

単純タイルワールドにおける *SIP* の実験的評価

Experiments for *SIP* in the Simplified Tile World

山田 誠二*¹ 磯田 佳徳*² 豊田 順一*³
 Seiji Yamada Yoshinori Isoda Jun'ichi Toyoda

- * 1 東京工業大学大学院総合理工学研究科知能システム科学専攻
 Interdisciplinary Graduate School of Science and Engineering Tokyo Institute of Technology, Yokohama 226, Japan.
- * 2 大阪大学産業科学研究所
 I. S. I. R., Osaka University, Ibaraki 567, Japan.
- * 3 日本電信電話(株)NTT ヒューマンインタフェース研究所
 NTT Human Interface Laboratories, Tokyo 180, Japan.

1995年8月23日 受理

Keywords: reactive planning, dynamic world, deliberation.

Summary

We have proposed the *SIP* planning method which interleaves planning with executions by using the success probabilities of plans. With *SIP*, an agent is able to control deliberation depending on the change of an environment. The formal framework has been developed using a plan Bayesian network, however we need to investigate the utility of *SIP* in a testbed for a dynamic environment. In this paper, we report results for various experiments made in a simplified Tile World. First we characterize a simplified Tile World with a few essential parameters: dynamics, cost for actions including observations and uncertainty. Various experiments are made, and we find out interesting properties: a single optimal execution threshold between reactivity and deliberation, robustness of the optimal threshold against the change of dynamics and observation costs. We consider that these properties significantly contribute to design of an agent in a dynamic environment.

1. はじめに

動的環境において熟考を制御する方法として、筆者の一人は、プラン実行の成功確率に基づいてプランニングと行為の実行をインタリーブするプランニング *SIP* を提案した[山田 96]。 *SIP* では、局所最適なプランの実行の成功確率がしきい値を下回ったときに、プランの実行が始まる。成功確率の減少の傾きは、環境変化の速さに依存するため、変化が緩やかなときは深いプランニングが、環境の変化が激しいときは浅いプランニングが行われ、熟考が環境の変化に応じて制御される。本論文では、動的環境の一般的な実験環境である単純タイルワールドにおいて *SIP* の実験を行い、その有効性を調べる。まず、簡単な環境モデリングについて述べ、次に従来の単純タイルワールドを少数の重要なパラメータで特徴づけ、そして、 *SIP* と環境の

パラメータをさまざまに変えて実験を行う。なお、環境に応じて熟考を制御できるシステムを用いて、さらに熟考のコストも考えた実験はこれまでに例を見ず、本研究は新規性を持つ。

2. *SIP* の概要

まず、実験で用いる *SIP* について簡単に説明する(なお、詳細は、[山田 96]を参照されたい)。 *SIP* とは、プラン実行の成功確率をもとに、プランニングとプラン実行を切り換えるインタリーブプランニングのアルゴリズムである。 *SIP* により行動決定するエージェント(*SIP* エージェントと呼ぶ)の環境には、価値が割り当てられた、出現消滅を非同期に繰り返す複数の目標が存在し、エージェントは、目標が消滅するまでにそれを処理してその価値を獲得する。また、エージェントの目的は、単位時間当りに獲得する価値をで

きる限り高くすることである。オペレータとして、一つの目標の処理に対応する STRIPS like な目標オペレータを用い、具体化された目標オペレータの系列をプランとする。ここで、*SIP* にとっての熟考とは、目標を処理する最適な順序を考えることであり、即応とは、処理順序を検討せずに即座に実行を開始することである。

SIP エージェントは、観測、実行、環境モデリング、*SIP* プランナの各モジュールから構成される。実行モジュールはプランナが生成したプランを実行し、観測モジュールはほかのモジュールとは並列に外界を常時観測しており、センサ情報を一定時間ごとに内部表現で記述された状態に変換する。環境モデリングモジュールでは、観測モジュールから得られる環境の情報に基づいて、環境のモデルを獲得するが、*SIP* では後述するように持続確率の推定のみを行っている。獲得された持続確率は、定期的に *SIP* プランナに渡される。*SIP* プランナは、*SIP* によるプランニングを行う。

