

人工知能におけるゲームの役割

Murray Campbell

IBM T.J. Watson Research Center
mcam@us.ibm.com

翻訳：丸山 宏

アイ・ビー・エム ビジネスコンサルティングサービス（株）
maruyama@jp.ibm.com

AI研究のテストベッドとしてのゲーム

人工知能研究の多くの領域において、ゲームは大変よいテストベッドとなり得る。ゲームは複雑な実世界を抽象化して見せるので、問題解決器 (problem solver) の研究においては、知識表現と探索におけるクリティカルな問題だけに意識を集中することができる。この事実は、計算機の黎明期から知られていて、たとえば Claude Shannon はコンピュータチェスに関する彼の 세미나⁴⁾の中で研究の対象としてチェスを選ぶ理由について以下のように述べている。

- (1) チェスでは、問題が、許される操作 (駒の動き) と最終的なゴール (チェックメイト) の双方において、きわめて明確に定義されている。
- (2) 問題がやさしすぎもせず (自明でなく)、まともな解が期待できないほど難しすぎもしない。
- (3) チェスを上手に指すには、「思考」が必要とされていると一般に信じられている—すなわち、この問題が機械で解ければ、我々は機械が思考することを認めるか、あるいは「思考」の定義を狭めなければならない。
- (4) チェスの離散的な構造は、現代のコンピュータのデジタル処理によく合っている。

現実世界の複雑さを考えるとき、AI研究のテストベッドとしてのゲームはさらに魅力的である。なぜなら現実世界の一般の問題解決では、以下のような難しさがあるからである。

1. (しばしば互いに矛盾する) 複数のゴールが存在し、それらのゴールは明確になっていないことが多い。
2. すべての可能な行動が明示的に示されていることはまれだし、それらは時とともに変化し、場合によっては事実上無限の選択肢があったりする。
3. 関連する情報が隠されていたり、多すぎて、すべてを受け入れ、処理することができなかつたりする。
4. 他の登場人物の意図や行動を認識したり予測したりすることが難しい。
5. 現実世界と交流するためのセンサやエフェクタを作成することはそれ自体で大きな技術的チャレンジである。

ゲームにおいては、これらの条件に対して、仮定を置くことによって簡単化することができ、1つか2つの重要な要因に対してのみ集中的に研究できる。そうでなければ、問題は難しすぎてどこから手を付けてよいのかも分からなくなってしまうだろう。

本稿では、なぜゲームがAI研究の優れた領域であったか、また今後もあり続けるかについて考える。知能の異なる側面に対して異なるゲームがあり、それぞれ



図-1 1997年の試合において、Deep BlueがGarry Kasparovを破った

がそれぞれの進歩を促すことを示す。また、史上初めて、規定試合において人間の世界チャンピオンを破った Deep Blue については、その探索と評価関数を重点に、より詳しく述べる。最後に今後の研究エリアについて考える。

ゲーム・システムにおける研究対象

AI研究においては、多岐にわたるゲームが研究されている³⁾。これらのゲームにおいて、高いパフォーマンスを得るには、探索の制御、実時間の意思決定、不完全情報の利用、ランダム性、学習、マルチエージェント環境など、さまざまな領域の研究が必要となる。

探索制御、すなわちゲームの状態の先読みを組織的に行うための手法は、ほぼすべてのゲームにおいて、重要な要素である。ゲームの探索空間は通常、盲目的な探索を行うには大きすぎるからである。たとえば、見込みのありそうな手をより深く探索する選択探索は1つの方法である。また、間接的な手法として、複数のゲーム状態を1つに畳み込む状態抽象化も探索の複雑さを減らす効果的な手法である。

実時間の観点も、ほとんどのゲームに現れる。限られた時間の中で、適切な決定を下すことは重要な研究エリアであり、数多くのテクニックが開発されている。たとえば、探索の深さを徐々に大きくしていく反復深化(iterative deepening)は、どこで探索が打ち切られてもリズナブルな決定をすることを可能にする。場合によっては、意思決定理論(decision theory)的なテクニックを使って、明示的に、追加の計算のコストとそれによって得られる利益について推論することもあり得る。

