人 工 知 能 86－7
（１９９3. 1. ２０）

インターリーブ・プランニングとその実験的評価

礒田佳徳　山田誠二　豊田順一

大阪大学産業科学研究所

isoda@ai.sanken.osaka-u.ac.jp

動的世界に対処することを指向したリアクティブ・プランニング実現の一手法として，常に１ステップのプランニング，実行を繰り返すインターリーブの方法がある．これに対して，我々は先に、外界の変化に依存したプラン成功確率を定義し、これを基にプランニング／実行を切り換えるタイミングを決定する方法を提案した。本報告では、インターリーブ・プランニング全体のアルゴリズムを提案する．また，タイル・ワールドという現実世界を反映した領域で実験を行い，インターリーブ・プランニングの有効な問題のクラスを探る．

Interleave Planning and Experimental Evaluation

Yoshinori Isoda　Seiji Yamada　Jun'ichi Toyoda

I.S.I.R, Osaka University
Mihogaoka 8-1, Ibarakai, 560 Osaka, Japan

For making a planning system adaptive to the dynamic world, various methods on reactive planning have been proposed. An interleave approach is a general framework in which planning and an execution/observation are mutually switched. We have proposed the method for determining the timing to switch planning into execution by the success probability of a plan. In this paper, we propose the whole algorithm for interleave planning. Furthermore, we report the experimental results that describes the class in which the interleave planning works better than the most reactive planner (interleaving planning into execution step by step) and the conventional planner for the static world. The experiments were made in the simplified tile world and we found out some classes in which the interleave planner is so effective.

1．はじめに

プランニングもオペレータという規則を用いた一種の推論を行なっている．では，プランニングと推論一般のどこが違うといえば，プランニングにより得られるプランは，後にアクチュエータをもった行為主体により必ず解釈され，行為主体が操作可能な対象世界に対し実行されることが前提となっているところではないだろうか．そして，行為主体と対象世界のわかりやすい例が，ロボットとそれを取り巻く物理世界(外界)だろう．こうして考えると，本来プランニングは，最終的には外界における目標の実現を目指す研究分野のはずである．しかし，従来のプランニング研究[1] [2]は，外界におけるプランの実行についてプランニング時に考慮することをやめ，外界の内部表現である記号の操作における探索問題の解決にそのほとんどの労力を費やしてきたと言えるだろう．このような方法論の根底には，外界でのプランの実行において生じる問題とそれらを無視した記号操作のプランニングとは別個に研究可能で，システム全体はその成果を線形的に構成することで実現されるであろうという前提がある．この前提は，プランニング中に外界は変化しない，プランの実行は失敗しない等の非現実的な仮定のもとでは十分に成り立つが，現実世界はプランニング中あるいは実行中にも状態が変化してしまう動的世界であり，そこではプランニングは不可避的に外界の変化の影響をうけ，それ自体を独立に取り扱うことはもはや困難である．

そこで，外界とのインタラクションを包括し，動的世界にも対応可能なプランニングであるリアクティブ・プランニング (reactive planning) の研究が近年活発になってきた[3]-[7]．本報告では，まずプランの成功確率を用いてプランニングと実行を交互に行なうインターリーブ・プランニング [山田91] の全体的なプランニング・アルゴリズムを提案する．そして，タイルワールドという現実世界を反映した領域でいくつかの興味深い実験を行ない，インタリーブ・プランニングが有効な問題のクラスを探る．

2．インタリーブ・プランニング

インターリーブ・プランニングとは，外界の変化に合わせてプランニングと実行を交互に行うリアクティブ・プランニングの一方式である。直観的にも明らかなように，この方式ではプランニングと実行の切り換えのタイミングをどのように決定するかが問題になるが，従来の研究では適切な解決がほとんどない．唯一とも言える解決案は，D. McDermot[McDermott 1978]によるもので，ワンステップ毎にプランニングと実行を繰り返すものである．しかし，本来は外界の変化が速ければ実行も速い時点で行ない，そうでなければゆっくりプランニングをした後で実行に移るというのが望ましい．このような外界の変化の度合いに適応したプランニングと実行の切り換えタイミング決定の方法をわれわれは既に提案しているので[山田91]，ここでは簡単に要点だけ説明していく．