SIP は、幅 w の前向きビーム探索を行う。1ステップの展開ごとに、展開されたプランそれぞれについて成功確率と期待値が算出され、期待値の上位 w 個のプランが選択されて展開が繰り返される。各レベルの展開において、期待値最大のプランの成功確率が実行のしきい値 $\tau \in [0, 1]$ 以下になるか、またはそのプランがすべての目標を処理できる完全プランである場合、プランニングを中止して、その期待値最大のプランを実行する。プランの実行後、再び観測モジュールから状態を受け取り、初期状態が更新され、プランニングが繰り返される。

3. 実験環境

3.1 単純タイルワールド

動的環境のテストベッドとして一般に用いられている単純タイルワールド [Kinny 91] とは、エージェント、障害物、そして点数の割り当てられた穴のある図1のような2次元格子である。エージェントは、一つのセルを占め、一度に上下左右に一つのセルだけ移動

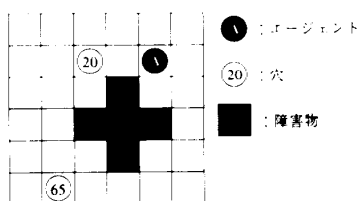


図1 単純タイルワールド

できる。穴の位置にエージェントが着くとその点数を得ることができ、また穴はある時間経つと消滅したり、新たに現れたりする。エージェントの目的は、単位時間当りにできるだけ高得点を得ることである。我々は、この単純タイルワールドにおいて本質的なパラメータを設定し、さらに観測・行為の不確実性を加えた拡張を行って、そのうえで *SIP* の実験を行う。

3.2 パラメータ設定

まず、実験環境の単位時間を設定する。1単位時間の実時間の変動することは好ましくないので、エージェント内の手続きで繰り返される定数オーダの処理を単位時間にする。別稿 [山田 96] で示したように、プラン1ステップ分の成功確率の計算は定数オーダなので、*SIP* のプランニング中の枝一つの展開コストも定数オーダである。よって、その展開コストをエージェントの1単位時間とする。

単純タイルワールドは、動的環境のための数少ない実験環境の一つであるが、十分に洗練されていない [Hanks 93]。よって、本節ではできるだけ少数の本質的なパラメータで、単純タイルワールドを特徴づける。まず、単一エージェントの動的環境にとって、重要な要素として以下のもの考える。

- (1) 環境変化の速さ
- (2) 観測を含む行為の不確実性
- (3) プランニングと観測のコスト

(1)は、エージェントとは独立な環境の変化である。

(1)~(3)を含むように単純化したパラメータを以下に示す。

〈パラメータ〉

- ・環境のダイナミクス d : 環境時間とエージェント時間の進む速度の比 [Kinny 91]。
- ・行為の不確実性 u : 観測を含むすべての行為が失敗する確率を一律に u とする。ここでは、エージェントが、穴、障害物の位置を正確に観測できる確率は $1-u$ であり、確率 u でランダムな上下左右1セルの近傍の位置に誤って観測してしまうとする。また、エージェントは、確率 u で移動できずに止まってしまう。
- ・観測コスト c : 1回の観測に要する時間であり、すべての観測について同一とする。

4. 単純タイルワールドでの実験

4.1 実験の目的

本実験の目的は、*SIP* について以下のことを調べ

ることである。

- (A) 実行のしきい値とのパフォーマンスの関係
- (B) 最適なしきい値の頑健性
- (C) 深さ固定の部分プランニングとの比較

(A)では、実行のしきい値を変えてパフォーマンスを調べたときに、はたして即応と熟考の間に最適なしきい値が存在するのかを調べる。(B)では、ダイナミクスと観測コストの変化により、最適なしきい値が変動するのかを調べる。また、*SIP*の特徴として、成功確率計算をオンラインで行っていることがある。このため、*SIP*のプランニングの深さ(=ステップ数)は概ねしきい値に依存するものの、個々のプランニングの深さは観測された目標に応じて変化する。例えば、*SIP*では平均としては深さ5のプランニングを行うしきい値の場合でも、高得点で短命の穴があると、1ステップで実行を開始することがある。一方、単純な熟考制御法として、プランニングの深さを固定することが考えられるが、その場合 *SIP*のように個々の目標に応じた調整はない。この差を調べるのが、(C)である。

さらに(C)では、ダイナミクス自体が変化する環境での実験が行われる。深さ固定プランニングは、ダイナミクスの変化に適応できるメカニズムを持たないのに対し、*SIP*は環境モデリングにより適応可能であることを示す。