不完全情報ゲームでは、情報はプレーヤから隠されて

いる。このようなゲームでは、ゲームが進むにつれ得られる情報から、隠されている情報を推察するのが普通である。たとえば、モンテカルロ・シミュレーションによって、適切な指し手を見つける、などの戦略を使うことができる。

ランダム性(randomness)は、ゲームにおいて、探索制御の問題の一部ととらえることができる。ここでは、未来の指し手に確率をつけることで、盤面の評価を行う。モンテカルロ・シミュレーションはそのための有効なテクニックである。

学習を行うことによって、ゲーム・プレーヤを強くすることは、古くから重要な研究エリアであり、1950年代のArthur Samuelによるチェッカーのプログラムにまでさかのぼることができる。多くの多様なテクニックが開発され、一部は成功し、その他はそうでもない。しかしながら、学習のみによって、何も無いところから有能なプレーヤに育てることは、きわめて狭いドメインだけでしか、今のところ成功していない。

3人以上のプレーヤによるゲームは、さらにゲームに複雑さを加える。ここでは、標準的な探索手法は実用的でないことが多く、むしろ、他のプレーヤのモデルを作って動きを予測し、その結果として探索空間を削減することが重要になる。

事例：チェス・システム 「Deep Blue」

コンピュータ・チェス Deep Blue¹⁾は、1990年代の中期にIBMの研究所で開発された。1997年に、Deep Blueは、規定試合で初めて人間のワールドチャンピオンを破ったコンピュータとなった(図-1)。

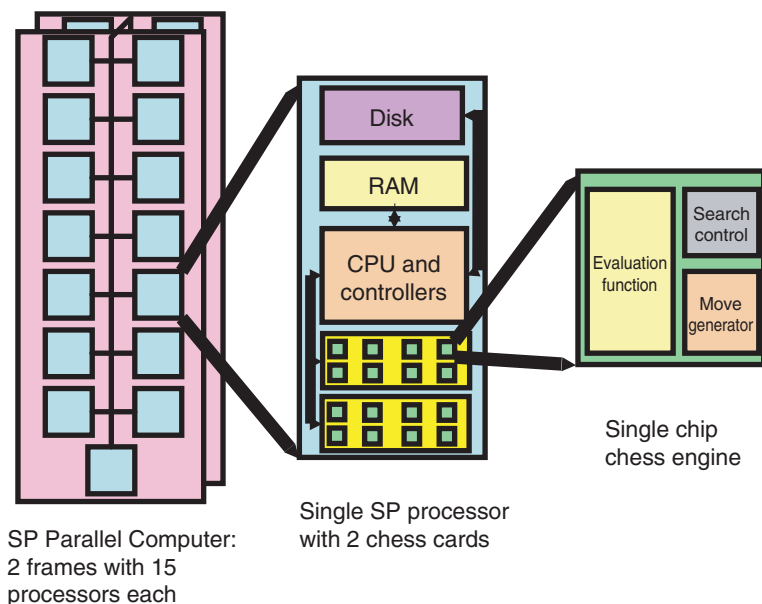


図-2 Deep Blue システムの構成

最も自明な Deep Blue の特徴は、その計算パワーにある。Deep Blue は、30 プロセッサの IBM RS/6000 SP コンピュータと、480 個のシングルチップのチェス探索エンジンとからなる。SP コンピュータ（図-2 の左端）は、15 個の汎用コンピュータを持つ 2 つのフレームからなる。これらの各プロセッサは標準的な SP コンピュータの構成に加えて、8 個のチェス探索エンジンチップを搭載した、2 枚のマイクロチャンネルカードを搭載している（図-2 の右端）。1 つの探索エンジンチップは、1 秒間に 200 万以上のチェスの盤面を検査できる（図-3）。

Deep Blue は、プロセッサの静的な木構造を用いて並列性を管理している。1 つの SP プロセッサがマスターとなり、残りの 29 個の SP プロセッサをコントロールする。マスターは、ルートに一番近いレベルの探索木の処理を行う。その結果、マスターは仕事をワーカ・プロセッサに分割する。それぞれのワーカは、さらに深いレベルの探索木の処理を行い、結果をシングルチップの探索エンジンに渡す（図-4）。

