グラフニューラルネットワークを用いた大貧民の評価関数の設計

福澤 詩音[1,a)]　松崎 公紀[1,b)]

概要：本研究では，多人数・不完全情報ゲームの一種であるカードゲーム「大貧民」を対象とする．大貧民では，同じランクの組や同じスートの階段など，複数のカードの組み合わせを役として提出する．手札中の一枚のカードは複数の役に同時に属しうるため，ある役を提出すると，カードを共有する他の役は提出できなくなる．著者らはこれまでに，方策と候補手の価値を直接評価する畳み込みニューラルネットワークを実装した．本研究では，新たにカード間の関係と，手札から構成できる役とカードの関係をグラフとして表現し，グラフニューラルネットワークにより方策と状態価値を出力する評価関数を設計した．入力に与える関係を変えた 3 種類のネットワークと畳み込みニューラルネットワークを，総パラメータ数をほぼそろえた条件のもとで比較した．その結果，役をノードとして与えたネットワークが最も高い性能を示した一方，スート間で情報を交換する機構を加えた畳み込みニューラルネットワークを対照に置くと，通常の局面では同程度の性能を示した．両者の性能に差が見られるのはカードの強弱が反転する革命状態の局面であり，順序を辺の向きとして表現することがこのような局面で有効であることが示唆された．

Design of Evaluation Function for Daihinmin Using Graph Neural Network

Shion Fukuzawa[1,a)]　Kiminori Matsuzaki[1,b)]

Abstract: In this study, we investigate Daihinmin, a multiplayer card game with imperfect information. Players play card combinations, referred to here as melds, such as sets of cards of the same rank or sequences of cards of the same suit. A single card may belong to multiple possible melds, but playing one meld makes any other meld containing that card unavailable. In previous work, we developed a convolutional neural network (CNN) that directly estimates a policy and the value of a played meld. In this study, we represent relationships among cards and between cards and candidate melds in a player's hand as a graph, and develop a graph neural network (GNN) that predicts both a policy and a state value. We compare three GNN variants that encode different types of relationships with a CNN baseline, using models with approximately equal numbers of parameters. The GNN incorporating meld nodes achieved the best overall performance, whereas a CNN augmented with a mechanism for exchanging information across suits achieved comparable performance in ordinary game states. However, the two models differed in their performance in positions following a revolution, which reverses the relative strengths of card ranks. This result suggests that representing rank-order information through directed edges is particularly beneficial in such positions.

1. はじめに

大貧民は日本で広く親しまれているトランプゲームであり，多人数・不完全情報ゲームに分類される．提出役に単体役・複数役・階段役という複数の種類があることに加え，カードの強弱が反転する「革命」などの独自ルールがあり，相手の手札を読みながら自分の手札を効率よく処理する戦略性が求められる．また，一局の結果に応じて次局の開始時にカードを交換するため，各局は独立ではなく，前局の

[1] 高知工科大学大学院 工学研究科
Kochi University of Technology
[a)] 305144m@gs.kochi-tech.ac.jp
[b)] matsuzakikiminori@kochi-tech.ac.jp

© 2026 Information Processing Society of Japan

順位が以降の局の初期条件を変える．これまで，UEC コンピュータ大貧民大会（UECda）[1] を中心に多くのコンピュータプレイヤの研究が進められ，とくに比較的単純な評価関数とモンテカルロ探索を用いたプレイヤが長く上位を占めてきた [2]．著者らはこれまでに，方策と候補手の価値を直接評価する畳み込みニューラルネットワークを実装した [3]．

大貧民では，手札中の一枚のカードが複数の役に同時に属しうる．このとき，ある候補役を選択すると，そのカードを共有する他の候補役は以後提出不可能となる．したがって，手札を評価する際には，カードと役の相互作用に着目することが有効だと考えられる．カードを 4 スート×13 ランクの格子に並べて畳み込みニューラルネットワーク（CNN）で処理する場合，こうした関係は格子上の位置関係として与えられる．そこで本研究では，カードの強弱の関係に加えて，カードがどの役に利用されるかの関係を明示的に表したグラフを入力とし，グラフニューラルネットワーク（GNN）により方策と状態価値を出力する評価関数を設計する．

ゲームの局面評価に GNN を適用した報告は限られており，オセロ [4] やチェス [5] など二人完全情報ゲームを対象としたものが中心である．不完全情報かつ多人数のゲームで同様の効果が得られるかは明らかでない．

以上を踏まえ，本研究では入力として与える関係を段階的に増やした 3 種類の GNN を CNN と比較し，次の二つを検証する．

(1) 不完全情報・多人数ゲームである大貧民において，ゲーム規則に由来する関係を明示的に与えることは，局面評価の精度と対戦成績を改善するか．
(2) その改善はグラフという表現形式に固有のものか，それとも畳み込みの枠内でも実現できるものか．