2．1　切り換えタイミングの決定

インタリーブ・プランニングでは，プランニングと実行の切り換えタイミングをどのような情報をもとに決定するかが問題となる．本節では，この問題をいかに扱うかを簡単に説明していく．より詳細な処理は，文献[山田91][yamada92]を参照．

インタリーブのタイミング決定を我々は，次のような方針で行なう．

「プランニングから実行への切り換えは，(外界の激しい変化等により)現在立てているプランが実行不可能になる危険性が強まったときに起こる．」

この考え方に基づいて切り換えタイミングの決定，プランニング全体を構成していくが，まずプランの実行可能性をどのように表現するかが重要である．あるプランを動的世界で実行して，その結果目標が達成できるか否かは，必然的に不確定な要素を含んでいる．そこでこのプランの不確定な実行可能性をプランの成功確率という確率で表現する．なお本研究では，プランナーに対し以下の条件を想定している．

- オペレータOP：STRIPS-like[13]とする．条件リスト$CL_i=[C_{i1},...,C_{ip}]$，削除リスト$DL_i=[D_{i1},...,D_{iq}]$，追加リスト$AL_i=[A_{i1},...,A_{ir}]$で構成され，各リストの要素は正のリテラルのみである．また，条件リスト中のリテラルはすべて削除リストに含まれるとする．
- 目標状態オペレータGOP：目標状態を条件リストとするオペレータで，このオペレータが適用されると目標が達成されたとする．削除，追加リストは空である．
- 信念B：状態を正のリテラルで表現．
- プランP：適用されたオペレータの系列．$P=[OP_1,\cdots,OP_n]$と表現される．

2．2　プランの成功確率

まず，完全プラン(a complete plan)と部分プラン(a partial plan)という概念を考える．完全プランとは，目標状態オペレータを含むもので，プラン中のすべてのオペレータが実行されれば目標状態を達成できるプランを意味する．また，部分プランとは目標状態オペレータを含まないプランのことで，それが実行されたとしても目標状態が達成されるわけではない．

プラン実行の成功とは，プランを外界で実行した結果，目標が外界において達成されることであり，プランの成功確率とはプラン実行が成功する確率を意味する．よって，完全プランにおいては目標状態が成り立つ確率，つまり目標状態オペレータが適用できる確率をプランの成功確率とみなすことができる．しかし，部分プランの場合，目標状態が部分的，あるいはまったく成り立っておらず，プランの実行した状態の定義が難しい．もちろん，部分プラン中に予め設定された副目標が含まれる場合はよいが，部分プランがサブゴールを含むタイミングで実行が始まる保証はない．よって，以下のように，我々は(サブ)ゴールに依存しないようにプラン実行の成功を定義する．

「プラン実行の成功とは，プラン中のすべてのオペレータが実行可能である状態を意味する．」

この定義によると，完全プランに対しては目標状態オペレータが含まれているから，従来の目標の達成に等価である．しかし，部分プランに関しては，プラン中のすべてのオペレータが適用されることが部分プランの目標であるというヒューリスティックに基づいている．

このような定義のもとでプランの成功確率を計算するが，まずオペレータの適用確率を定義する．オペレータの適用確率は，プラン中のすべてのOP_i $(1 \leq i \leq n)$について割り当てられる．このプランのオペレーションに対応するアクション $ACT_1,\cdots,ACT_n$を外界で実行し，それぞれのアクションの実行直後に観測により得られた信念を$S'_2,\cdots,S'_{n+1}$とする．あるS'_i $(1 \leq i \leq n)$に対しOP_iを適用するという試行をおこなう．この試行の結果得られる事象は，適用可能か不可能かのいづれかである．そこで適用可能である事象を$Ap(OP_i)$，OP_iの適用確率を$Pr(Ap(OP_i))$とする．適用不可能な確率は，$1-P(Ap(OP_i))$ (≥ 0)である．

プランPlan中のオペレーション系列を$OP_1,\cdots,OP_n$とすると，プランPの成功確率SP(P)は以下のようになる．

$$SP(P) = \Pr(Ap(OP_1),\cdots,Ap(OP_n))$$
$$= \Pr(Ap(OP_1))\cdot\Pr(Ap(OP_2)\mid Ap(OP_1))\cdots\Pr(Ap(OP_n)\mid Ap(OP_1),\cdots,Ap(OP_{n-1})) \quad \cdots(1)$$