システム全体での Deep Blue の処理速度は、毎秒 1 億 2,600 万盤面に達する。これは、1 秒に 1 から 2 盤面のスピードで考える、人間のグランドマスターに比べると過酷なまでの差である。明らかに、人間の思考過程は Deep Blue のアプローチとまったく異なるのである。

コンピュータの持つ大変な速度上のアドバンテージにもかかわらず、単純な全数探索 (brute force) のアプローチでは、人間のグランドマスターを破ることはできなかっただろう。その理由の 1 つは、チェスや他の多くの

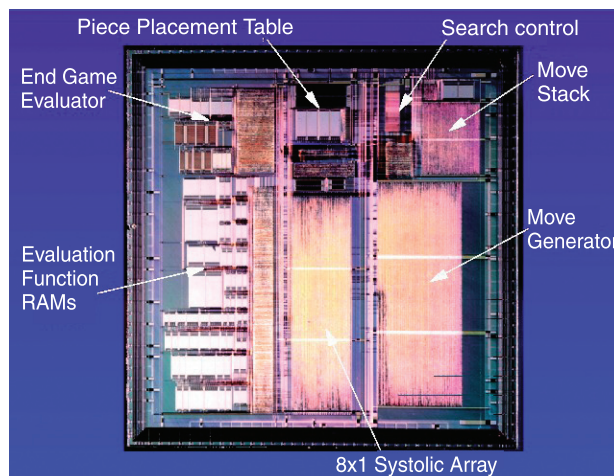


図-3 Deep Blue に使われている、シングルチップ・チェス探索エンジン VLSI

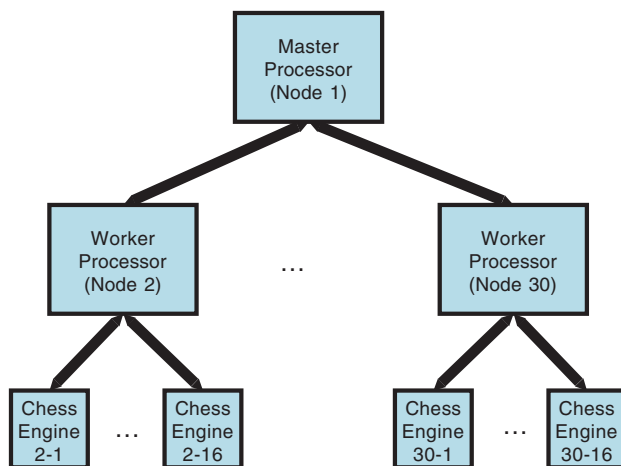


図-4 Deep Blue は、3 レベルの階層構造を持つ

ゲームでは、全数探索を行うと、探索の深さに対して指数関数の手間がかかることである。このことが、機械による探索の深さを制限している。一方、人間のグランドマスターは、全数探索に基づく探索システムが到達できると思われる深さをはるかに超えて先読みをすることができる。ここでの違いは、人間が選択的探索 (selective search)、すなわち探索木の場所によって、探索の深さを変えていることである。選択的探索を正しく行えば、より「重要な」探索の枝を深く読み、そうでない枝については探索の深さを浅くすることができる。図-5はこのアイデアを示している。図の左側では、全数探索はあらかじめ決められた一定の深さまでしか先読みできないことを示している。右側は選択的探索であり、探索中に見つかった重要な局面 (図中☆で示されている) に関して、優先的に計算資源を割り当て、より深く先読みする。これによって、選択的探索では、重要度の高い探索枝をより深く読むことができる。