4・2 設 定

前節で設定したパラメータを変化させて実験を行うが、まず、[Kinny 91]を参考に固定されたパラメータを説明する。以下の設定は、実験全体を通して固定される。

(1) 環境の設定：(時間は、環境時間)

- ・単純タイルワールドのサイズ：20×20で障害物なし。
- ・穴の点数の範囲：[60, 100]。
- ・穴の寿命の範囲：[1 200, 5 200]。
- ・穴の生成周期の範囲：[100, 300]。

生成周期ごとに新しい穴が一つ生成され、寿命が経つと消滅する。穴の点数、生成周期、寿命は、各範囲からランダムに設定される。この設定における、時間

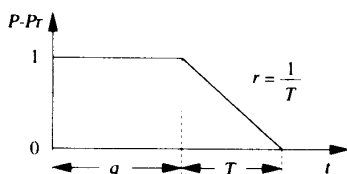


図2 生存関数

($t=0$ で穴の発生)と穴の存在する確率の関係は、

$$P \Pr(L, t) = \begin{cases} 1 & (t < q) \\ rt - rq + 1 & (q \leq t < q + 1/r) \\ 0 & (q + 1/r \leq t) \end{cases}$$

であり、そのグラフは図2のようになる。図中で、 q 、 $q+T$ は、それぞれ穴の寿命の最小値、最大値である。寿命の平均と生成周期の平均の比(ここでは、16)が、エージェントがいない場合の任意の時点で現れている穴の数の平均となる[Kinny 91]。Kinnyらの研究[Kinny 91]と同様に、ここでは移動のパスは既与とする。このような前提では、障害物は意味がないので設定しない。また、ビーム幅を変えた実験も行ったが、変化がなかったため、ここでは後述のように固定している。

(2) エージェントの設定

- ・目標オペレータ：別稿[山田 96]参照。図3に、prologで記述された目標オペレータを示す。
- ・目標オペレータの実行：目標までのマンハッタン距離による山登り法で移動。
- ・ビーム幅：4で固定。
- ・一つのセルの移動コスト：2単位時間で固定。
- ・入力確率：穴とエージェントの位置を観測するときの観測確率は、 $1-u$ であり、マンハッタン距離 b の移動が成功する効果確率を $(1-u)^b$ とする。また、持続確率は、前述の $P\text{-}\Pr(L, t)$ の式が与えられる。

以上が固定パラメータである。なお、上記の入力確率はエージェントが自ら環境を観測し、統計的に推測することも可能であり、この環境モデリングについては後述する。

(3) エージェント時間の進行：エージェントが、1回の観測、 i 個のノードの展開による目標系列のプランニング、 j 個のセルの移動という処理を繰り返すごとに、エージェント時間が $c+i+2j$ だけ進む。

(4) パフォーマンス評価：

$$\text{得点率 } \epsilon = \frac{\text{エージェントの得点}}{\text{最大得点}}$$

最大得点は、出現した穴の延べ点数である。

(5) 実験方法：次節のようにパラメータを変えて実験を行った。得点率を1000エージェント時間ごとに調べ、その最近の10個の値の標準偏差が0.01

```
ope(1, goal_operator(agent1(Ax, Ay), hole(Name, Value, Hx, Hy)),
[cell(agent1, Ax, Ay), cell(hole(Name, Value), Hx, Hy),
score(Score), NewScore is Score + Value],
[cell(agent1, _, _), cell(hole(Name, Value), Hx, Hy), score(_)],
[s(cell(agent1, Hx, Hy)), score(NewScore)]).
```

図3 目標オペレータ

以内に収束すれば、実験を終了した。また、同じパラメータ設定について、エージェントの初期配置を変えた実験を5回行い平均をとった。その結果、大体数万台エージェント時間で収束した。

以降では、実行のしきい値が0.1以下と0.9以上を、それぞれ便宜的に**熟考**と**即応**と呼ぶ。また、パラメータのデフォルトの設定を、 τ が0.1刻みで0~1、 d が1刻みで1~8、 $u=0.01$ とする。

4.3 実験A：ダイナミクスと観測コストの変化

目的(A), (B)を調べるため、デフォルト設定と観測コスト $c=5, 50, 100, 200$ からなる1760問で実験を行った。それぞれの結果が図4~図7である。

