

## 競争巡回形知能ロボット制御用先読エンジンの設計

Design of the Search Engine for Competitively Traveling  
Intelligent—Robot

苫米地 宣裕

Nobuhiro Tomabechi

八戸工業大学

Hachinohe Institute of Technology

キーワード：競争 (competition), 巡回 (traveling), 知能 (intelligent),  
ロボット (robot), 先読 (search), エンジン (engine)

連絡先：〒031-8501 八戸市大字妙字大開 8 8—1 八戸工業大学 苫米地 宣裕

Tel 0178-25-8051, Fax 0178-25-1430, E mail: tomabech@hi-tech.ac.jp

### 1. はじめに

本研究では、自分と利益の反する相手があり、相手と争いながら何事かを自律的に実行するような種類のロボット（対戦ロボットという）を対象とする。対戦ロボットに特徴的な課題は、相手に勝つために、相手の行動を予測し適切な対策を選択するという作業（先読みという）にある。先読みには、相手の行動を読む、作戦を立てる、知識を生かすなど人工知能的課題<sup>(1)</sup>が多く含まれている。一般に、先読みは読みの深さが深くなると処理時間が飛躍的に大きくなる<sup>(2)</sup>。しかも、対戦ロボットにはリアルタイム応答が要求される。本研究では、対戦ロボットに共通する先読み

機能を並列ハードウェアによって高速に実行できるようなプロセッサ（先読エンジンという）の構成法を検討する。

先読エンジンに要求される機能は、先読みという点では共通性はあるが、具体的な点是对戦の内容によって異なる。本研究では、ある区域に散在する目標物を自律的に移動しながら争って収集するロボット（競争巡回形知能ロボットという）について検討する。対戦の内容は、結局、一種のゲームとしてモデル化されることになる。また、実ロボットの動作を取り扱うには、ロボットの動力学モデルを明らかにする必要がある。本稿では動力学モデルは省略し、もっぱら、ゲームモデルのみを取り扱うこととする。

本稿では、まず対戦のゲームモデルを示し、次

に本ゲームモデルに適した最善手探索方法として、目標物に直行する動作を単位とする探索方法を提案している。また、多数のエージェントプロセッサを用意し、先読みを分担して実行するという先読エンジンの構成法を提案している。

## 2. 対戦のゲームモデルと性質

### 2. 1 ゲームモデル

本研究では、ある区域内の目標物を競争しながら収集するロボットについて検討する。このとき、区域の広さ・形状、障害物の有無、ロボットの機能（進行速度、方向転換に要する時間、双方のロボットの機能が同じか異なるか）、目標物の価値・分布などを明らかにする必要がある。本稿では、以下のようにゲームモデルを設定する。

[ルール1] 区域は $M \times N$ のマス目で表わされる。

[ルール2] 区域内に障害物はない。

[ルール3] 双方のロボットの機能は等しい。

[ルール4] ロボットの位置、および、目標物の位置はマス目で表わされる。1個のマス目には、一時に1個のロボットだけが存在できる。

[ルール5] ロボットは一つのタイミングで、現在位置のマス目の8近傍の任意のマス目に移動できる。

[ルール6] 2つのロボットが同一のマス目に入ろうとすると、いずれか一方が優先権を有する（優先権を有する側を先手、その反対側を後手という）。

[ルール7] 目標物の数 $B$ は変数とする。

[ルール8] 目標物は価値を表わす点を有する。

[ルール9] 目標物の分布はランダムとする（1回の対戦ごとに乱数によって決定する）。

[ルール10] ロボットが目標物の置かれたマス目に入るとその目標物の価値点はそのロボット

の得点に加算される。そして、目標物は削除される。

[ルール11] 2つのロボットのスタート位置は、相互に対称な最も離れた2つのマス目をとる。

[ルール12] ある回数の対戦を行い、合計得点差 $> H$ のとき先手の勝ち、そうでないときは後手の勝ちとする。ただし、 $H$ はハンデキャップの値を表わしており、多くの対戦結果に基づいて適切に定める。

図1に、領域が $8 \times 8$ 、目標物の数 $B = 6$ の場合の初期局面の例を示している。図において、 $A$ は先手ロボット、 $B$ は後手ロボット、分数式 $n/m$ は目標物を表わしている。 $n$ は目標物の番号を、 $m$ はその価値点を表わしている。