よって，以前のオペレータがすべて適用可能であるという条件の下における条件つき適用確率$\Pr(Ap(OP_1)), \Pr(Ap(OP_2)\mid Ap(OP_1)),\cdots,\Pr(Ap(OP_n)\mid Ap(OP_1),\cdots,Ap(OP_{n-1}))$を求めればよい．これらの条件つき確率の詳細な算出法は，既に報告している[山田91][Yamada92]ので，ここでは実験方法及び結果の理解に必要となる，成功確率算出の入力についてのみ以降で触れる．

2．3　成功確率計算の入力

(1) リテラルの追加確率A-Pr(OP, A)

プラン中のあるオペレータOPが外界で実行された結果，追加リスト中のリテラルAに対応する効果が外界において実現されている確率を追加確率A-Pr(OP,A)とする．また，外界においてリテラルに対応する状態が存在するとき，そのリテラルは外界に存在するという．また，追加確率A-Pr(OP,A)は，OPが適用可能か否かだけに依存しているとする．この追加確率は，オペレータの削除リスト中のリテラルには割り当てられない．これは，オペレータが条件リスト中に負のリテラルを持たないため，プランの成功確率の計算に削除が成功するか否かが無関係なためである．

(2) 状 態不変確率NC-Pr(L, T)

ある時点に外界で存在していて，なおかつオペレーションにより削除されないリテラル Lが，T秒後にも外界で成り立っている確率を状態不変確率NC-Pr(L,T)とする．リテラルが外界に存在することが，オペレータの実行によるものか否かは問わない．状態不変確率は，リテラルと時間の関数となる．

(3) 観測確率OB-Pr(L)

信念を構成するリテラルは，外界を観測(知覚も含む)することにより得られるが，その観測及び知覚はノイズや観測誤差による不確定性を常に含んでいる．そこで，観測されたリテラル Lがそのまさに観測の瞬間に外界において存在する確率を観測確率OB-Pr(L)とする．この確率により，観測の精度を表現できる．

(4) 実行及びプランニングの時間的コスト

本研究では，プランの実行やプランニング中でも外界が変化する動的世界を扱っており，またその変化を時間の関数としてとらえるので，オペレータの実行やプランニングにかかる時間的コスト，具体的には所要時間を設定しなければならない．このコストは，各オペレータの実行時間とプランニングの基本単位について割り当てられる．

以上の入力とプランから，そのプランの成功確率が計算される．簡潔にいうと，まずどのリテラルがどのオペレータにより追加されたのかという因果関係を表した因果ネットがプランから生成される．この因果ネットは，ベイジアン・ネットワーク[Pearl88]で形式的に表現可能であり，前述の各条件つき確率もベイジアン・ネットワークの遷移規則により算出可能である．また，一般にベイジアン・ネットワークは，ループを含んでいると更新の計算量が指数的に爆発するが，証明はないものの因果ネットワークではプランのステップ数の線形オーダで更新可能ではないかと考えている．より詳細な計算手続きは，文献[山田91][yamada92]を参照されたい．

3．インターリーブ・プランニング・アルゴリズム

3．1　プランニング

本節では，プランの成功確率を用いたインターリーブ・プランニング全体のアルゴリズムについて説明する．

インターリーブ・プランニングの場合，全体のアルゴリズムは必然的に，記号操作のプランニングと実行／観測というアクションの双方を組み込んだものとなる．前節で，実行するタイミングを決定するためのプランの成功確率の計算法を簡単に述べたが，全体のアルゴリズムではプランを探索する手続きが必要となる．

プランニングは，基本的には前向きのビーム・サーチで行う．前向き探索を用いるのは，動的世界でのプランニングでは，部分プランの実行を行う必要があり，生成されるプランは初期状態に適用可能なプランでなければならないためである．後向きに生成された部分プランは，プランの適用条件が観測されて初期状態で満たされている保証はない．また，ビーム・サーチを用いるのはプラン探索の際の，探索空間の組み合わせ的爆発を抑制するためであり，各レベルにおいて生成された各々のプランに対して算出された効用を用いて，一定数にプラン候補の枝刈りを行い，探索木の幅を一定に制限する．

