑健性の観点でモデルを強化する。遠い目標である言葉での説明のためには、実際の判断の思考記録だけでなく強化学習の過程の分析が重要と考えられる。その緒を探すために、「何か」を言葉にする技術である、統計的機械翻訳や、逆辞書の研究との関連を模索する。

近い将来に AI システムが社会の様々な場面で導入されることはほぼ間違いないが、AI が社会に受け入れられまた活用されるためには、信頼できる AI システムの作成技術を確立する必要がある。信頼を得る努力には大きく分けて 2 種類があり、性能の良いシステムを作るという点と、AI の個々の提案を検証できる情報を理解可能な形で提供する方向がある。将棋や囲碁では、前者が達

成されつつあるが、後者については、AI の提案がこれまでの人間の常識を変える良い手なのか、あるいは暴走であって棄却すべきなのかを判断することは大変難しいと示されている。現状でも思考ログを提供することは可能だが、それはそのドメインの専門家が把握できるものではないため実質的に何もできていない。本研究は、前者の方向について複数のドメインで実証し発展させつつ、後者の技術への展開を模索するという点で、大きな目標に貢献するものである。大規模な強化学習については、Alpha 碁に限らず複数の成功が報告されはじめているが、囲碁や将棋、あるいはポーカーなどの不完全情報ゲームなどの複数のドメインを比較した事例はまだない。また学習で得た内容をコンピュータに説明させようという試みはいくつかなされているが、単体の予測器の範囲がほとんどで、グラフ探索と組み合わせで動作する場合の研究はない。したがって、これらについて本研究で取り組む意義がある。

### 3. 当拠点公募型共同研究として実施した意義

人工知能システムが高性能な判断力を備えるためには、大規模な学習が必要である。規模は年々大きくなり、一般の研究者が個人で所有する設備では難しくなっている。まず、思考ゲームの分野では 1 年前の時点で、3,000 万局の棋譜が必要であると報告されている。棋譜は 1 手ごとのグラフ探索を行いその結果を記録したものである。このグラフ探索は、機械学習で得た評価関数 (value network) による状況判断に基づく。1 手あたり 0.1 秒で着手したとしても、それだけ棋譜を揃えるためには 10 年単位必要な計算であり、それだけでも並列分散処理や他の工夫が必須である。さらに本研究で

扱う強化学習は、試行錯誤と経験からの学習を繰り返すものであり、(1)自身の対局により上記の規模の棋譜データを作成する(2) その経験から学習し自身の思考方法を調整するということを、10 から 100 回程度繰り返すだけの計算資源が必要となる。本研究には、自作のプログラムだけでなくこの分野で標準的なオープンソースソフトウェアも利用するため、Intel CPU, Linux システム, Nvidia Pascal 世代の GPU といった他の研究者が標準的に用いているハードウェアで実験を行う必要がある。総合して、Reedbush-{U, H} システムを活用することが適している。

#### 4. 前年度までに得られた研究成果の概要

新規課題である。

#### 5. 今年度の研究成果の詳細

##### 5-1 中間報告までの成果

将棋においては、gpsfish という将棋プログラムを大規模に運用し、様々な設定で約 3,000 万局の棋譜を得た。現在それらを持ちいて様々な実験を運用中である。

囲碁においては、aya という準トップレベルのプログラムをお借りして、19 路盤と 9 路盤のそれぞれで、約 100 万局の棋譜を作成した。囲碁では GPU を用いるため、将棋よりも棋譜の作成に時間がかかっている。現在それらを持ちいて様々な実験を運用中である。自然言語処理との連携を採る課題においては、単語ベクトルの改良と目的単語のカテゴリ推定機構の導入により (Hill et al. 2016) の検索システムの精度を改善することに成功した。この成果は研究成果リスト d) として発表されている

##### 5-2 最終成果

中間報告以降の主要な成果として、「研究成果 b) RankNet for evaluation functions of the game of Go と c) 局面の組合せを用いた囲碁評価関数の学習」としてまとめた研究と「研究成果 a) Building Evaluation Functions for Chess and Shogi with Uniformity Regularization Network」としてまとめた研究の二点について報告する。

はじめに、研究成果 b) RankNet for evaluation functions of the game of Go と c) 局面の組合せを用いた囲碁評価関数の学習 . について述べる。研究成果 c) を進めてより良い結果を得たものが、研究成果 b) である。