まず、全体に得点率はダイナミクスの増加とともに単調に低下する。図4では、観測コストが低いため、頻繁に観測を行う即応でコストがかからず、熟考の弊害だけが出て、ダイナミクスの増大に伴って熟考の得点率が急激に低下している。次に図5では、ダイナミ

クスが小さいとき ($d=1, 2$) は得点率は一様に高く、しきい値を変えても差がない。しかし、 $d=3$ になると即応と熟考の両端の得点率が下がりはじめ、 $d=4, 5$ のときには、得点率の最大値と最小値の間に大きな差が生じ、特に $d=4$ でその差は、0.5にもなっている。以上のことは、即応と熟考の間に最適なしきい値がある問題の存在を示している。さらに、そのピークが一つという単峰性も概ね満たしている。この単峰性は、得点率を評価関数としてエージェント自身がしきい値のチューニングを行う場合、単純な山登り法で最適値を探索できるという意味で重要である。さらに、図6、図7でも、同様の特長が見られる。

また、以上の実験では、観測コストが、熟考と即応の間における最適なしきい値の存在を決定している。図4のように観測コストが低いときはダイナミクスの増加とともに、熟考の得点率が下がり、即応が最適となる。しかし、図5~図7のように観測時間がある程度大きくなると、ダイナミクスの増加とともに即応と

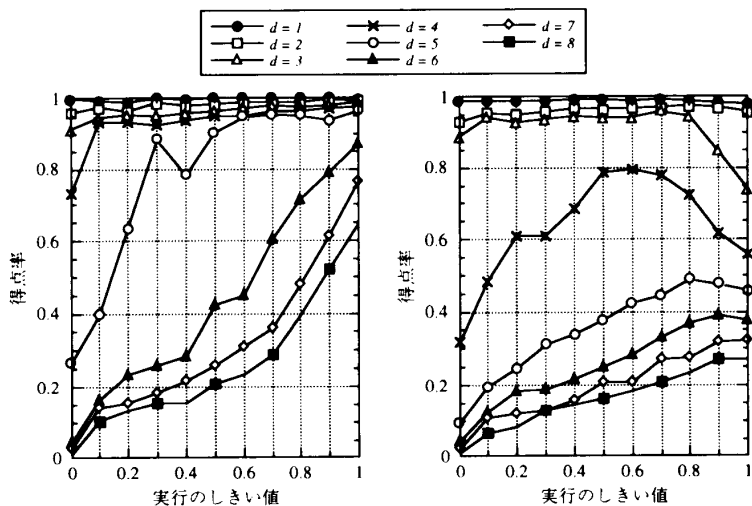


図4 実験1の結果 ($c=5$)

図5 実験1の結果 ($c=50$)

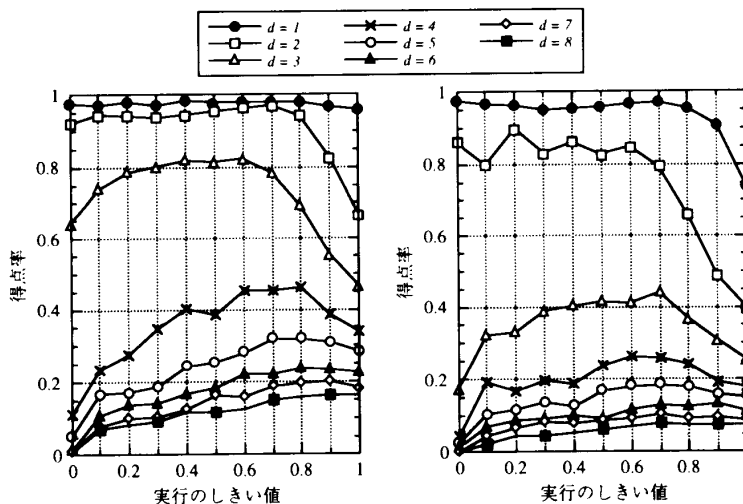


図6 実験1の結果 ($c=100$)

図7 実験1の結果 ($c=200$)

熟考の両方の得点率が下がり始め、その間に最適なしきい値が現れる。即応では、頻繁に観測が行われるため、観測コストが大きいと即応の得点率の低下が起こる。逆に、熟考では、長いプランが生成実行されるため、観測のインタバルが即応よりも長くなるが、環境の変化を取り込む周期が長いいため、変化に対応できず実行不可能なプランを多く生成してしまう。このような要因で、熟考と即応の間にピークが生じる。また、ダイナミクスを十分大きくすると、環境の変化によるプランニングの弊害が十分大きくなることによりトレードオフがなくなり、ピークは小さくなる。