また，観測はプランニングとは並列に行うとするが，観測情報はプランニング開始時点のみプランナに取り込まれるとする．

3．2　プラン実行のタイミング決定

ビーム・サーチにより絞り込まれたプランについて，次に各々の成功確率が計算される．これらのプランの成功確率を基にプラン実行のタイミング決定を行う．また，前述のように，プランニングと実行の切り換えタイミングの決定は，「プランの実行が危うくなったら実行に移る」ということである．したがって，プランニングを進めていき，各レベルのプラン成功確率の中に実行のしきい値を下回るプランが一つでも生成された場合に，プランニングから実行の切り換えは起こるとする．また，プランの成功確率は，そのプランを構成するオペレータの適用確率の積となるため，その値はプラン展開レベル毎に単調減少する．よって，動的世界では，[0,1]のしきい値を設定すれば，必ずプランの成功確率は有限のプランニングによりそのしきい値を下回ると考えてよい．

プランニングと実行の切り換え時に実行のしきい値を下回るプランが複数ある場合，次に実行するプランの選択を行う．実行のしきい値を下回ったプランの中で成功確率とプランの効用の積が最大であるプランが選択されて，実行される．この選択は，意思決定理論と同様である．以上の方法により実行されるプランは，プランニング中の実行候補として展開されているプランの中に失敗の可能性が，あるレベルまで高まったプランが生成された時に，失敗しそうになったプランの中からプラン成功確率が高く且つプランの効用も高いプランが選択されることになる．また，ビームサーチで行われる枝刈りにより，効用の低いプランが実行されることはない．しきい値を下回るプランがない場合は，次のレベルのプランの展開が行われ，実行のしきい値を下回るプランが生成されるか，もしくは完全プランが作成されるまでプランの展開が続けられる．

ここで実行のしきい値について考えてみると，実行のしきい値は０から１まで変化させることが可能である．実行のしきい値を１に設定すると，常に１ステップのプランニングで成功確率とプランの効用が最大のプランを選択し，実行するということを繰り返すことになる．逆に実行のしきい値を０に設定すると，成功確率が下がるのを無視して，とにかく完全プランを作成し，実行する．一般に動的世界では，完全プランの成功確率はかなり低く，その実行はほとんど成功しない．

4．実験

今まで説明したインターリーブ・プランニングを用いて動的世界において問題解決を行うエージェントについて実験を行う．この実験の目的は，動的世界においてはたしてインターリーブ・プランニングが有効な問題があるのか，より具体的にはワンステップでも完全プランでもない中間の部分プランを実行した方が有効な場合があるのかということを探ることである．また，実行のタイミングを決定するプラン成功確率のパラメータが環境の性質に対して，どのような傾向を示すかも調べる．

4．1　対象領域

エージェントが問題解決を行う領域は，タイル・ワールド[Pollack and Ringuette,1990]と呼ばれる仮想的な環境である．ただし，実際に実験を行ったのは，これを更に簡略化したタイルのない領域である[Kinny and Georgeff]．このタイル・ワールドは２次元の格子からなる領域で，エージェントは格子を移動し，ホールに到達するとそのホールに付けられたスコアを得るというものである．個々のホールは，一定の存在／消失期間を有しており，出現／消失時刻からの経過時間により消失／出現する．

出現の際の座標は，ある分布のもとに決定される．また問題を複雑にする要素として，固定された障害物を設定可能である．このタイ・ワールドは，単純ではあるが動的世界の一種であり，上記の実験の目的に沿ったものであると考える．

4．3　エージェント

インターリーブ・プランナー，そして生成されたプランの実行と環境の観測を行う主体を，ここではエージェントとよぶ．エージェントは，タイル・ワールドの中を上下左右に1格子移動するためのオペレータをもっている．エージェントはプランニングの際，各展開プランについて目標とするホールを決定し，そのホールに近づくようにプランニングを行う．この場合，目標とするホールに到達するまでに偶然別のホールを経由することもあるため，プランを構成するオペレータは，単に格子を移動するオペレータ，目標とするホールに到達するオペレータ，目標以外のホールに到達するオペレータの3種類がある．これらのオペレータを組み合わせることで格子を移動し，存在するホールを巡回するパスを作成する．尚，タイル・ワールドの場合の完全プランは初期状態で観測されたホールを全て巡回するプラントする．