### 2. 2 ゲームの数学的性質

はじめに用語を定義する。

[定義1] 距離：地点 $P$ から地点 $Q$ に至る最短経路に含まれるマスの数を $P$ と $Q$ の距離といい、 $|P - Q|$ と表す。

[定義2] 勢力圏：「ロボット $A$ との距離」が「相手との距離」より大きいマスの集合をロボット $A$ の勢力圏という。2つのロボットからの距離が等しいマス目は先手側の勢力圏に含める。

[定義3] 直面目標：ロボットの現在位置からある目標物までの最短経路中に他の目標物が存在しないような目標物を直面目標という。（言いかえると、他の目標物の位置を経由してその目標物に到達すると、直行するよりも距離が長くなるような位置にある目標物を直面目標という）。

[定義4] 非切迫度：ある目標物について（自分との距離－相手との距離）をその目標物の非切迫度という。

本ゲームについて次のような性質が成り立つ。

[定理1] ロボットを $A$ 、 $B$ 、2つの目標物を $X$ 、 $Y$ と表す。 $X$ 、 $Y$ ともに $A$ の勢力圏内にあり、か

つ、Aに対してもBに対しても直面目標であると  
する。いま、 $|A-Y| > |A-X|$ とする。こ  
のとき、次の条件が成り立つならば、BはX、Y  
のいずれか一方を取ることができる。

$$|A-X| + |X-Y| > |B-X| \quad \text{AND} \\ |A-X| + |X-Y| > |B-Y|$$

(証明) AがXに直行するならば、BはYに直行  
しAより先にYに到着できる。AがYに直行する  
ならば、BはYに直行しAがYの地点に入ったと  
きXに直行すれば、BはAよりも少なくとも1タ  
イミング早くXに到着することができる。□

[定理2] ロボットをA、B、2つの目標物をX、  
Yと表す。X、YともにAの勢力圏内にあり、かつ、  
Bに対してYは直面目標であるがXは直面目  
標でない(BはYを経由してXに到着する)とす  
る。このとき、AはX、Yの両方を取ることがで  
きる。

(証明) AがまずYに直行し次にXに向かえば、  
AがY、Xの両方にBより先に到着できるのは明  
白である。□

定理1、定理2は、相手の勢力圏内にある目標  
物でも、配置によっては一部を取ることができ  
ることを示している。

### 3. ゲームを進める指針と最善着手の 探索

#### 3.1 ゲームを進める指針

次のような指針でゲームを進めるのが有効と  
考えられる。

[指針1] 一つの直面目標を選択しそれに向かっ  
て最短距離で直行する。

[指針2] 目標を決める優先順位は次のようにす  
る。

- ① 双方の勢力圏の境界線付近にあるもの。
- ② 自分の勢力圏内にあり、かつ、評価点の高い

もの。

- ③ 自分の勢力圏内にある目標物の価値の合計点  
が大きな値になるような位置にあるもの。

#### 3.2 最善着手探索アルゴリズム

本ゲームモデルでは、ロボットの動作の最小単  
位は1マスの移動となるが、この1マス単位の探  
索を行うと、探索空間が広くなり先読みに要する  
時間がかかり過ぎるので現実的でない。そこで、  
目標物へ直行する動作を単位とする探索方法を  
検討する。まず、着手を次のように定義する。

[定義5] 着手：ある目標物に向かって最短距離  
で直行することを着手という。

着手を目標物に直行する動作を単位とすると、  
着手の回数は目標物の個数Bと等しくなる。着手  
の回数をBとすると悉皆探索を取った場合の探  
索の数は $B!$ となる。従って、Bが大きくなると  
悉皆探索は非現実的となる。本ゲームモデルには、  
将棋の詰みや連珠の5連のような「突然の終局」  
がないので、探索方法としては $\alpha - \beta$ 探索<sup>(1)、(2)</sup>  
が有効となると考えられる。ただし、上記のよう  
に目標物に直行する動作を単位とすると、先手、  
後手が常に交互に着手するとは限らない(相手の  
1回の着手に対し、その着手の到達距離以内なら  
ば何回でも連続着手できる)。また、着手の変更  
があり得る。このように変形された $\alpha - \beta$ 探索を、  
以下、「非交互着手 $\alpha - \beta$ 探索法」とよぶことと  
する。そのアルゴリズムは以下のようになる。た  
だし、 $\alpha$ 、 $\beta$ 値による枝刈りに関する記述は、通  
常の $\alpha - \beta$ 探索と同様なので省略している。

[アルゴリズム1] 非交互着手 $\alpha - \beta$ 探索法

(手順1)  $M \leftarrow \infty$

(手順2) 探索の深さが指定された値になるか、  
または、目標物がなくなったときは、局面評価を  
行い、かつ、探索の深さを1レベル戻して手順6  
へ、さもなければ手順3へ。