図8に観測コスト $c=200$, $d=1, 3, 5, 7$ のときの実行時のプラン長と標準偏差の平均を示す。 $d=1$ のときは、1~2ステップのプランが実行されている。これは、穴がなくなり、次に出てくるのを待って、出たらすぐにとるというような定常状態になっており、得点率は高い。 $d=3, 5, 7$ のときは、しきい値の低下に従って、実行プランの長さが単調に長くなり、また同じしきい値でも、ダイナミクスが大きいことは実行プラン長が短くなることから、熟考が制御されているのがわかる。また、プラン長にばらつきがあるのは、**SIP** が局面ごとに成功確率をオンラインで計算しているためである。

さらに重要な点は、観測コストがある程度大きい場合(特に、 $c=100, 200$ のとき)、ダイナミクスと観測コストの変化に対し、最適なしきい値がそれほど変化しないことである。図9に図4~図7の各観測コストでの得点率のダイナミクスでの平均を示す(ただし、 $d=1, 7$ は変化があまりないため除いてある)。 $c=5$ より大きな各観測コストでは、熟考と即応の間で単峰性のピークを持ち、かつ図4~図7で同一観測コストにおける任意の二つのダイナミクス間の順位相関係数

(rank correlation coefficient) $\in [-1, 1]$ が正の値 $0.2 \sim 0.5$ であることより、最適なしきい値がダイナミクスの変化に頑健であるといえる。また、図9において最適なしきい値 ($0.6 \sim 0.8$) が、観測コストが変化してもそれほど変化しないことより、観測コストに対しても頑健であることがわかる。

これらの頑健性は、**SIP** が環境の変化と観測コストを考慮して成功確率を計算することに起因する。つまり、ダイナミクスの大きな環境では、成功確率の傾きも大きくなるため、しきい値が固定でもプラン長の調整が可能であること、さらに観測コストを含む時間を考慮して成功確率の計算が行われることで、最適なしきい値を固定することが可能となる。

以上の実験から得られる、単純タイルワールドにおける **SIP** の特性を以下に示す。

- ・熟考と即応の間に最適なしきい値が存在する問題がある。
- ・最適なしきい値は、ダイナミクス、観測コストの変化に対し頑健である。

実験Aと同じパラメータで、しきい値を変えるのではなく、深さ固定プランニングにして、その固定深さ(プラン長)を1~10で変化させて実験を行った。この深さ固定プランニングは、プラン長が与えられた定数になったときに停止すること以外は、**SIP** と同じである。

$c=100$ の結果を、図10に示す。図6と比較すると、最高得点率はほとんど変わらない。このことは、他の c の値でも同じであった。**SIP** では、しきい値が固定でもプラン長は局面に応じて変化するが、残念ながらその効果が出ていない。

しかし、**SIP** との違いについて、直観的に考えられることは、制御パラメータの最適値の頑健性である。ここで、**SIP** と深さ固定プランニングの制御パラ

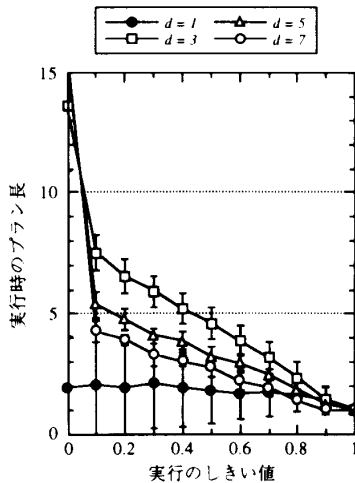


図8 実行時のプラン長 ($c=200$)