ビーム・サーチで用いるプランの効用関数は，目標とするホールまでの距離とそのホールに到達するまでに作成プランが経由したホールのスコアから算出する次式を用いる．

$$
\begin{aligned}
Utility(Plan) &= \frac{Score(Goal-Hole)}{Manhattan\text{-}distance(Agent, Goal-Hole)+1} + Score(Plan) \quad (Agent \neq Goal-Hole) \\
&= Score(Plan) \qquad\qquad (Agent = Goal-Hole)
\end{aligned}
$$

エージェントは，観測をプランニング／実行と並列に行っているとし，すべてのホールの出現時刻，出現座標等の外界変化の情報は観測可能である．観測された情報は，プランニング開始時点にエージェントに取り込まれ，この情報を基にプランニング及びプランの成功確率の計算が行われる．実験を行ったタイル・ワールドの成功確率計算の入力は，以下のようになる．

○観測の成功確率：外界を観測したときのホールの出現時刻，出現座標等の精度．今回の実験では，観測の成功確率は1.0とした．

○状態不変確率：実験を行うタイル・ワールドの場合，外乱により変化するのはホールのみである．従って，ホール以外の状態不変確率は1.0となるが，ホールの場合は，ある時刻にホールが存在している確率を表す．ホールは一定の存在期間を有しているため，状態不変確率は出現時刻と経過時間，ホールの種類によって決まる．この状態不変確率は以下の式で算出される．エージェントは，ホールの存在時間を予め知っているとする．

$$
\begin{aligned}
NC\text{-}Pr(L,T) &= 1 - \frac{\Delta T}{2 \times Appearing\text{-}interval(L)} \qquad (\Delta T \geq Appearing\text{-}interval) \\
&= 0 \qquad\qquad (\Delta T < Appearing\text{-}interval)
\end{aligned}
$$

○追加確率：格子の移動もしくはホールに到達するオペレーションに失敗する可能性のある場合，追加確率を1.0より小さい値に設定するが，今回の実験では追加確率は1.0とした．

以上のように，本研究では時間経過と共に変化する外界において問題解決を行うエージェントを考えているが，外界の変化を時間の関数を用いて扱っている．従って，基準となる時間の単位が必要であるが，ここではプランを1レベル展開するのに要する時間を1単位時間としている．これから1つのオペレータの実行コスト等の他のコストは，この1単位時間の整数倍で表現する．また，観測のコストはプランニング／実行とは並列に行っているとしているので，観測の時間的コストは0とした．

エージェントの性能評価は，一定時間内に獲得したスコアの合計で評価すると，環境変化の激しさが異なる場合，直接比較することができない．よって，エージェントが得た得点の合計を，一定時間に出現するすべてのホールのスコアの合計（得点の上限）で割ることにより正規化して得られた得点率を用いて評価する．

5．　実験方法及び結果

タイル・ワールドにおいて実験を行ったが，以下に示す実験は外乱として初期状態に存在しているホールが消失するだけで，新たにホールは出現しないという場合である．ホールが出現／消失する実験も行ったが，その場合1ステップのプランニング／実行を繰り返す，実行のしきい値が1.0の場合が概ね得点率が高くなった．これはエージェントが観測情報をプランニング開始時点でしか取り込まないとしているために，1ステップのプランニング／実行を繰り返す場合は頻繁に観測を行うことができるためである．

実験１　ホール存在期間を変更

問題によりホールの存在期間を変化させることで環境変化が大，中，小の場合について実験を行った．各パラメータの設定は，以下の通りである．

[ホール]：初期状態で15個のホールが存在．ホールは消失するのみで新たにホールは出現しない．
　　　　初期状態での出現位置及び、出現した時刻はランダム

[ホールの存在期間 int]　問題１：250単位時間，問題２：350 単位時間，問題３：450 単位時間

[領域サイズ]：10×10

[実行コストecost]：5　，[ビーム・サーチ幅]：25

問題１は，環境変化が激しい場合で実行のしきい値が1.0の時が最高の得点率となっている．これは１ステップのプランで成功確率と効用の積が最大のプランの実行を繰り返す場合である．よって，環境変化が激しい場合に最も確実なプランを常に実行することになり得点率が高くなると考えられる．逆に環境変化が激しい場合に，実行のしきい値を下げることでよりステップ数の多いプランニングを行うことは，失敗する無駄なプランニング／実行を行うことが多くなり，かえって得点率を低下させることになる．従って，得点率が最低となるのは，最も多いステップ数のプランニングを行う実行のしきい値が0.1の場合となっている．