この研究では東京大学も含めて国内の通常の研究機関で利用可能な計算機資源は、Google/DeepMind 社よりはるかに劣るという状況に対応するために、比較的少ない教師データから知識を吸収するという点での学習効率の向上に取り組んだものである。局面の組合せを利用した囲碁評価関数の学習法について提案した。深層学習において、複数の入力を持つネットワークは近年盛んに研究されており、様々な応用例が提案されている。そのような複数の入力を持つネットワークの学習を注意深く応用することにより、教師例の数を二乗のオーダーで増やすことが可能と期待されるものである。近年の人工知能の発展には、探索手法の向上とならんで深層学習 (deep learning) 技術の進化が背景にある。深層学習によって以前は困難だと考えられてきた囲碁の評価関数の学習が成功し、人間と同等以上の性能を発揮した。その評価関数の学習においては人間の熟達者が残した棋譜を用いて評価関数を学習し、一定の強さをもつエージェントを作成したのちに強化学習によって

さらに性能のよい評価関数を学習している。一方で、近年様々な深層学習の応用研究がなされており、そのうち複数の入力を持つ深層ニューラルネットワーク (DNN) の研究が近年注目されている。特に入力の数が増える DNN は Siamese ニューラルネットワークと呼ばれ、古くから研究されており、最初期の例では署名の検証の研究において提案されている。近年では画像の類似度の判定や、テキストの類似度判定などで利用されている。ゲームにおいて Siamese ネットワークを用いた先行研究として、チェスの評価関数を Siamese ネットワークを用いて作成した DeepChess が存在する。いずれの例でも、入力が二つ存在するネットワークの訓練では、 $N^2$  個の教師データを組合せを用いることで  $O(N^2)$  に増やすことができ、既存の知見が十分に蓄積されていない領域でも教師あり学習を効率的に行える可能性がある。本研究では上記のような、二つの入力を持つ DNN を利用し、囲碁の評価関数を作成することを目的とする。具体的には入力として二つの局面を受け取り、どちらがどれだけ優れているを判定する DNN の学習を目的とする。このような特徴を持つ DNN の先行研究として、ランキング学習 (learning to rank) で用いられている RankNet が挙げられる。RankNet は入力を二つ受け取る Pairwise な学習手法であり、そのどちらが優れているかどうかを学習することが可能である。実験では学習に用いる棋譜の数を変化させて DNN を訓練し、正答率、交差エントロピー損失、そして対戦成績という側面から性能を評価し、棋譜の数が少ない状況で、既存手法より高い勝率を得た。この研究実施時点での関連研究としては AlphaGo, DeepChess, RankNet が特に重要で

あった。

AlphaGo は Google DeepMind が開発した囲碁プログラムであり、モンテカルロ木探索と DNN による評価関数によって次の着手を決定している。AlphaGo は 2015 年 10 月に行われた、当時欧州王者である Fan Hui 二段との対局において五勝〇敗の成績を残した。またそこから改良された AlphaGo は 2016 年に行われた、当時最も強い棋士の一人である Lee Sedol 九段との対局で四勝一敗の成績を収めた。

AlphaGo の評価関数は 1) 既存の棋譜からの方策関数の学習を行った後、2) 自己対局による価値関数の学習を行うことによって作成されている。方策関数の学習では既存の棋譜から (局面, 着手) のペアを教師例として用いて次の行動の確率を学習する。価値関数の学習では自己対局による棋譜から (局面, 終端局面の勝敗) のペアを教師例としてある局面の勝率を学習する。この学習は非常に大規模であり、既存の棋譜として 3000 万局面、自己対戦から抽出した局面も同程度の局面数が作成された。

DeepChess は、ゲームの評価関数を局面の組合せを用いて行った先行研究のひとつである。DeepChess はコンピュータチェスプログラムであり、本稿ではそれが用いている DNN を指す。DeepChess は入力として局面对を受け取り、そのどちらがより好ましい局面かを二項分類問題として出力する。入力局面を二つ受け取るため、DeepChess は二又のネットワークという、通常は使わない構造を採用している。

DeepChess では全結合層を多用していることもあり、重みの初期値を事前に自己符号化器 (autoencoder) で調整してから学習を行っている。具体的には、積層自己符号化器 (stacked autoencoder) として学習している。自己符号化器によって重みの初期値

を得た後、学習を次のとおりに実行する。まず、用いる棋譜に含まれる局面を白番から見た勝ち局面集合  $W$  と負け局面集合  $L$  に分割する。その後  $W, L$  から一局面ずつ  $w, l$  を乱択抽出し、順番をランダムに入れ替えてネットワークに入力する。つまり確率 0.5 で  $(w, l)$ , 0.5 で  $(l, w)$  の順番でそれぞれの下部ネットワークに入力する。ネットワークは勝ち局面がどちらであるかを正しく判定するように重みを更新する。