二つめの問いのために，GNN が持つ関係を畳み込みの枠内で模した対照モデルを設計し，両者が一致する条件と分かれる条件を特定する．

2. 本研究で用いる大貧民のルール

本研究で用いる大貧民のルールは，UEC コンピュータ大貧民大会 [1] で配布されているサーバ実装の標準設定に従う．

ゲームは 5 人で行い，ジョーカー 1 枚を含む 53 枚のカードを用いる．各プレイヤは自分の手札と場に提出されたカードのみを観測でき，他プレイヤの手札は分からない．提出できる役は，1 枚の単体役，同一ランクの複数枚からなる複数役，および同一スートで連続する 3 枚以上の階段役である．ジョーカーは任意のカードの代用として使用できるほか，単体で最強のカードとして提出できる．また，同じスートが連続して提出されると，以降その場ではそのスートの札しか出せなくなる「縛り」が発生する．

一局が終了すると，手札をなくした順に大富豪，富豪，平民，貧民，大貧民の身分が与えられる．また，得点として，大富豪に 5 点，富豪に 4 点，平民に 3 点，貧民に 2 点，大貧民に 1 点がそれぞれ与えられる．次局の開始時には，大貧民は大富豪へ最も強い 2 枚を渡し，大富豪は大貧民へ任意の 2 枚を返す．同様に貧民と富豪は 1 枚を交換する．したがって各局は独立ではなく，前局の順位が次局の初期条件を決める．ただし本設定では 100 局ごとに身分をリセットし，その局ではカード交換を行わず，最初の手番を無作為に選ぶ．

プレイヤの強さは 1 局あたりの平均得点で比較する．5 人全員の期待値は 3.0 点であり，3.0 点が互角の水準となる．

3. 提案手法

3.1 グラフニューラルネットワーク

グラフニューラルネットワークは，ノードと辺からなるグラフを入力とし，各ノードが辺で結ばれた隣接ノードの特徴を集約して自身の表現を更新する操作を繰り返すことで，グラフ構造を反映した表現を得る手法である．この更新を 1 層とし，複数層重ねることで，各ノードの表現には辺をたどって到達できる範囲の情報が順次取り込まれる．

畳み込みニューラルネットワークとの違いは，参照先の与え方にある．畳み込みは入力を格子状に並べ，各要素がカーネルの大きさで定まる固定の近傍を参照する．近傍は格子上の位置によって決まるため，問題に応じて変えることはできない．これに対しグラフニューラルネットワークでは，どのノードとどのノードが情報をやり取りするかを辺として明示的に与える．すなわち近傍の定義そのものが設計の対象となる．辺に向きを与え，向きごとに別の種類として扱えば，情報が自分より上位のノードから来たのか下位から来たのかを区別できる．

3.2 局面のグラフ表現

大貧民では，手札中の一枚のカードが複数の役に同時に含まれる場合があり，ある役を提出すると，そのカードを共有する他の役は提出できなくなる．本研究では，こうしたカード間の関係と，カードと役の関係をグラフとして明示的に与える．図 1 に，スペードの 3, 5, 6, 7, 9 とハートの 3 を手札とする場合のカードグラフを示す．

ジョーカーを除く 52 枚のカードをノードとし，2 種類の辺を張る．同ランク辺は，同一ランクで異なるスートのカードどうしを結ぶ無向辺であり，複数役の構成関係に対応する．各ランクについて 4 枚が完全グラフをなす．強弱辺は，同一スートでランクが隣接するカードを弱い側から強い側へ向かう有向辺であり，階段役および強さの順序に対応する．強弱辺は向きごとに別の関係として扱い，それぞれに異なる線形変換を与える．革命状態では辺の向きを反転させる．これら 2 種類は手札の内容によらず全 52 枚

© 2026 Information Processing Society of Japan

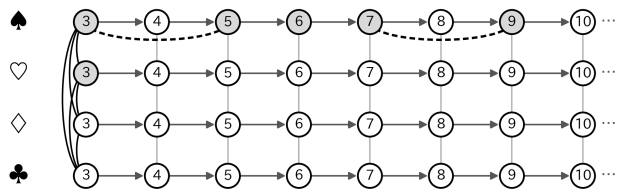


図 1 カードグラフの例．手札はスペードの 3, 5, 6, 7, 9 とハートの 3 であり，灰色の丸が手札のカード，白い丸が手札にないカードを表す．横向きの矢印が強弱辺，縦の線が同ランク辺，破線の弧が不連続辺である．強弱辺と同ランク辺は手札によらず全 52 枚に対して定義されるが，紙面の都合でランク 3 から 10 までを示し，同ランク辺はランク 3 の列のみ完全グラフとして描いた

に対して定義し，手札に含まれるか否かは辺の有無ではなくノードの入力特徴として与える．

これに加えて，手札に依存する辺を導入する．同一スート内の所持カードをランク順に並べ，その並びで隣り合いながらランクが連続していない 2 枚を結ぶ．以下これを不連続辺と呼ぶ．ランクが連続する 2 枚はすでに強弱辺で結ばれているため，不連続辺はそれを補うものであり，階段役を構成する際の欠落に対応する．この辺も向きごとに別の関係として扱う．

さらに，手札から構成できる役をノードとして追加し，構成カードとの間で，カードから役，役からカードへの伝播をそれぞれ別の重みで 1 回ずつ行う．ここでいう役とは，単体役，同一ランクの複数役，同一スートの階段役であり，ジョーカーを代用として含むものも区別して数える．役ノードは場に対する合法手ではなく手札のみから定まるため，現在は場に出せない役も表現に含まれる．これは，直ちに提出できない役であっても，それを構成するカードを温存すべきかどうかの判断に寄与すると考えたためである．学習データ上での個数は 1 局面あたり平均 13 個であった．ジョーカーにはノードも辺も与えず，ジョーカーを含む役は通常カードの構成員のみに接続する．