それに対して環境変化がある程度緩やかな問題２，３の場合，得点率が最高となるのは，ある程度のステップ数のプランニングを行った場合となる．これはFig.2に示すようなホールの位置関係及び出現時刻の関係にある場合が典型的で，このときいずれのホールの効用も等しい．よって，実行のしきい値が1.0の時は必ず成功確率が高い，すなわち消えにくいホールから先に移動する．従って，２つ目のホールに移動しようとしてもホールに到達する前に消失してしまっていて得点を得ることができない場合が生じる．それに対して，しきい値を下げてプランニングを行うと，先ず消えそうなホールに移動し，次に消えにくいホールに移動するプランを作成することが可能となるため得点率が上昇すると考えられる．しかし，環境変化が中程度の場合は，実行のしきい値を0.1近くまで下げて非常に長いステップ数のプランニングを行うと，１ステップのプランニング／実行の場合に比べて得点率は低くなる．しかし，問題３のように環境変化が小さい場合は，長いステップ数のプランニングによる効果が大きくなり，得点率が最小となるのは，しきい値を高く設定した場合となっている．

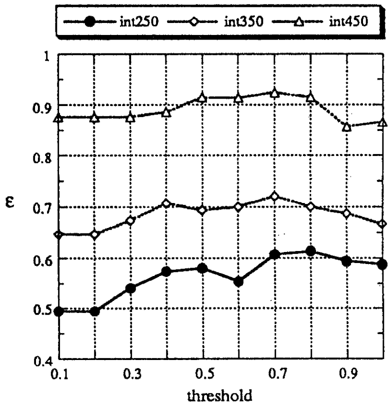

Fig.1 ホール存在期間の変更

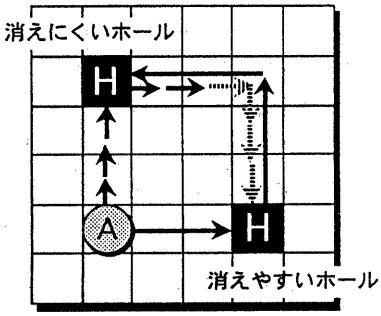

Fig.2 1stepのプランニング／実行が失敗する典型的な例

実験２　ホールの出現位置に偏りがある場合

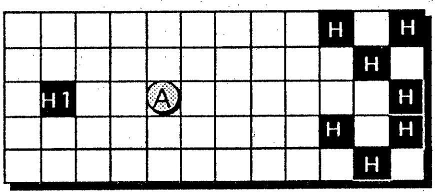

Fig.3 ホールの出現位置の偏り

[ホールの存在期間 int]：80 単位時間，Hole1は時刻-10に出現，他のホールは時刻-10以前に出現

[領域のサイズ]：12×5

[実行コストecost]：3 単位時間，5 単位時間

[ビーム・サーチ幅]：25

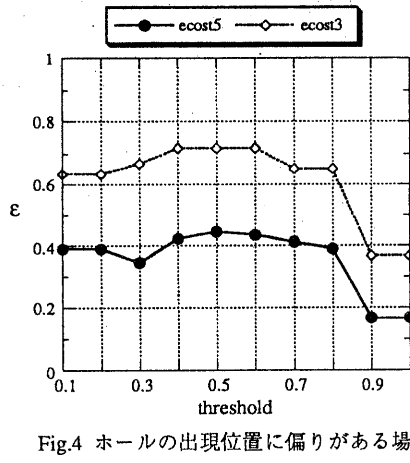

Fig.4 ホールの出現位置に偏りがある場合

この実験は，Fig. 3 に示すようにホールの出現位置に偏りがある場合の実験である．左に１つ孤立しているホールは，他のホールに比べて出現時刻が遅く成功確率が高いホールである．この問題では，成功確率は相対的に低いがホールの密集した右に先ず移動するか，成功確率は高いが１つしかホールのない左に移動するかが得点率を大きく左右する．