DeepChess の学習では  $|W|, |L|$  とともに百万程度であり、手番の対称性なども考慮した場合、組合せを考えた場合にはおよそ  $2 \times 10^{12}$  通り程度作成することができる。実験ではテスト集合における予測精度は 98% 程度であると報告されている。対戦においても効果的だったと報告されているが、ニューラルネットワークの特殊な形状から応用が限定されると判断して、本研究では別の方向を目指した。

それが RankNet であり、ランキング学習 (learning to rank) の手法の一つである。ランキング学習は入力されたクエリに対してそれぞれのクエリの好ましさを入力するような学習器を構築することが目的である。ランキング学習の手法は入力する特徴ベクトルの数によって分類されることが多く、単一の入力の場合 Pointwise、二つの場合 Pairwise、そしてリストの場合 Listwise と呼ばれる。RankNet はランキング学習に DNN を用いた手法で、Pairwise なランキング学習に分類される。

RankNet には二つのクエリ  $U_i, U_j$  のそれぞれの特徴ベクトル  $x_i, x_j$  が入力される。入力されるクエリ  $U_i, U_j$  において  $U_i$  が  $U_j$  より好ましいという事象を  $U_i \rightarrow U_j$  と表記する。RankNet は特徴ベクトルを実数値に写像する DNN  $f$  を用いて、それぞれのクエリのスコア  $s_i = f(x_i), s_j =$

$f(x_j)$  を計算する。このとき、事象  $U_i \rightarrow U_j$  である確率  $P_{ij} = \Pr(U_i \rightarrow U_j)$  をシグモイド関数とスコアを用いて

$$P_{ij} = \frac{1}{1 + e^{-\sigma(s_i - s_j)}}$$

であると仮定する。ここで  $\sigma$  はパラメータである。この確率を用いて、交差エントロピー損失関数  $C$  を以下のように表記できる:

$$C = -\bar{P}_{ij} \log P_{ij} - (1 - \bar{P}_{ij}) \log(1 - P_{ij}).$$

ここで  $\bar{P}_{ij}$  は  $U_i$  が実際に  $U_j$  よりも優れている確率である。

学習時には  $U_i, U_j$  に加えて  $S_{ij} \in \{1, -1\}$  を教師例として入力する。

このとき  $U_i \rightarrow U_j$  であるならば  $S_{ij} = 1$ ,  $U_j \rightarrow U_i$  であるならば  $S_{ij} = -1$  とする。

そうすると、既知の確率  $\bar{P}_{ij}$  は  $S_{ij}$  を用いて  $\bar{P}_{ij} = \frac{1}{2} (1 + S_{ij})$  と計算できる。よって損失関数は

$$C = 1/2 (1 - S_{ij}) \sigma(s_i - s_j) + \log(1 + e^{-\sigma(s_i - s_j)})$$

となる。

RankNet の学習とは DNN  $f$  の学習であり、二つの入力  $x_i, x_j$  から  $s_i = f(x_i), s_j = f(x_j)$  を計算した後、 $s_i, s_j$  と  $S_{ij}$  を用いて損失  $C$  を計算し、誤差伝搬法によって  $f$  の重みを更新する。提案手法は、局面の組合せを利用した RankNet を用いて囲碁評価関数の学習手法をおこなうことである。

つまり、二つの局面を受け取って、どちらの局面が優れているかを入力するネットワークの学習法について提案した。まずはじめに用意した棋譜集合を手番の勝利局面集合  $W$ , 手番の敗北局面集合  $L$  に分割し、それぞれの集合を訓練集合 ( $W_{\text{train}},$

$L_{\{train\}}$  とテスト集合 ( $W_{\{test\}}$ ,  $L_{\{test\}}$ ) に分割する。学習の際には、 $W_{\{train\}}$ ,  $L_{\{train\}}$  からランダム一つずつ局面  $w_i$ ,  $l_j$  を抽出し、順番もランダムに入れ替えた局面对と教師例 ( $x_i$ ,  $x_j$ ,  $S_{\{ij\}}$ ) を作成する。ここで  $x_i$ ,  $x_j$  は局面の特徴ベクトルであり、どちらかが勝利局面  $w_i$  を表す特徴ベクトルで、もう一方が敗北局面  $l_j$  を表す特徴ベクトルである。ここで、 $x_i$  が  $w_i$  に対応するならば  $S_{\{ij\}} = +1$ , そうでないならば  $S_{\{ij\}} = -1$  となる。その後、それぞれの入力のコス  $s_i$ ,  $s_j$  を DNN を用いて計算し、スコアから損失  $C$  を求め、誤差伝搬法によって DNN の重みを更新する。学習した DNN は局面の特徴ベクトルから実数値へと写像する関数  $f$  となり、なおかつ二つの局面  $s_i$ ,  $s_j$  に関して  $s_i$  が  $s_j$  よりも優れているならば  $f(s_i) > f(s_j)$  であると期待することができる。