図 2 に，図 1 と同じ手札について，カードと役を結ぶ二部グラフを示す．スペードの 3 や 5 のように，1 枚のカードが複数の役に属する．

各カードノードの入力は 3 次元であり，自分の手札に含まれるか，まだ誰も提出しておらず自分も持っていないか，現在場に出ているかを表す．役ノードの入力は 33 次元であり，ランク 14 次元，種類 4 次元，枚数 14 次元，ジョーカー使用の有無 1 次元の one-hot 表現を連結したものである．これとは別に，グラフに属さないスカラー入力を 17 次元与える．内訳は，ジョーカーの所在 6 次元，各プレイヤの残り手札枚数 5 次元，各プレイヤの身分 5 次元，革命状態か否か 1 次元である．

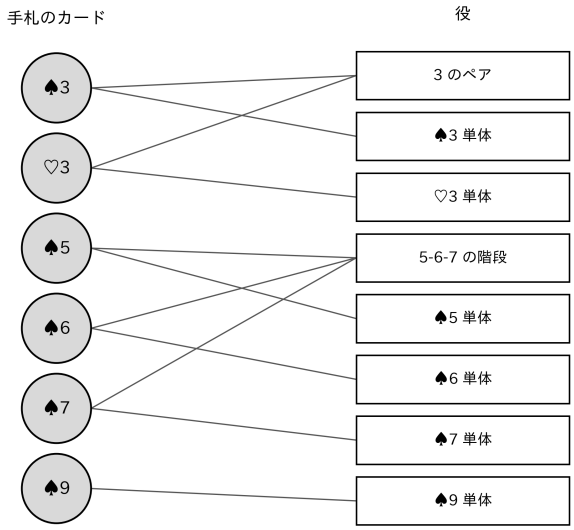


図 2 図 1 と同じ手札における，カードと役の二部グラフ．左の丸が手札のカード，右の箱が役であり，この手札から作れる役 8 個をすべて示した．スペードの 3 とハートの 3 はそれぞれ単体役とペアの双方に，スペードの 5, 6, 7 は単体役と階段の双方に属する．役ノードどうしを結ぶ辺は設けていない

3.3 ネットワークの構成

前節のグラフに対し，関係の種類ごとに異なる重みを持つグラフ畳み込み [6] を 4 層重ねる．関係は，同ランク辺 1 種類，強弱辺 2 種類，不連続辺 2 種類の計 5 種類である．カード i の第 l 層における表現 $x_i^{(l)}$ は次式で更新する．

$$x_i^{(l+1)} = \mathrm{ReLU}\left(W_0 x_i^{(l)} + \sum_r \frac{1}{|N_r(i)|} \sum_{j \in N_r(i)} W_r x_j^{(l)} \right)$$

ここで $N_r(i)$ は関係 r におけるカード i の隣接ノード集合であり，W_0 は自己変換，W_r は関係ごとの変換を表す．隣接ノード数で除しているのは，関係によって次数が異なるためである．カードの初期表現 $x_i^{(0)}$ は，3 次元の入力を線形変換により 64 次元へ写像して得る．

この更新を 4 層適用した後に，カードから役，役からカードへの伝播を 1 回行う．役 m の構成カードの集合を $C(m)$，カード i を含む役の集合を $M(i)$，役 m の入力特徴を f_m とすると，次式となる．

$$z_m = \mathrm{ReLU}\left(W_s \,\mathrm{ReLU}(W_e f_m) + \frac{1}{|C(m)|} \sum_{i \in C(m)} W_c x_i \right)$$

$$\tilde{x}_i = \mathrm{ReLU}\left(W_x x_i + \sum_{m \in M(i)} W_m z_m \right)$$

役からカードへの集約のみ総和とするのは，多くの役に参加するカードほど大きな信号を受け取るようにし，そのカードの使い回しの利きやすさを表現に反映させるためである．

最後に，52 枚のカードの 64 次元の表現を固定順序で平

© 2026 Information Processing Society of Japan

表 1 比較するモデル

モデル	構成	パラメータ数
CNN	畳み込みのみ	896,289
スート混合 CNN	CNN ＋ スート間の平均	896,309
拡張入力 CNN	CNN の入力を拡張	896,615
GNN-A	同ランク辺，強弱辺	894,704
GNN-B	GNN-A ＋ 不連続辺	890,815
GNN-C	GNN-B ＋ 役ノード（提案）	889,357
GNN-D	GNN-C ＋ 役ノード間の辺	897,677
拡張入力 GNN-C	GNN-C の入力を拡張	895,605
融合モデル	CNN と GNN-C の統合	895,055

坦化した 3,328 次元にスカラー 17 次元を連結し，全結合 2 層を経て 33 次元を出力する．正規化層，残差接続，プーリング，ドロップアウトは用いない．バッチ正規化を用いないのは，対戦時の推論をバッチサイズ 1 で行うためである．活性化関数は ReLU とする．

