

特集 「エージェントの基礎と応用」

エージェントのプランニング

Planning for an Agent

山田 誠二*
Seiji Yamada

* 大阪大学産業科学研究所
ISIR, Osaka University.

1995年5月16日 受理

Keywords: planning, uncertainty, reactivity, deliberation.

1. はじめに

明確な定義もないままに、「エージェント(agent)」なる概念が、人工知能、ロボット工学、ソフトウェア工学などのさまざまな分野で広まりつつある。エージェントに抱くイメージはかなりの幅があると思われるが、エージェントが持つべき要素の一つとして、環境を観測して自分で行動を決定し、実行するという性質があげられるであろう。この行動決定を適切に行い、目標を達成するためには、エージェントがプランニング(planning)を行う必要があり、その点でプランニングはエージェントにとっての必須要素と考えられる。本解説では、このエージェントのプランニングについて、その特徴、課題、現状での解決方法について説明していく。

2. エージェントと環境

まず、ここで対象とするエージェントの概観を確認し、そのエージェントが取り囲まれている環境について考察する。

2.1 エージェントの構成

本稿で考えるエージェントのイメージを図1に示す。ここで、エージェントの構成単位は、機能的に定義され、その機能の実現方法がソフトウェアであるのか、物理的要素であるのかを問わない。つまり、ソフトウェアエージェントであろうが、ロボットであろうが構わない。図のように、エージェントを構成する機

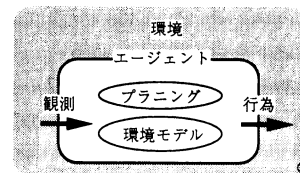


図1 エージェントの構成

能的要素は、観測、環境モデル、プランニング、行為の四つであり、エージェントと環境とは、観測と行為を通じてインタラクションが生じる。以下、各機能について簡単に説明する。

〔1〕観測

観測とは、環境からの情報の獲得である。エージェントが計算機(または、計算機ネットワーク上)で閉じた環境にいるソフトウェアの場合で、観測と行為が単なる通信の場合もあるし、エージェントがインタラクションを持つ環境が、人間や他のロボットなどの物理世界を含むときは、物理世界の情報の獲得およびその解釈を伴うことになる。いずれにせよ、これらの処理の性能およびコストが重要になる。また、物理世界を含む場合、後述する不確実性が顕著に現れる。

〔2〕環境モデル

環境内の必要な情報が、エージェントにとって内部で操作可能な表現で記述されたものを環境モデルと呼ぶ。その表現としては、STRIPS[Fikes 71]に代表される一階述語論理が古典的かつ一般的である。

〔3〕プランニング

プランニングとは、エージェントが環境モデルを操作して、自分の行為による環境の変化を予測し、適切な行動を決定することである。典型的には、後述する古

古典的プランニングのように、現在の状態を目標状態まで導くような操作系列(行為系列)を生成することを意味する。このようなプランニングで得られた操作系列をプランと呼ぶ。

〔4〕 行 為

エージェントの行為は、ロボットなどの場合の物理的行為、ソフトウェアエージェントによる通信、メッセージパッシングなどの非物理的行為のすべてを含む。ここでも、観測と同様に行為のコストおよび不確実性が問題になる。

2・2 エージェントを取り巻く環境

すでに述べてきたように、典型的なエージェントの環境としては、以下のようなものが考えられる。また、下の2種類の環境を両方含むこともあり得る。

(1) 物理世界：ロボットの環境を考えた場合、それは物理世界である。さらに、他のロボット、人間などの他エージェントを含む場合もある。

(2) 計算機あるいは計算機ネットワーク：ソフトウェアエージェントの場合、環境が一つの計算機上、あるいはインターネットなどの計算機ネットワークに限定される場合があるだろう。この環境においても、他のソフトウェアエージェントなどの他エージェントを含む場合がある。

上記の環境は、いずれもエージェントの関与しないところで、状態が時々刻々変化する動的環境である。特に、実世界とのインタラクションを持つエージェント、つまり物理世界と干渉するロボットや、他エージェントと干渉がある状況では、その環境は静的ではあり得ないといえる。

エージェントを取り巻く環境が動的であることは、本質的には環境内にエージェントにとって予測不可能な部分が存在することであり、よって、完全でかつ決定論的な環境モデルを持つことはできない。このような動的環境においては、エージェント内部に完全で決定論的な環境モデルを持つことを前提とした古典的プランニングは、破綻することが容易に想像できる。例えば、古典的プランニングは、最初に1回観測を行い、そのとき得られた情報に基づいて、時間のかかるプランニングを行うが、そのプランニング中に最初に観測した環境そのものが変化してしまうということが頻繁に起こってしまい、プランの実行は不可能になる。