実験の結果，実行のしきい値が高い場合，必ず左の成功確率の高いホールへ移動するプランの実行が行われ，得点率は低くなっている．それに対して，実行のしきい値を低くすると作成するプランのステップ数が増加して，数ステップのプランでは右に移動するプランの効用の方が高くなり，先ず右に移動するプランの実行が行われる．また，実行のしきい値が0.8以下の場合の得点率にあまり変化が見られないのは，右の領域に非常にホールが密集していることで実行のしきい値を変化させても大きな違いが生じないためと考えられる．

実験３　高い効用を持つホールが１つ存在する場合

[ホール]：15個，１つのホールだけが他に比べてスコアが100倍で消失しない
[領域サイズ]：10×10，　[実行コストecost]：5
[ビーム・サーチ幅]：25

この実験では，消失しないがゆえに成功確率及びユーティリティの高いホールが１つ存在することにより，最低得点率0.8772が保証されている．実行のしきい値が1.0の場合は，Fig.6のようにユーティリティの高いホールが存在するために，先ずこのホールに向かうプランを実行することになる．しかし，ユーティリティは高いが消失しないホールであるから，このホールに移動している間に他のホールが消失してしまう可能性がある．従って，しきい値が1.0の場合は得点率は高くない．対照的に，実行のしきい値を下げると他の消えそうなホールを考慮しながらプランニングを行うため，得点率は高くなる．

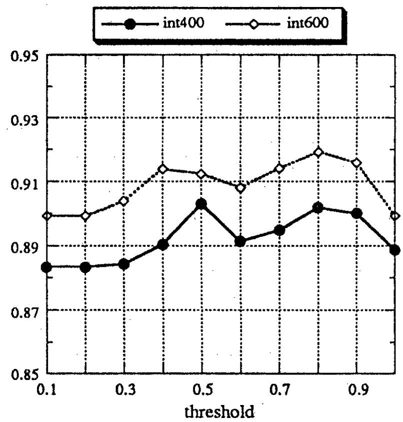

Fig.5 高い効用を持つホールが存在する場合

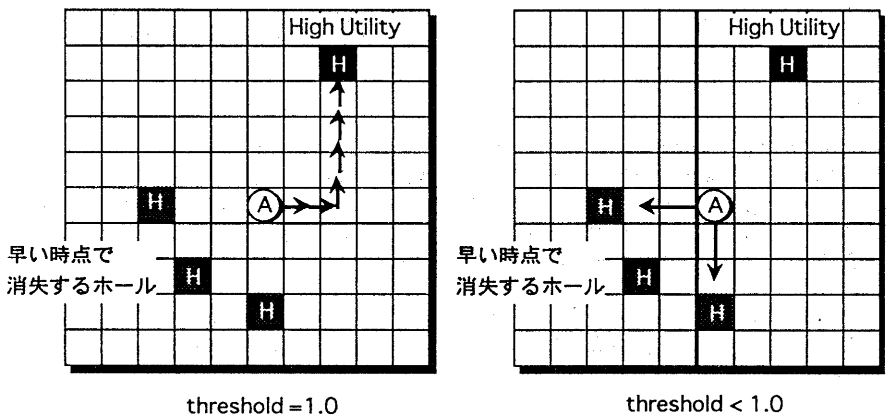

Fig.6 効用の高いホールが１つ存在する場合

実験４　障害物がある場合

[障害物の配置]：固定　Fig.7
[ホール存在期間 int]：問題１：200単位時間，問題２：300単位時間
[領域のサイズ]：10×10，　[実行コスト]：5，　[ビーム・サーチ幅]：25

障害物のある環境で，ホールの存在期間の異なる２種類の問題について実験を行った．問題１はホールの存在時間が200単位時間と環境変化が激しい場合で，問題２は存在期間が300と比較的環境変化が緩やかな場合である．

プランニングの際に使用するプランの効用関数が，目標ホールまでのManhattan距離を用いているため，領域中に障害物が存在すると目標とするホールとエージェントの位置関係によっては，効用関