出力 33 次元の内訳は，状態価値 1 次元と，着手の属性を表す 3 つの独立したソフトマックス関数，すなわちランク 14 次元，種類 4 次元，枚数 14 次元である．種類はパス，単体役，複数役，階段役の 4 通りとする．状態価値の教師は，局の最終的な得点（大富豪の 5 点から大貧民の 1 点）を -1 から 1 へ線形に写像した値である．

着手は，場の状況から合法手を列挙し，各手の最小ランク・種類・枚数に対応する 3 つの周辺確率の積が最大となる手を選ぶことで決定する．合法性の判定はネットワークの外で行うため，縛りや革命による制約はこの段階で処理される．ただしこの方式では 3 属性が同一の合法手を区別できない．例えば同一ランクのカードを 3 枚所持している場合，そのうち 2 枚を選ぶ 3 通りの複数役は 3 属性がすべて等しい．学習データ上でこのような局面は 8.9%であり，着手一致率の上限は 97.0%となる．本研究のプレイヤは探索を行わず，着手の決定には方策のみを用いる．

3.4 比較するモデル

提案手法の各要素がどれだけ寄与しているかを調べるため，表 1 に示す 9 種類のモデルを比較する．出力部と着手の決定方法は，いずれも前節と共通である．

GNN-A，GNN-B，GNN-C：GNN-C が提案手法そのものである．GNN-B はそこから役ノードを除いたもの，GNN-A はさらに不連続辺を除き，手札によらず定まる同ランク辺と強弱辺のみを用いるものである．下位のモデルが持つ関係は上位のモデルがすべて含むため，隣り合うモデルの差が各要素の寄与に対応する．

CNN：カードを 4 スート× 13 ランクの格子に並べ，ランクの軸に沿ってカーネル幅 3 の 1 次元畳み込み（conv1D）を 4 層適用する．本研究の実装では，conv1D はカードの特徴（入力では 3 次元）をチャネルとして畳み込みを行い，4 つのスートには同一の重みを別々に適用する．したがって，スート間の相互作用は全結合層で平坦化されるまで起こらない．中間層のチャネル数は順に 32，64，64，64 であり，padding を 1，ストライドを 1 とし，プーリングは行わない．

スート混合 CNN：カーネル幅 3 の畳み込みは，ランクが一つ下と一つ上のカードに別々の重みをかけて足し合わせる操作であり，GNN-A の強弱辺による伝播と一致する．したがって CNN と GNN-A の違いは，同ランク辺によるスート間の情報交換の有無に帰着する．この違いだけを切り分けるため，CNN の各層の畳み込みの直後に，同一ランク位置における 4 スートの平均を線形変換して全スートへ加算したものを用意した．第 l 層の出力を $u^{(l)}$ とすると次式となる．

$$u^{(l)} = \mathrm{ReLU}\left(\mathrm{Conv}(u^{(l-1)}) + W_{\mathrm{mix}}\,\frac{1}{4}\sum_{s=1}^{4}\mathrm{Conv}(u^{(l-1)})_s\right)$$

ここで s はスートの添字である．スート間を結合する方法には任意性があり，スートをチャネルとして扱えば任意の線形結合を学習できる．ここでは，スートが本来対称で順序を持たないことを踏まえ，置換に対して対称な平均による結合を用いた．これは GNN-A の同ランク辺が行う集約に対応する．

融合モデル：CNN と GNN-C の両枝を持ち，カードごとに両者を重み付けして統合する．両枝が出力する 52 枚ぶんのカード特徴を h^{cnn}, h^{gnn} とし，$g = \sigma(W_c h^{\mathrm{cnn}} + W_g h^{\mathrm{gnn}})$ により $h = g \odot h^{\mathrm{cnn}} + (1-g) \odot h^{\mathrm{gnn}}$ とする．ここで σ はシグモイド関数である．両枝は同時に学習する．

GNN-D：GNN-C に，役ノードどうしを結ぶ辺を加えたものである．役とカードの接続行列を A とすると，$AA^{\top}$ の非対角成分が 2 つの役が共有するカードの枚数を与える．これを正規化して役ノード間の伝播の重みとする．カードを共有する役は同時に提出できないという関係を表す意図である．

拡張入力 CNN，拡張入力 GNN-C：カードの入力を 3 次元から 8 次元に増やして，どのプレイヤが捨てたカードかを区別し，スカラー入力を 17 次元から 43 次元に増やして場に出ている役のランク・枚数・種類を与えたものである．これらは合法手の列挙では用いていたが，ネットワークの入力には含めていなかった情報である．

いずれのモデルも，総パラメータ数は全結合中間層の幅によって調整し，9 モデルで 0.9%以内にそろえた．表 2 に主要な 4 モデルの層ごとのパラメータ数を示す．全結合第 1 層が全体の 85.8%から 95.6%を占めるため，中間層の幅を数十変えるだけで，関係を増やしたことによる増分を吸収できる．本研究では意図的に，関係を多く用いるモデルほど総パラメータ数が小さくなるようにした．

© 2026 Information Processing Society of Japan

表 2 層ごとのパラメータ数

層	CNN	GNN-A	GNN-B	GNN-C
カード埋め込み	320	256	256	256
畳み込み・グラフ畳み込み	30,912	66,560	99,840	99,840
役埋め込み	—	—	—	2,176
役・カード間	—	—	—	16,640
全結合第 1 層	856,576	819,770	782,964	762,888
全結合第 2 層	8,481	8,118	7,755	7,557
計	896,289	894,704	890,815	889,357