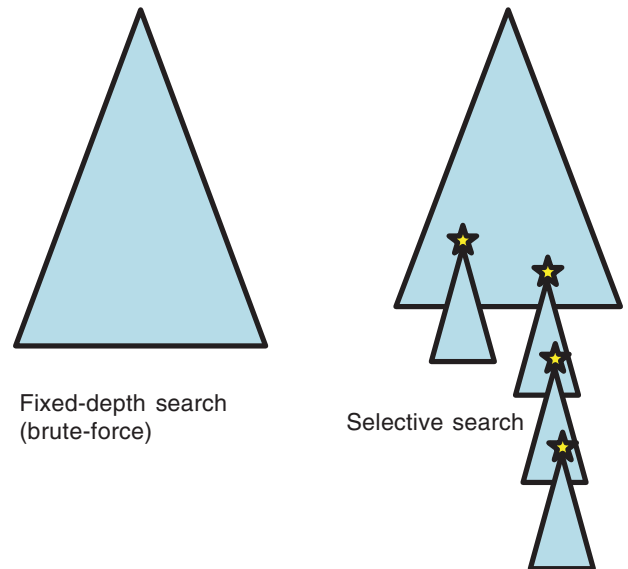


図-5 全数探索と選択的探索

Deep Blueにおいては、選択的探索の改善に大きな努力が費やされた。選択的探索におけるアプローチの1つは、興味深い枝をより深く伸ばすことであり、それと対になるもう1つの作業は、つまらない枝を早めに切り取ることである。Deep Blueにおいては、選択的探索を伸張戦略のみによっていて (早期の枝狩りはリスクが大きすぎると考えた)、「遅延伸張を伴う双対評価」(dual credit with delayed extensions) と呼ぶ新しいアルゴリズムに基づいている。このアルゴリズムは、以下の原則を持つ。

1. 強制された指し手 (forced moves) を伸ばす。人間が探索に用いる最重要な戦略は、強制された指し手の列をより深く読むことである。強制された指し手の列とは、きわめて限られた手しか、片方あるいは双方のプレーヤーにとって意味のない (すなわちそうでない手はただちに負けにつながる) ような、指し手の列である。我々は、このような列を発見するのに、我々が開発した特異手拡張 (singular extensions) アルゴリズムに基づいた。
2. 強制された指し手は期待値に依存する。もし、プレーヤーがその手によって有利になると期待すれば、その指し手は強制された指し手と考える。もし、その手によっても盤面の有利性が変わらなければ、強制されたとは考えない。
3. 部分的な伸張。強制された列をすべて完全に伸張するのは、探索の爆発を招くので通常うまくいかない。したがって、指し手列の「強制の度合い」(forcedness) を計算し、それに基づいて伸張の深さを制御する。
4. 伸張の遅延。個別の強制された手よりも、強制された

指し手の列の方がはるかに興味深い。したがって、強制された指し手を伸張する前に、その先にさらに強制された差し手が続くかどうかを判断することは大変効果がある。

5. 双対評価。チェスは2プレーヤーによるゲームであり、列によっては双方が強制された指し手になってしまうこともある。双方の評価を考えることによって、さらに探索空間を狭めることができる。
6. 探索済みの空間を保持する。探索済みの盤面を注意深く管理していないと、ある状況の下では、選択的探索は、ループに陥ることがある。これらの盤面は、ハッシュ表に記録される。

この結果得られた探索は、高度に選択的であった。Deep Blueが全数探索を行った場合、探索の深さはおおよそ14層 (ply, 1層は白または黒の1つの指し手) であったと思われる。一方、選択的探索によって、すべての可能な手について12層の先読みが保証されたのに加えて、重要な指し手列については30層から40層の先読みが可能になった。このことがDeep Blueの1997年の勝利に貢献した重要な要素になっていたはずである。

選択的探索に比べて見えにくいだが、同様に重要な要素として、Deep Blueの評価関数 (evaluation function) がある。評価関数は、盤面が与えられたときにその盤面の良さを表すスカラー値を返す。正の値は白にとって有利な盤面を示し、負の値は黒にとっての有利な盤面を表す。Deep Blueの評価関数は、おおよそ8,000のパターンを認識するかなり複雑なものである。それぞれのパターンには、評価の元となる値が付加されている。

Deep Blueの評価関数が複雑であることは、良い面も悪い面もある。チェスの盤面の評価に関するさまざまな側面を考慮することによって、この評価関数は盤面の良さをより正確に与えることができる。一方、それぞれのパターンに適切な値を付与することは、きわめて難しい問題である。特に、盤面のさまざまな状況に応じて、この値が変わることに問題がある。我々は自動的にこれらの値を調整するための多くの手法を試してみたが、限定的な成功しか得られなかった。依然として、チェスにおける複雑な評価関数を調整する最良の方法は、解決されていない研究テーマということができよう。人間のグランドマスタが持つ「評価関数」は、多くの場合においてDeep Blueよりも優秀なのである。