よってこの  $f$  そのものを評価関数として利用できることを期待できる。

実験はすべて九路盤で行った。実装には python 3 を、深層学習のフレームワークとして chainer mn を使用した。実験は RankNet を用いた DNN の学習の性能評価、学習した DNN の対戦における強さの測定、及び着手の予測性能について行った。

まずニューラルネットワークの学習には山下宏氏によって作成されたコンピュータ囲碁プレイヤーによって生成された棋譜を用いた。この囲碁プレイヤーは CGOS の BayesElo で 2,900 程度の棋力を持つ。また対局において、一方が投了した後には目数差が最大になるような方策に切り替えて着手を行い、死石をすべて打ち上げ、双方がパスを選択するまで行動選択を行っている。早めに投了する代わりにこのように最後まで打ち続けることで、終盤の棋譜も学習対象にする

ことができ、学習が安定すると期待される。コミは 7.0 目であり、よって引き分けが起こりうる。学習に用いる際には、一つの対局からランダムに一局面のみを抽出したものをを用いた。これは、局面と勝敗を関連付ける学習において、局面の独立性をなるべく高めるためである。

提案手法で用いる DNN は、特徴ベクトルを実数値に写像するものであれば任意のものを使用できる。ここでは AlphaGo および AlphaGo Zero にそそえて実験を行った。具体的な実験結果については、出版された論文を参照されたい。

つづいて、研究成果 a) Building Evaluation Functions for Chess and Shogi with Uniformity Regularization Networks について述べる。

2017 年後半に AlphaZero の論文が、arXiv に投稿された。DeepMind の AlphaGo のシリーズは、AlphaGo、AlphaGo Zero、Alpha Zero と進化して現時点ではこれが最新である。AlphaGo から AlphaGo Zero では、人の棋譜を使わず、自己対局による強化学習のみに手法が簡素化された。一方で、同時に、より莫大な計算機資源が自己対局のために必要となり、大量の TPU が投入された。ニューラルネットワークの構造では、画像認識の課題では標準的な ResNet が囲碁にも導入された。同時に、Policy ネットワークと Value ネットワークの畳み込み層を共有させ、同時に訓練することにより学習の安定性を実現している。Alpha Zero では、この AlphaGo Zero の手法が、チェスト将棋でも有効であると示したものである。DNN が、チェスや将棋でも有効であるという予備的な知見は以前から得られていたが、実際にトッププログラムに大きく勝ち越すことは研究者の予想を越えたもので世界に大きな衝撃が走っ

たと言える。

この研究の進展を受けて、当グループの研究計画も若干変更した。すなわち、当初は深層ニューラルネットワークの研究は囲碁を対象に行い、将棋では超高次元の線型結合を用いる予定だった。計画段階では、それぞれの手法がそれぞれで適していると、一般に考えられていたためである。しかしながら、Alpha Zero の成功を受けて、将棋およびチェスでも深層ニューラルネットワークの研究を行い知見を蓄積することの重要性が非常に高まったと判断し、深層ニューラルネットワークの研究により力を注いだ。当研究は、そのような背景で、将棋及びチェスを対象に、深層ニューラルネットワークによる形勢判断の学習に取り組んだものである。

前半の研究成果でも述べた研究と同様に、この研究でも計算機資源の制約から、比較的少数の(多数ではあるが現実的な範囲の)棋譜からの学習において、吸収できる強さを向上させるという意味での効率化に取り組んだ。具体的には、Alpha Zero で行われている「Policy ネットワークと Value ネットワークの畳み込み層を共有させ、同時に訓練することにより学習の安定性を実現している」部分についての改善に取り組んだ。Policy ネットワークとは、チェスや将棋のようなゲームでは囲碁よりもずっと構造が複雑になり、学習の負担となる。そこで代わりに、敵対的学習 (Generative Adversarial Networks) の Discriminator にヒントを得た識別機を導入し、Value ネットワークと畳み込み層を共有させて、同時に訓練することで学習を安定させる手法を提案した。