(手順3) 着手の候補をリストアップする。着手の対象とする目標物は直面目標とする。ただし、相手が同一目標物に対して着手しており、かつ、その目標物が相手の勢力圏内にある場合は着手候補としない。

(手順4) 一つの着手候補(優先順位の最も高いもの)を選択し着手する。

目標に到着するまでの距離をカウントする。自分の着手の距離が相手の着手の距離を超えたときは手順5へ、さもなければ手番を交代せずに、探索の深さを1レベル深くして手順2へ。

相手が同一目標物に対して着手しており、かつ、その目標物が自分の勢力圏内にある場合は、相手の着手を修正させる。すなわち、自分が目標物に到着した時間に相手が進行した位置で停止するという着手に修正させる。

(手順5) 手番を交代し、探索の深さを1レベル深くして手順2へ。

(手順6) 着手を戻す。

先手番なら $M < \text{局面評価値}$ 、後手番なら $-M < \text{局面評価値}$ のとき、 $M \leftarrow \text{局面評価値}$ 、かつ、最善着手候補 $\leftarrow$ 着手。

(手順7) 未探索の着手候補があれば手順4へ、さもなければ手順8へ。

(手順8) 根局面ならば終了、さもなければ探索の深さを1レベル戻して手順5へ。 □

#### (1) 局面評価

手順1における局面評価は先読みの性能に大きな影響を与える。本ゲームモデルでは、局面評価に寄与するファクタとして次が考えられる。①自己の既得点、②相手の既得点、③自分がこれから得点できる可能性、④相手がこれから得点できる可能性。

本稿では、局面評価を次式のように行う。

局面評価値 = 自己の既得点 +  
自己の勢力圏内にある目標物の点の合計値

- (相手の既得点 + 相手の勢力圏内にある目標物の点の合計値)

#### (2) 着手の優先順位

手順3において着手の優先順位は次のように決める。

① まず、自分の勢力圏内にある直面目標に対する着手で(価値 $\times$ ( $P - \text{非切迫度}$ ))の大きいものを優先する。ただし、 $P$ の値は経験に基づいて定める。

② 次に、相手の勢力圏内にある直面目標に対する着手で(価値 $\times$ ( $P + \text{非切迫度}$ ))の大きいものを優先する。

なお、探索は着手単位で行うが、ロボットの移動はあくまで1マス単位で行う。従って、1マス移動するたびに移動した位置を基準に探索をし直すこととなる。

図2に、図1の局面に対応する先読み探索のゲームの木を示している。ただし、一部のみを示している。図において、○は先手のノード(着手)を、□は後手のノードを表わしている。□□と同じ記号が連続するのは、□側の連続着手を示している。ノードの中の分数式は、分子で目標物の番号を、分母で得点を表わしている。5' / 0は、5番の目標物に向かって着手したが相手の方が先に到着したので途中で進路を変更したことを表わしている。

## 4. 先読エンジンの基本構成

非交互着手 $\alpha - \beta$ 探索アルゴリズムにおいて、処理時間は、主として次によって決定される。①着手候補リストアップと優先順位決定に要する時間、②局面評価時間、③探索のくり返し回数。従って、先読みの処理速度を向上させるためには、次のような方策が有効となると考えられる。①着手候補リストアップ、着手候補の優先順位決定、

および、局面評価をハードウェアによって行う、  
②前項のハードウェアから構成される部分探索ハードウェアを多数用意し全体の探索を並列に実行する。ただし、全体の探索を並列に分割する方法としては、枝刈りがあるため、先頭の着手ごとに分割し一つの着手は独立のハードウェアで処理しさらなる分割はしないとするのが有効となる。

本稿では、以下のような先読エンジンの構成法を提案する。

先読エンジンを、ある深さの探索を独立に実行するエージェントプロセッサ（APという）複数と各APに探索局面を割り当てる主プロセッサで構成する。本構成を図3に示している。

1個のAPは、探索制御回路、フィールド回路、および、局面評価回路から構成される。探索制御回路は、① フィールド回路、優先度計算回路、局面評価回路の制御、② 最善手の決定の機能を有する。フィールド回路はゲームの行われる区域に対応づけて構成され、① ロボットの移動、② 目標物の消去、回復、③ 直面目標の検出の各機能を有する。APの構成を図4に示している。