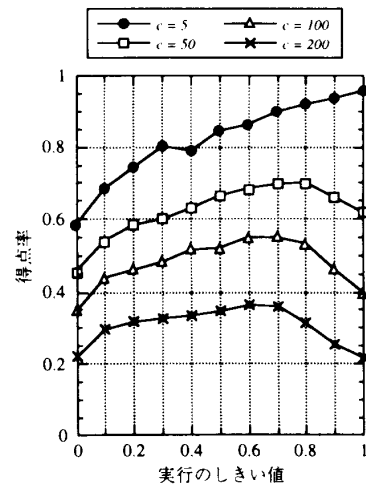


図9 ダイナミクスでの平均

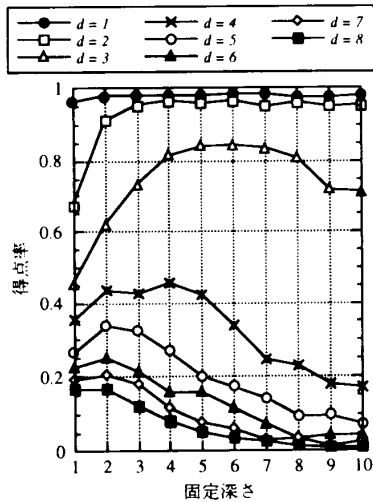


図 10 固定深さプランニング ($c=100$)

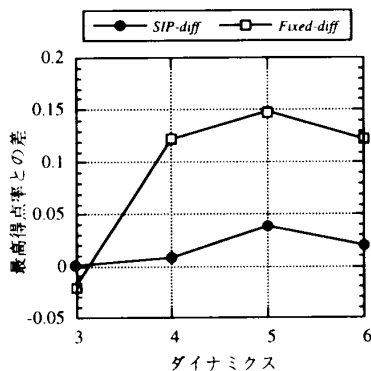


図 11 最適しきい値の頑健性

メータは、それぞれ実行のしきい値とプラン長である。先に見たように、*SIP*の最適なしきい値は、ダイナミクスに頑健であるのに対し、固定長プランニングは、環境の変化の速いときは短いプランニング、遅いときは長いプランニングというように、最適なプラン長が大きく変化することが予想される。そのような結果が、図 10 の $d=3$ と 5 の場合であり、最適なプラン長が 6 から 2 へ変化している。

また、*SIP* と深さ固定プランニングの制御パラメータを $d=3$ ($c=100$) における最適値(図 6 より $\tau=0.6$ と図 10 よりプラン長=6)に固定して、ダイナミクスを 3, 4, 5 と変化させた場合の得点率の変化のグラフを、図 11 に示す。縦軸は、図 10 における $d=2, 3, 4, 5$ での最高得点率(これらは、図 6 のものとはほぼ等しい)と、図 6 と図 10 の対応する得点率との差である。ダイナミクスの変化に伴い、*SIP* よりも深さ固定プランニングのほうが差が大きくなっているのがわかり、このことは、ダイナミクスの変化に対し、深さ固定プランニングの最適プラン長よりも *SIP* の最適しきい値が頑健であることを示している。よって、ダイナミクスの変化に追従するために、*SIP* はしきい値を再調

整する必要がないのに対し、深さ固定プランニングでは、実際のパフォーマンスを評価関数にしたプラン長の再調整が必要となり、それには多くのコストがかかる。ただし、*SIP* も維持確率の更新を行う必要があるが、次節で示す例のように比較的簡単な計算ですみ、さらにプランニングなどとは並列に処理可能なため、そのコストは実際のパフォーマンスを評価関数として制御パラメータを調整するよりは十分に小さい。

最後に、環境のダイナミクスそのものが変化した場合に対し、*SIP* エージェントが自身で適応する実験を行う。そのためには、これまでの実験ではその真値が与えられていた環境のダイナミクスを、*SIP* エージェントが環境を観測して得られた情報から随時推定する必要がある。つまり、[山田 96]における環境モデリングモジュールを実現しなければならない。次節でその環境モデリングについて説明し、続いて実験を行う。

4.4 環境のモデリング

SIP エージェントの行う環境モデリングとは、持続確率[山田 96]の推定であり、非常に単純なものである。持続確率とは、目標が経過時間に対し存在している確率分布であり、厳密には統計的推定が適用できるが、*SIP* エージェントでは、下式[Dean 89]で図 2 の q と r を計算することにより、持続確率の推定を行った。入力は、最近の n 個の寿命のサンプル集合 L である。新しいサンプルが得られるごとに、推定値が更新されるため、ダイナミクスの変化に適応可能である。また、観測確率および効果確率も、観測の真値が与えられれば、推定可能である。

$$q(L) = L \text{ の最小値}$$

$$r(L) = \frac{-0.5}{L \text{ の平均値} - q(L)}$$

この手法は、線形に減衰する持続確率をモデルにしており、真値が線形でない場合は誤差が大きくなるが、収束が速いことから採用した。目標が出現してから消滅するまでの寿命を観測モジュールがサンプリングして、環境モデリングモジュールに渡すが、当然ながらこの寿命は、目標が消滅する前に獲得されてしまうと得られない。よって、非常に高い得点率の状況では、推定不能である。また、サンプル数をいくらにするかは、精度と変化への即応性のトレードオフがある。