さらに、単に同時に訓練するだけでなく、この共同訓練を行ってはじめて導入可能となる、Uniformity 正則化という新たな手法

を提案している。これはミニバッチ内の Discriminator の成功率を揃える方向に制約を設けるものである。計算機実験では、様々な損失関数で Value ネットワークの訓練を行い、どの損失関数に追いても、提案した共同訓練で学習したネットワークが安定して良い性能を持つことを示している。そのうえで、Uniformity 正則化は、共同訓練で向上した性能を更に高めている。この安定化は、過学習を防止する効果が高いことが、意図的にデータ量を減らした際の性能減衰度合いを測る計算機実験によって確かめられ、提案手法の良さを示している。学習したネットワークの性能は、既存の深層学習の研究と、局面の合法手を全て評価して最大値を取る局面が棋譜の指し手と一致する率により比較された。さらに、探索を加えた場合の一致率についても、既存の将棋プログラム (Bonanza) やチェスプログラム (stockfish) を対象に比較した。具体的な実験結果については、出版された論文を参照されたい。この研究項目については、現在投稿中のものしか存在しないが、万一不採択になった場合も、arXiv 等も含めた適切な方法で速やかに出版する予定である。

## 6. 今年度の進捗状況と今後の展望

進捗状況:

研究成果の詳細に示したとおり、複数の国際会議論文を産出し研究グループの規模としては妥当な成果を挙げられた。この点で、代表者として満足している。

また、報告時点では未完成でありまだ論文にすることはできていないが、さまざまな基礎的な研究を継続して実施中である。究極的には、将棋、囲碁(9路盤、19路盤)で強いプログラムを作成することがプロジェクトの目標であるが、その途上で様々な課題

を解決する必要があるためである。今後も適切な観点から、論文として切り出して、成果発表を行う予定である。

研究計画を立ててから一年間の間にも、ゲームを題材とした強化学習の研究は大きく進み、特に Google DeepMind 社が研究成果のインパクトの点で大きく先行されている。日本に限らず他の研究機関はほとんど引き離されているが、中国のテンセント社など、大規模な計算機資源を持つグループが比較的近い成果を出しつつある。日本の研究機関としては、(仮に ReedBush や他の資源を使ったとしても 5,000TPU で実験を行う DeepMind よりは遥かに)少ない計算機資源でどこまで巨大企業に迫れるかという観点で課題を明らかにすることが重要であり、本研究課題での目標の一つであった。この点でも、当研究グループの知見は大きく広がっており、今後に向けて十分な前進を得たと評価できる。各課題を検討して知恵を絞り、また研究資源を効果的に投入することで、距離を縮めてゆくことが出来ると期待される。

今後の展望:

強化学習の研究は今年も大きく進み、画像のキ説明生成、質の異なる仕事の金額評価と言った、ゲームとは大きく異なる応用にまで使われている。したがって、本研究課題である大規模な強化学習は、これからの日本社会に不可欠であり、現時点では Google DeepMind 社のように潤沢な資源が使った派手な応用成果があげられなくとも、地道に継続して基礎的な技術と知見を国内に蓄積することが重要と、報告者は信じるものである。研究を続けていれば、将来計算機資源がより安価になった際に、さまざまな応用を実現することにつながると期待できる。一方でもし、すぐに「勝てない」

からと、研究を避けて放置してしまえば世界との差が開くばかりとなる。総合してこの分野を離脱することは、将来の発展を阻害するというリスクのほうが大きく得策でないと報告者は考えている。

当グループはこの報告の翌 2018 年度も HPCI でない課題として鋭意研究を進めている。現時点で具体的に書くことは控えるが、強化学習の課題として curiosity (新規事象を探索する動機づけ)の活用、学習の効率化として joint learning, transfer learning, mix-in の活用、人間への説明という観点でニューラルネットワークの判断の可視化技術などが重要と、考えている。

## 7. 研究成果リスト

### (1) 学術論文

#### (2) 国際会議プロシーディングス

a) Building Evaluation Functions for Chess and Shogi with Uniformity Regularization Networks.

Shanchuan Wan and Tomoyuki Kaneko.  
IEEE Conference on Computational Intelligence and Games (投稿中)

b) RankNet for evaluation functions of the game of Go.

Yusaku Mandai and Tomoyuki Kaneko.  
The 10th International Conference on Computers and Games (投稿中)

### (3) 国際会議発表

### (4) 国内会議発表

c) 万代 悠作, 金子 知適.  
局面の組合せを用いた囲碁評価関数の学習.  
ゲームプログラミング・ワークショップ



2017 年 11 月 箱根 pp. 8-14

d) 森永雄也、山口和紀

カテゴリ情報を付与した文の分散表現による逆引き辞書の精度向上

インタラクティブ情報アクセスと可視化マイニング 第 16 回研究会 2017 年

(5) その他（特許，プレス発表，著書等）

なし