4. 実験設定

4.1 学習データ

UECda に出場した既存クライアント同士を対戦させ，その棋譜から学習データを生成した．評価に用いる公開クライアント switch は学習データに含めない．対戦する 5 席の構成を 2 種類用意し，一方は Blauweregen，Ganesa，Rappa，jn17，res.kou2 の 5 種を 1 席ずつとし，他方は Blauweregen，wisteria，default，Grape の 5 種を 1 席ずつとした．いずれの構成にも Blauweregen を含め，その手番を方策の教師とする．2 種類を用いるのは，特定の対戦相手の構成に固有の傾向が学習結果に強く反映されることを避けるためである．

学習データは 50 万試合，約 537 万局面である．これを 8:1:1 に分割し，1 つめは重みの更新に，2 つめはどのエポックの重みを採用するかの判断に，3 つめは最終的な測定にのみ用いる．以下それぞれを学習用，検証用，評価用と呼ぶ．同じ局の局面が複数の分割にまたがらないよう，分割は局単位で行った．

方策の損失は Blauweregen の手番の局面のみを対象とし，状態価値は全プレイヤの全局面で回帰する．強いプレイヤの着手のみを模倣しつつ，価値の学習には全プレイヤの経験を用いるためである．方策の教師となる局面は全体の 21.1%であり，状態価値の教師はその約 4.7 倍となる．

4.2 学習条件

学習条件は全モデルで共通とし，モデルごとのハイパーパラメータ調整は行わない．最適化には Adam を用い，学習率の初期値を 1×10^{-4}，バッチサイズを 256，エポック数を 50 とした．検証用データでの着手一致率が 4 エポック連続で改善しない場合に学習率を半減させる．損失は次式で与える．

$$L = \mathrm{CE}(\mathrm{rank}) + \mathrm{CE}(\mathrm{type}) + \mathrm{CE}(\mathrm{count}) + \mathrm{MSE}(\mathrm{value})$$

ここで 3 つの交差エントロピー CE は Blauweregen の手番の局面のみで計算し，二乗誤差 MSE は全局面で計算する．4 項の重みはすべて等しい．採用するモデルは，検証用データでの着手一致率が最大となったエポックのものとする．学習には NVIDIA GeForce RTX 2060 SUPER（8GB）を用いた．

表 3 switch との対戦成績（4000 局 ×8 回）

モデル	点/局
CNN	2.280 ± 0.022
スート混合 CNN	2.542 ± 0.035
拡張入力 CNN	2.353 ± 0.032
GNN-A	2.526 ± 0.054
GNN-B	2.506 ± 0.020
GNN-C	2.556 ± 0.045
GNN-D	2.541 ± 0.027
拡張入力 GNN-C	2.689 ± 0.018
融合モデル	2.554 ± 0.033

3.0 点が互角の水準．± は 8 回の反復の標準偏差．

4.3 評価方法

着手一致率は評価用データで測定する．ランク・種類・枚数の 3 属性がすべて一致した割合を着手一致率と呼ぶ．3 属性がすべて一致することは，着手そのものが一致することと同義である．

対戦成績は，学習したモデル 1 体と公開クライアント switch 4 体を対戦させ，1 局あたりの平均得点で評価する．カード交換により各局の結果は前局に依存し，また対戦相手の接続順によっても結果が変わるため，乱数の種を変えた 4000 局の対戦を 8 回行い，その平均と標準偏差を示す．

5. 実験結果

5.1 対戦成績

各モデルを公開クライアント switch 4 体と対戦させた結果を表 3 に示す．

結果は 4 つの水準に分かれる．最も低いのが CNN の 2.280 点，次いで拡張入力 CNN の 2.353 点である．スート混合 CNN，GNN-A，GNN-B，GNN-C，GNN-D，融合モデルの 6 モデルは 2.51 点から 2.56 点の範囲に集まり，相互の差は反復間のばらつきと同程度である．最も高いのは拡張入力 GNN-C の 2.689 点である．

CNN とこの 6 モデルを分けているのは，スート間で情報を交換する経路の有無である．GNN-C は CNN を 0.276 点上回るが，スート間の経路を持つモデルどうしの差は対戦成績では判別できない．

入力情報を拡張した 2 モデルは，いずれももとのモデルを明確に上回った．拡張入力 GNN-C は GNN-C を 0.133 点，拡張入力 CNN は CNN を 0.073 点上回る．前者の差は反復間の標準偏差に対して十分に大きく，GNN-C に加えた変更のうち対戦成績を明確に改善した唯一のものである．

なお，3 回の反復で測定していた段階では，GNN-C と GNN-A の差は 0.055 点あり，差があるように見えていた．対戦相手の接続順が実行ごとに変わるため，同一の乱数種でも結果は再現せず，少数の反復では差を過大に見積もり