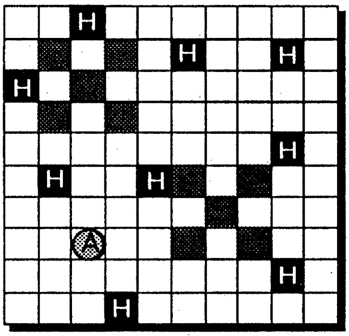

Fig.7 障害物のある環境

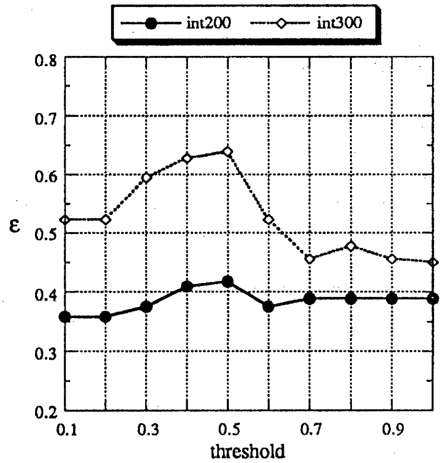

Fig.8 障害物がある場合

数の極小値に落ち込んでしまうことがある[石田：知性をもったMTS]．この障害物による凹部の大きさに対して十分なステップ数のプランニングを行う場合は，効用関数の極小値に落ち込むことはないが，少ないステップ数のプランニング／実行を繰り返す場合は効用関数の極小値に陥ることが多くなり，結果として得点率が低くなる．

問題１のように環境変化が激しい場合は，実行のしきい値を低く設定し，数ステップ・プランニングをすることで障害物の窪みに落ち込むことはなくなるが，環境変化が激しいために，得点率は比較的低い値となっている．実行のしきい値が高く設定されている場合は，極小値に落ち込むことにより得点率は低いが，環境変化が激しい場合は少ないステップ数のプランニングが有効である．このことは，実験１の結果からも分かる．結果として，環境変化が激しい場合は実行のしきい値の変化による得点率の大きな変化はないが，実験１の障害物のない場合と比較した場合，得点率のピークは実行のしきい値が低い方に移動している．

問題２のように環境変化が緩やかな場合は，実行のしきい値を下げて，数ステップのプランニングをすることで効用関数の極小値に落ち込むことがなくなり，かつ環境変化も緩やかであるため得点率は最高値をもつようになる．また，しきい値を変化させた場合の得点率の変化は大きくなる．

６．まとめ

成功確率を用いたプラン実行のタイミング決定を行うインターリーブ・プランニングの全体的なアルゴリズムを提案し，またその有効性を検討するために幾つかの実験を行った．そして，環境変化及び環境の激しい場合は少ないステップ数のプランの実行が有効であり，逆に環境変化が緩やかな場合は，ある程度のステップ数のプランの作成，実行が有効であるという直観にあう結果を得た．特にそのような結果が著しいのは領域中にホールの位置や数，プランの成功確率等の偏りがある場合及び障害物のある場合であった．

今後は，体系的な実験を行い環境変化及び環境の特徴に応じてプラン成功確率を決定するパラメータ及び実行のしきい値が，どのように変化するのかを調べ，しきい値決定の方法論を確立することを目指す予定である．

＜参考文献＞

[石田1990]石田亨：知識表現と動的世界ー最近のプランニングの研究からー，人工知能学会誌，Vol.5,No.2,pp.146-153(1990)

[Pollack and Ringuette1990]M.E.Pollack,M.Ringuette:"The Tileworld:Experimentally Evaluating Agent Architectures".AAAI-90,pp.183-189(1990)

[McDermott 1978]D.McDermott:Planning and Action,Congnitive Science 2,pp.71-110(1978)

[Ishida1992]Toru:Ishida:Moving Target Search with Intelligence.AAAI-92(1992)

[山田91]山田誠二：インターリーブによるリアクティブ・プランニング，情報処理学会第79回人工知能研究会，91-AI-79-8 (1991)

[Yamada92] Yamada, S.: Reactive Planning with Uncertainty of a Plan, pp.201-208, In Proceedings of The Third Annual Conference on AI, Simulation and Planning in High Autonomy Systems (1992)

[Pearl88] Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference (1988)