4.5 実験 B: ダイナミクスの変化する環境

深さ固定プランニングとの違いを明確するために、ダ

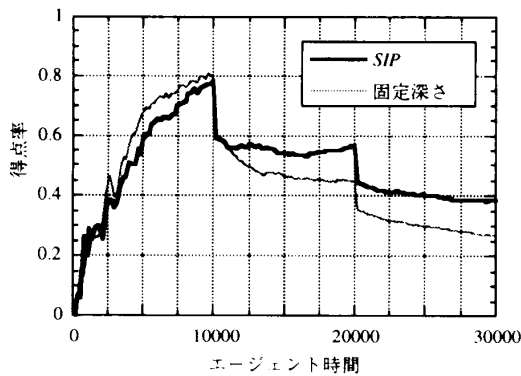


図 12 ダイナミクスの変化 ($c=100$)

イナミクスの変化する環境で実験を行った。観測コスト c は、二つのプランニングの違いが顕著である 100 に設定し、ダイナミクス d をエージェント時間 t が、0~10 000 では $d=3$ 、10 000~20 000 では $d=4$ 、20 000~30 000 では $d=5$ と変化させた。また、持続確率推定のサンプル数は 20 とし、**SIP** と深さ固定プランニングのそれぞれの制御パラメータを、図 11 の実験同様に、 $d=3$ 、 $c=100$ における最適値に固定した。その他の設定は、実験 A と同じである。

実験結果を図 12 に示す。 $t=10 000$ までのダイナミクス一定の期間では、深さ固定プランニングのほうが若干良いパフォーマンスを示しているにもかかわらず、 $t=10 000$ 以降でダイナミクスが変化した期間では、深さ固定プランニングはその変化に追従できず、**SIP** よりも得点率が下がっている。深さ固定プランニングでは、ダイナミクス変化の直後で急に得点率が下がり、さらに徐々に下がって収束値に近づいていくようすが見て取れる。それに対して、**SIP** は、ダイナミクス変化直後は落ち込むが、それ以後は迅速な(サンプルが 20 でよい)持続確率の推定により、変化後のダイナミクスに適応して、パフォーマンスの維持を図る。この傾向が、特に $t=10 000\sim 20 000$ で顕著である。よって、**SIP** が最も有効であるクラスの一つは、ダイナミクスの変化する環境であると考えられる。ちなみに、いずれのプランニングの値も、時間が十分でないため、図 6、図 10 の収束値とは一致していない。

なお、すべての実験で用いたエージェントおよび単純タイルワールドは、*SICSTUS Prolog v.2.1#7*(コンパイラ)と X ウィンドウインタフェース *XWIP* を使って、WS 上で実装された。

5. 考 察

一般に熟考と即応のトレードオフは、熟考による解(ここでは、プラン)の質の向上と熟考による時間の消

耗とのトレードオフである。しかし、図 4 から明らかのように、残念ながら単純タイルワールドにおいて、**SIP** の熟考、つまり目標の処理順序の最適化は、ほとんど質的向上をなしておらず、むしろそのトレードオフは、観測頻度と即応性の間にある[Kinny 92]。よって、単純タイルワールドでは、少なくとも **SIP** の熟考のメリットはあまりない。

また、単純タイルワールドでは、各目標(=穴)間の因果関係が単純すぎる。本実験では、図 3 からわかるように、エージェントがプラン上で直前の穴の位置にいることを条件とするという、非常に単純な因果関係のみを設定している。そのため、これから計算される成功確率は状況に依存せずに、プラン長の線形関数になる傾向が強い。このため、実験 A の結果のように、深さ固定プランニングとの差がほとんどないという結果となる。この観点から、単純タイルワールドにより複雑な因果関係(例えば、目標間の半順序関係)を導入しなければ、**SIP** における熟考のメリットは少ないと思われる。