© 2026 Information Processing Society of Japan

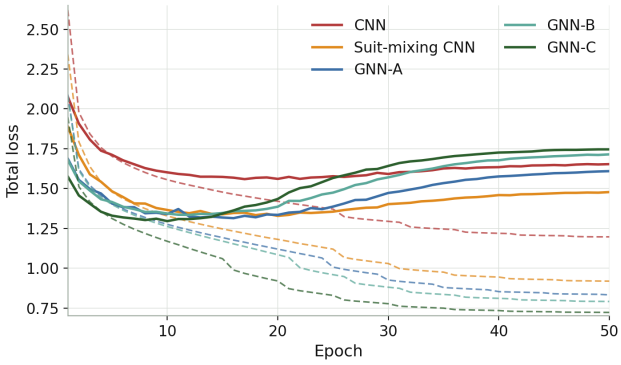


図 3 合計損失の推移（実線：検証用データ，破線：学習用データ）．合計損失は 3 つの交差エントロピーと価値の二乗誤差の和である

やすい．

身分の分布を集計すると，GNN-C は CNN に対して大富豪となる割合が 2.6 ポイント高く，大貧民となる割合が 9.5 ポイント低い（いずれも百分率の差）．差は最下位を回避する能力に現れている．前局が大貧民であった場合に続けて大貧民となる確率は，CNN で 52.2%，GNN-C で 44.5% であった．

switch 以外の 5 種のクライアント（Blauweregen，default，res.kou2，Ganesa，jn17）とも，CNN，スート混合 CNN，GNN-A，GNN-B，GNN-C，融合モデルを 4000 局ずつ対戦させた．いずれの相手に対しても CNN が最下位であり，他の 5 モデル間の差は 0.1 点以内であった．教師とした Blauweregen を相手とした場合，最良のモデルでも 2.05 点にとどまる．

本研究のプレイヤは方策のみで着手を決定するが，参考として，各合法手を打った後の局面の状態価値を方策と等重みで組み合わせて選ぶ条件も測定した．CNN で 2.326 点，スート混合 CNN で 2.493 点，GNN-A で 2.488 点，GNN-C で 2.552 点であり，いずれも方策のみの場合と同程度で，モデルの順位も変わらなかった．

5.2 学習の挙動

図 3 に損失の推移を示す．全モデルで検証用データの損失は途中で最小となり，以降は上昇した．検証用データでの着手一致率が最大となるエポックは，CNN で 26，スート混合 CNN と拡張入力 CNN で 20 であるのに対し，GNN-C と拡張入力 GNN-C では 10，GNN-D では 15，融合モデルでは 13 であった．また 50 エポック目における検証用データと学習用データの損失の差は，CNN の 0.456 に対し，GNN-C で 1.024，融合モデルで 1.472 と大きい．CNN 系のモデルは遅く緩やかに過学習するのに対し，GNN 系と融合モデルは早く強く過学習する．

損失を項別に分解すると，上昇するのは方策の交差エントロピーのみであり，状態価値の二乗誤差は上昇しない．

表 4 評価用データにおける着手一致率

モデル	全体	通常時	革命状態	低下
CNN	0.6386	0.6467	0.5451	0.1016
スート混合 CNN	0.7028	0.7078	0.6458	0.0620
拡張入力 CNN	0.6644	0.6729	0.5668	0.1061
GNN-A	0.7030	0.7048	0.6825	0.0223
GNN-B	0.6995	0.6999	0.6939	0.0061
GNN-C	0.7098	0.7101	0.7065	0.0036
GNN-D	0.7112	0.7120	0.7015	0.0105
拡張入力 GNN-C	0.7391	0.7403	0.7247	0.0156
融合モデル	0.7087	0.7098	0.6959	0.0140
件数	113,094	104,052	9,042	

低下は通常時から革命状態への一致率の減少幅．標準誤差は全体と通常時で約 0.0014，革命状態で約 0.005．

最小値から 50 エポック目までの変化は，方策の交差エントロピーが CNN で 6.6%増，GNN-A で 24.9%増，GNN-C で 38.5%増であるのに対し，価値の二乗誤差は横ばいまたは改善である．4.1 節で述べた教師量の非対称，すなわち方策の教師が価値の約 4.7 分の 1 であることによると考えられる．GNN-C では方策の交差エントロピーが 38.5%増加する一方，着手一致率の低下は 0.017 にとどまる．予測が崩れたのではなく，確率の値が過信の方向へ偏っていると解釈できる．着手を最大値で選ぶ限り影響は小さいが，方策を探索の事前確率として用いる場合には不利に働く．

5.3 着手一致率

評価用データにおける着手一致率を表 4 に示す．検証用データで選んだ値と評価用データの値はほぼ等しく（GNN-C: 0.7101，0.7098），エポックの選択による偏りは見られない．

着手一致率は対戦成績よりもばらつきが小さく，対戦では区別できなかった差のいくつかが現れる．GNN-C は GNN-A を 0.0068 上回る．一方，スート混合 CNN と GNN-A の差は 0.0002 であり，一致率でも区別できない．GNN-D の GNN-C に対する差は 0.0014 にとどまる．最も大きな改善は入力情報の拡張によるもので，GNN-C で 0.0293，CNN で 0.0258 向上した．拡張入力 GNN-C の 0.7391 は全モデル中で最も高い．

5.3.1 革命状態の局面

革命が起きるとカードの強弱の順序が反転する．GNN では強弱辺の向きを反転させることでこれを表現するのに対し，CNN では格子上の配置が固定されているため，反転はスカラー入力の 1 ビットとしてのみ与えられる．この違いが現れるかを確かめるため，評価用データを革命状態の局面（9,042 件，8.0%）と通常の局面に分けて一致率を測定した．