主プロセッサは次の機能を有する。

- ① 初期設定、② 局面の表示、③ ゲームの進行、④ APの機能、⑤ 各APへの探索局面の分配

本先読エンジンは次のように動作する。

- ① 主プロセッサは、1～2手先までの可能な着手候補をすべて求める。
- ② 主プロセッサは、各APに着手候補を割り当てて探索を命令する。
- ③ 各APは、与えられた着手候補について指定された深さまで探索する。探索を終えたAPは主プロセッサに指定された着手の評価値を知らせる。
- ④ 主プロセッサは空いたAPが生ずれば、別な着手候補を与えて探索を命ずる。

- ⑤ 主プロセッサは、すべての着手候補についての評価値が得られたら、最善着手を決定する。

## 5. むすび

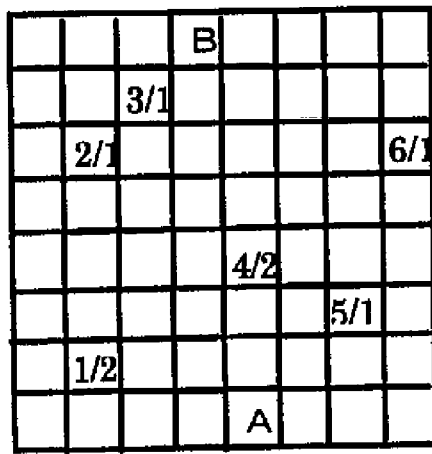
本稿では、ある区域に散在する目標物を自律的に移動しながら争って収集するロボット（競争巡回形知能ロボット）を取り上げ、対戦ロボットに共通する先読み機能を並列ハードウェアによって高速に実行できるプロセッサ（先読エンジン）の構成法を検討した。まず、対戦のゲームモデルを設定した。次に、本ゲームモデルに適した最善手探索方法として、目標物に直行する動作を単位とする探索方法（非交互着手 $\alpha - \beta$ 探索法）を提案した。また、多数のエージェントプロセッサによって先読みを分担して実行する先読エンジンの基本的な構成法を提案した。

今後、次のような課題について検討する予定である。

- ① シミュレーションソフトウェアを作成し、非交互着手 $\alpha - \beta$ 探索法の有効性を検証する。
- ② エージェントプロセッサと先読エンジン全体の詳細な論理設計を行う。
- ③ 先読エンジンを集積したVLSIチップの設計・製作を行う。

## 参考文献

- (1) 人工知能学会編：人工知能ハンドブック、オーム社 (1990).
- (2) 小谷、吉川他：コンピュータ将棋、サイエンス社 (1990).