人工知能に関連して、Deep Blueのもう1つの興味深い面は、「拡張定石」(extended book)である。拡張定石は、過去の人間のグランドマスタの経験を取り入れるためのメカニズムである。すなわち、70万にも及ぶ過去のグランドマスタ戦の記録を処理し、それらに現れる盤面のうちいくつかのものについて、グランドマスタがその盤面で指した手について、その評価値を推定した。この推定には、その手が同じ盤面で何回指されたか、そのプレーヤーはどのくらいの強さだったか、そしてその手を指した結果をどうだったかなど、多くの要素が考慮されている。Deep Blueはこれらの結果によって、より評価の高い手を選好するなど探索空間に重みを与えるために使うことができる。もちろん、Deep Blueが今まで指されたことのない、新しい良い手を発見した場合には、そちらを指すこともある。

まとめると、Deep Blueは、人間のグランドマスタとは大きく違うやり方でチェスを指しているといえる。Deep Blueは人間のプレーヤーほど洗練された評価関数を持っていないが、それを驚異的ともいえる徹底的で深い探索によって補っている。これを、他のさまざまなテクニックと組み合わせることによって、Deep Blueは世界クラスレベルのチェスを指すことができるのである。

今後の方向性

PC上でのチェス・プログラムは、着実に進化している。Deep FritzやDeep Juniorなどの商用プログラムはいまや世界トップのグランドマスタたちに挑戦しようとしている。一方、ゲームの研究そのものの焦点は、多くの他の分野に移りつつある。

有効な学習システムを作ることは、多くのゲームにとって重要なチャレンジである。たとえばチェスにおいては、グランドマスタの過去の棋譜を調べることや強化学習のアプローチによって、評価関数の特性値 (feature

value) を学習することが試みられてきた。評価関数をまったくのゼロから学習によって得ることにはまだ多くのやるべきことがある。特に、新しい特性値を発見することには難しい (Gerry TesauroのTD-Gammonは特筆すべき例外である)。

探索の領域においては、分岐数 (各盤面において可能な手の数) の大きいゲームはコンピュータにとって難しいことが分かっている。今までチェスやチェッカーのようなゲームで成功した探索のテクニックは、これらの分岐数の大きいゲームにはうまく適用できない。高度に選択的な探索、すなわちより人間のやり方に近い探索が、これらのゲームに関しては必須である。

不完全情報やランダム性を持つゲームはまったく別の問題を持つ。コントラクトブリッジのようなカードゲームでは、Matthew GinsbergのGIBが、伏せられたカードを扱うための洗練された手法を持っており、世界クラスのレベルに近づきつつある。別のカードゲームであるポーカーでは、はったりを扱うために相手方のモデリングが重要な要素となる。Alberta大学のグループは、やはり世界レベルに到達しつつあるプログラムを開発している。

チェスや碁などの、伝統的な戦略ゲームの研究が進む一方、人工知能を用いたゲームの研究の将来は、対話的ゲーム (interactive game) の方向に進むと見られる²⁾。対話的ゲームをよりリアルに、また面白くするために知的な登場人物や複雑な仮想世界を作ることには大きな利点がある。対話的なコンピュータゲームのそれぞれのジャンル (アクション、ロールプレイング、アドベンチャー、戦略、スポーツ) にはそれぞれ特有のチャレンジがある。人工知能の研究が貢献できる場面としては、それらしい戦術を持つ敵、味方、他のキャラクタなどがある。対話的ゲームは、当面、人工知能研究者にとって、さまざまな研究テーマを提供する場となろう。

謝辞 Deep Blueの共同研究者である、Feng-hsiung HsuとA. Joseph Hoane, Jr.に感謝する。

参考文献

- 1) Campbell, M., Hoane, A.J. and Hsu, F.H.: Deep Blue, Artificial Intelligence, Vol.134, pp. 57-83 (2002).
- 2) Laird, J.E. and van Lent, M.: Human-level AI's Killer Application: Interactive Computer Games, AAAI National Conference on Artificial Intelligence (2000).
- 3) Schaeffer, J.: A Gamut of Games, AI Magazine, Vol.22, No.3, pp.29-46 (2001).
- 4) Shannon, C.: Programming a Computer for Playing Chess, Philosophical Magazine, Vol.41, pp.256-275 (1950).

(平成15年10月5日受付)