革命状態における低下は，CNN で 0.1016，スート混合 CNN で 0.0620 と大きい．一方 GNN 系のモデルはいずれも低下が 0.025 以内にとどまり，GNN-C では 0.0036 とほ

© 2026 Information Processing Society of Japan

表 5 正解手の種類別の着手一致率

モデル	パス	単体役	複数役	階段役
CNN	0.598	0.664	0.626	0.530
スート混合 CNN	0.682	0.694	0.752	0.603
GNN-A	0.668	0.702	0.740	0.655
GNN-B	0.667	0.695	0.749	0.599
GNN-C	0.669	0.715	0.737	0.665
件数	21,750	61,411	26,659	3,274

標準誤差はパス 0.0032，単体役 0.0018，複数役 0.0027，階段役 0.0083.

とんど低下しない．スート混合 CNN と GNN-A は全体では区別できなかったが，通常時の差が 0.0030 であるのに対し，革命状態では GNN-A が 0.0367 上回る．GNN-C とスート混合 CNN でも，通常時の差は 0.0023，革命状態では 0.0607 である．革命状態の局面は全体の 8.0%に過ぎないため，この差は全体の一致率にはほとんど現れない．

拡張入力 CNN は全体の一致率を 0.0258 改善したが，革命状態での低下は 0.1061 であり，もとの CNN と変わらない．捨て札や場の役の情報を加えても低下は補償されず，この差は入力情報の不足ではなく，順序の反転を構造として扱えるかどうかに由来すると考えられる．

5.3.2 着手の種類別

表 5 に，正解手の種類別の着手一致率を示す．GNN-A と CNN の差は複数役 0.114，階段役 0.124 と，複数枚で構成される着手において大きい．一方 GNN-C とスート混合 CNN を比べると，階段役で 0.062，単体役で 0.021 だけ GNN-C が上回る反面，複数役では 0.015，パスでは 0.014 だけスート混合 CNN が上回る．なお不連続辺は階段役の欠落を表す意図で導入したが，階段役の一致率は GNN-A の 0.655 から GNN-B の 0.599 へ低下した．強弱辺と同じ伝播経路に重なることで，ランクが連続しているか否かの区別を曖昧にしたと考えられる．

複数役は同一ランクの集合であり，スート方向の平均集約が扱う対象そのものである．これに対し階段役は，スート内のランクの連続性，すなわち順序と方向を要する．革命状態の結果と合わせると，スート方向の平均集約は集合としての関係を扱うのに適し，辺の向きを持つ伝播は順序としての関係を扱うのに適する，という対応が得られる．

6. 考察

パラメータ数を統制した条件で，ゲーム規則に由来する関係を明示的に与えることが，関係を持たない CNN に対して着手一致率と対戦成績の双方を改善することを示した．ただし 5.3 節と 5.3.1 節の結果は，その改善の中身を分解して捉える必要があることを示している．

3.4 節で示したとおり，GNN-A の強弱辺はカーネル幅 3 の畳み込みと式の上で一致し，両者の差は同ランク辺の項に帰着する．実際，その項を畳み込みの枠内で与えたスート混合 CNN は，通常時の局面において GNN-A と区別できない性能を示した．したがって，CNN と GNN の差として平均の尺度で観測されたものは，グラフという機構そのものではなく，スート間の情報流の有無に帰属する．この対応は実験後の解釈ではなく，設計の段階で予測できるものであった．

これに対し，カードの強弱の順序が反転する局面では両者が分かれる．GNN は順序を辺の向きとして表現するため，反転は辺の向きの反転として構造的に扱える．CNN では格子上の配置が固定されており，反転はスカラー 1 ビットの条件付けとしてしか与えられない．この違いは，革命状態の局面における着手一致率の低下として現れた．スート混合 CNN の 0.0620 に対し GNN-C は 0.0036 であり，両者の差は革命状態に限れば 0.0607 となる．さらに 5.3.1 節で示したとおり，CNN の入力情報を増やしてもこの低下は改善されない．すなわちこれは情報の不足ではなく表現形式の問題である．

着手の種類別の分解も同じ構図を与える．スート方向の平均集約は複数役において優位であり，辺の向きを持つ伝播は階段役において優位である．集合としての関係と，順序としての関係とで，適した表現形式が異なる．

役をノードとして与える構成は，固定された格子と固定されたカーネルによる畳み込みでは表現できない．役ノードは局面ごとに個数が変わり，その存在自体が手札に依存し，可変長の集合を 1 つの単位として集約するためである．ただしその寄与は，平均得点および全体の着手一致率という尺度では検出できなかった．役ノードの効果は，5.3.2 節で示した単体役の判断と，5.3.1 節で示した革命状態における安定性という，限定された形で現れている．

一方，本研究で最も大きな改善をもたらしたのは，関係の設計ではなく入力情報の拡張であった．捨て札をプレイヤごとに区別し，場に出ている役の属性を与えるだけで，着手一致率は GNN-C で 0.0293 向上し，対戦成績は 0.133 点向上した．これらはいずれも合法手の列挙では用いていた情報であり，ネットワークに渡していなかっただけである．構造の設計を細かく変えるよりも，利用可能な情報を漏れなく与えるほうが効果が大きかった．