A : 先手ロボット、B : 後手ロボット、  
 n/m : 目標物の番号/価値点

図1 ゲームの初期局面の例

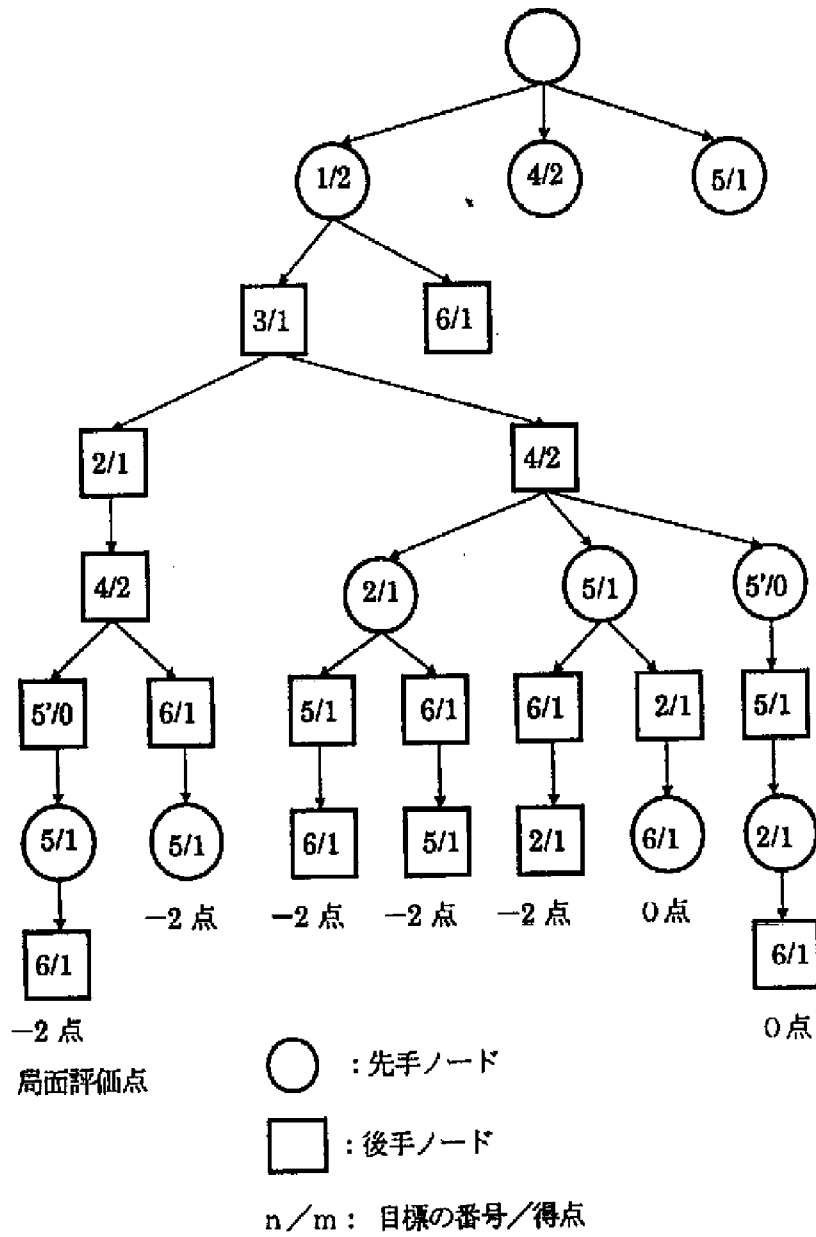


図2 ゲームの木 (図1の局面に対応)

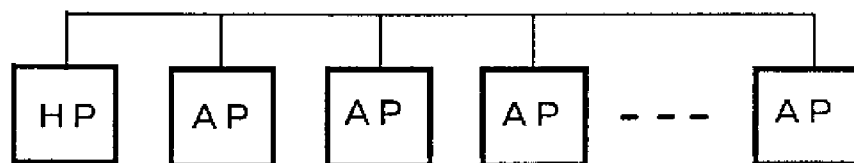


図3 先読エンジンの基本構成

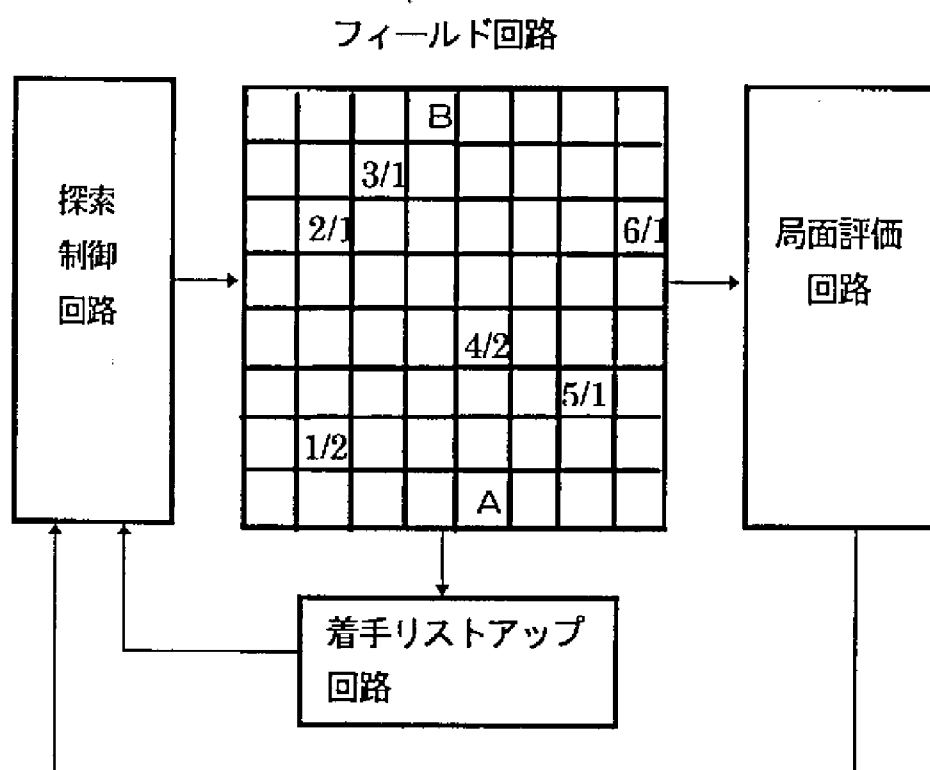


図4 エージェントプロセッサの構成