次に、関連研究との比較を行う。タイルワールド上の熟考性と即応性に関する実験として、すでに Pollack らの研究[Pollack 90]と Kinny らの実験[Kinny 91]がある。Pollack らの研究は、エージェントアーキテクチャ IBMS の評価実験をタイルワールド上で行った。IBMS における熟考とは、次に処理すべき目標を決定することであるが、その熟考は、穴の位置と点数から優先順位を決定する単純なものであり、その制御は行われていない点、本研究と本質的に異なる。また、Kinny らは、主に環境の変化速度と、決定された次の目標にどれだけ固執するかというコミットメントとの関係[Kinny 91]、さらに、観測コストと観測の周期がパフォーマンスに及ぼす影響[Kinny 92]を調べている。これらの研究も、Pollack らと同様に、熟考を制御しているのではない点、本研究との違いである。

実時間探索におけるプランニングのステップ数を変えて、エージェントのパフォーマンスを調べる実験がある[Russell 91]が、**SIP** のようにステップ数の調整機構を持つエージェントを用いた実験ではない。

6. ま と め

プランの成功確率を用いて熟考の制御を行う **SIP** の性能評価を、単純タイルワールドによる実験により行った。その結果、単純タイルワールドにおいて、以下のことがわかった。

- 熟考と即応の間に最適なしきい値が存在する問題がある。
 - *SIP* の最適なしきい値は、ダイナミクス、観測コストに対し頑健である。
 - *SIP* の最適なしきい値は、ダイナミクスの変化に対し、深さ固定プランニングよりも頑健である。
- 今後は、目標間にもっと複雑な因果関係のある領域、目標の発生に時空間的に偏りがある領域[磯田

93], 熟考による解の質的向上がある領域などでの評価, さらには, *SIP* マルチエージェント系の研究を考えている。

謝 辞

MARK (Multi-Agent Research Community in Kansai) における議論が役に立ちました。記して感謝いたします。

◇ 参 考 文 献 ◇

- [Dean 89] Dean, T. and Kanazawa, K.: Persistence and Probabilistic Projection, *IEEE SMC*, Vol. 19, No. 3, pp. 374-385 (1989).
- [Hanks 93] Hanks, S., Pollack, M.E. and Cohen, P.R.: Benchmarks, Test Beds, Controlled Experimentation, and the Design of Agent Architectures, *AI Magazine*, Vol. 14, No. 4, pp. 17-42 (1993).
- [磯田 93] 磯田佳徳, 山田誠二, 豊田順一: インターリーブ・プランニングとその実験的評価, 情処学会人工知能研究会, 93 AI 86-7 (1993).
- [Kinny 91] Kinny, D. and Georgeff, M.: Commitment and effectiveness of situated agents, *IJCAI 91*, pp. 82-88 (1991).
- [Kinny 92] Kinny, D., Georgeff, M. and Hendler, J.: Experiments in Optimal Sensing for Situated Agent, *PRICAI 92*, pp. 1176-1182 (1992).
- [Pollack 90] Pollack, M.E. and Ringuette, M.: Introducing the Tileworld: Experimentally Evaluating Agent Architectures, *AAAI 90*, pp. 183-189 (1990).
- [Russell 91] Russell, S. and Wefald, E.: *Do the Right Thing*, MIT Press (1991).
- [山田 96] 山田誠二: 動的環境における成功確率を用いた熟考の制御, 人工知能学会誌, Vol. 11, No. 4, pp. 645-652 (1996).

[担当編集委員: 諏訪 基, 査読者: 梅山伸二]

—— 著 者 紹 介 ——

山田 誠二(正会員)は, 前掲(Vol. 11, No. 4, p. 652)参照。



磯田 佳徳(正会員)

1967年6月28日生まれ。1991年大阪大学基礎工学部制御工学科卒業。1993年同大学院基礎工学研究科修士課程修了。同年、日本電信電話(株)入社。現在、同社ヒューマンインタフェース研究所にて知能ロボットのセンシングシステムに関する研究に従事。情報処理学会会員。



豊田 順一(正会員)

1961年大阪大学工学部卒業。1966年同大学院博士課程単位取得退学。同年、大阪大学基礎工学部助手。1969年助教授。1982年大阪大学産業科学研究所教授。工学博士。現在、概念の形成, Visual fidelity, マルチエージェントのわかりにくさに関する研究に従事。情報処理学会, 電子情報通信学会, 日本認知科学会各会員。