5.1 節で示したとおり，関係の明示的な表現は最下位を回避する方向に強く現れた．大貧民ではカード交換により前局の順位が次局の初期条件を決めるため，一度大貧民になると強いカードを失った状態から次局を始めることになる．この悪循環から抜け出せるかどうかが長期的な平均得点を左右する．手札の中にどのような役が構成できるかを把握することは，弱い手札から最大限の順位を確保する判断に寄与すると考えられる．

本研究にはいくつかの制約がある．第一に，対戦評価の検出力である．1 回の測定の標準誤差は約 0.033 であり，CNN とそれ以外の差である約 0.25 点は十分に検出できる

© 2026 Information Processing Society of Japan

が，GNN の変種どうしの差である 0.01 から 0.05 点は 8 回の反復でも判定できない．実際，3 回の反復では差があるように見えたものが，8 回では消えた．対戦成績は最終的な目的に最も近い尺度であるが，モデルの設計を比較する尺度としては検出力が足りない．本研究では着手一致率を条件ごとに分解することでこれを補った．第二に，最良のモデルでも 2.69 点であり，互角の水準である 3.0 点に届かない．教師とした Blauweregen を相手とした場合は 2.05 点であり，模倣が不完全であることを直接に示している．本研究は探索を持たない模倣学習の範囲での比較である．第三に，スート間の結合として置換対称な平均のみを検討しており，スートをチャネルとして扱う実装など，他の結合方法との比較は行っていない．

7. まとめ

本研究では，大貧民の局面をカード間の関係と，カードと役の関係を表すグラフとして表現し，グラフニューラルネットワークにより方策と状態価値を出力する評価関数を設計した．入力として与える関係を段階的に変えた 3 種類の GNN と畳み込みニューラルネットワークを，総パラメータ数を 0.8%以内にそろえ，同一の学習データと学習条件で比較した．

その結果，3 種類の GNN のうち，役をノードとして与えたものが，総パラメータ数が最小ながら，着手一致率と対戦成績のいずれでも最も高い値を示した．一方，GNN の強弱辺がカーネル幅 3 の畳み込みと式の上で一致することを示し，差が同ランク辺の項に帰着することを明らかにした．実際にその項を畳み込みに与えたスート混合 CNN を対照に置くと，通常時の局面では両者を区別できない．平均の尺度で CNN と GNN の差として観測されたものは，グラフという機構ではなくスート間の情報流に帰属する．

これに対し，カードの強弱の順序が反転する革命状態の局面では両者が分かれる．順序を辺の向きとして表現し，反転をその向きの反転として扱う GNN は革命状態でも精度を保つのに対し，反転を 1 ビットの条件付けとして与える CNN は大きく精度を落とす．この低下は CNN の入力情報を増やしても改善されず，情報の不足ではなく表現形式に由来する．着手の種類別の分解も同じ構図を示し，スート方向の平均集約は集合としての関係に適し，辺の向きを持つ伝播は順序としての関係に適する．グラフによる表現が畳み込みで代替できないのは，この順序と方向に関わる部分である．

加えて，構造の設計変更よりも，ネットワークに与える入力情報を増やすほうが大きな改善をもたらした．捨て札をプレイヤごとに区別し，場に出ている役の属性を与えた拡張入力 GNN-C は，着手一致率 0.7391，対戦成績 2.689 点で，比較した 9 モデル中で最も高い．

今後の課題として，入力情報をさらに拡張すること，状態価値を葉評価に用いる探索との組み合わせ，および方策の教師を複数のクライアントに拡張することが挙げられる．また，GNN において学習の後半で方策の損失が上昇する現象について，強化学習を併用することでその課題を克服できるかを試行する必要がある．

参考文献

[1] 電気通信大学：UECda コンピュータ大貧民大会，`https://flute.u-shizuoka-ken.ac.jp/daihinmin/2023/document_rules.html`（2026 年 7 月閲覧）．

[2] 大渡勝己，田中哲朗：方策勾配を用いた教師有り学習によるコンピュータ大貧民の方策関数の学習とモンテカルロシミュレーションへの利用，情報処理学会研究報告ゲーム情報学（GI），Vol.2016-GI-35，No.10，pp.1–8，2016.

[3] 福澤詩音，松崎公紀：大貧民におけるポリシー・バリューネットワーク評価関数の学習法と利用法の検討，情報処理学会第 88 回全国大会講演論文集，2026.

[4] Ben-Assayag, S. and El-Yaniv, R.: Train on Small, Play the Large: Scaling Up Board Games with AlphaZero and GNN, arXiv:2107.08387, 2021.

[5] Rigaux, T. and Kashima, H.: Enhancing Chess Reinforcement Learning with Graph Representation, *Advances in Neural Information Processing Systems 37* (NeurIPS 2024), 2024.

[6] Schlichtkrull, M., Kipf, T. N., Bloem, P., van den Berg, R., Titov, I. and Welling, M.: Modeling Relational Data with Graph Convolutional Networks, *The Semantic Web (ESWC 2018)*, pp. 593–607, 2018.

© 2026 Information Processing Society of Japan