3. エージェントのプランニングにおける課題

本章では、人工知能においてこれまで研究されてき

た古典的プランニング(classical planning)を説明し、エージェントのプランニングとして見た場合の古典的プランニングの課題を明確に示す。

〈古典的プランニング〉

STRIPS[Fikes 71]に代表される古典的プランニングでは、環境が一階述語論理などで表現された環境モデルで記述される。そして、環境モデルを変換する規則であるオペレータの集合と、環境モデルで記述された初期状態と目標状態が、問題として与えられる。古典的プランニングのタスクは、初期状態を目標状態に変換できるようなオペレータ系列を求めることである(図2)。この系列が、プランと呼ばれる。プランの実行は、オペレータに対応する行為をプラン中の順序で実行していく。そして、プランが完全に実行されれば、目標状態が環境において達成されることになる。

一般に、プランを探索することは、非常に多くの計算量を必要とし、積木の世界においても計算量が指数関数的に爆発する。STRIPS以降も、独立に達成できない複数の目標に対処するプランナなどが研究されてきたが、いずれも環境から完全な情報が得られ、またそれらの情報を完全に記号に変換できるという仮定のもとで、いかに効率良くプランを探索するかに主眼が置かれている。

前章のようなエージェントを囲む動的環境において、従来の古典的プランニングでは無視されてきた問題が表面化してきた。以降では、エージェントのプランニングを実現するための課題について述べる。これらは、古典的プランニングの枠組みでは、ほとんど扱われてこなかったものである。

3・1 不確実性の扱い

エージェントを取り巻く環境においては、以下のような不確実性(uncertainty)が生じる。これらにどう対処するかが重要である。

(1) 観測の不確実性：エージェントが、人間を観察したり、周りの物体の位置を計測したりするような物理世界の観測を行う必要のある場合に、顕著に観測の不確実性が生じる。また、この不確実性

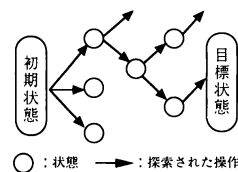


図2 古典的プランニング

は、計算機ネットワーク上でも、微妙なタイミングのずれなどにより、容易に起こり得るものと考えられる。

- (2) **行為の不確実性**：ロボットの物理的行為はもとより、計算機上のエージェントであっても、自分の行為が必ずその効果を達成することができる保証はない。例えば、ロボットが前進したつもりでも、実は動いていなかったり、ソフトウェアエージェントが、遠隔地の他エージェントにメッセージを送ったとしても、実は届いていなかったりという場合が考えられる。よって、多くの行為が不確実性を含むことになる。

3・2 再プランニング

古典的プランニングにおいて生成されたプランを実行するとき、先に述べた不確実性のため、その実行が最後まで成功裏になされるとは限らない。つまり、プランの実行が途中で止まることが起こるわけで、そのとき現在のプランに固執してその続きを実行するのか、初めからプランニングをやり直す(再プランニング)のか、などの決定を行わなければならない。このような再プランニングに関する研究は、重要であるにもかかわらず、研究例は少ない。再プランニングにおける課題は、以下のようなものが考えられる。

- (1) **再プランニングすべきか否かの決定**：途中で実行が止まってしまった場合、現在のプランの続きを何とかして実行すべきか、あるいは観測をやり直して再プランニングすべきかの決定をどのようにするか。
- (2) **再プランニングの手続き**：再プランニングを行う際に、どのような情報(例えば、現在のプラン)が必要かつ有効か。また、それらの情報を使って、いかにうまく再プランニングをするか。
- (3) **再プランニングに適したプランの表現**：再プランニングに用いるために有効な表現は何か。

3・3 熟考と即応のトレードオフ

この問題は、古典的プランニングの課題ではなく、熟考を行う古典的プランニングと即応を実現するリアクティブプランニングをいかに統合するかということである。古典的プランニングは、目標までの完全なプランを立てるので、行動の合目的性、あるいは最適性は高い。しかし、反面、変化する環境に対する即応性が低い。それに対し、次章で説明するリアクティブプランニングは、古典的プランニングより、即応性は高いものの、合目的性あるいは行動の最適性は低い。このように、熟

考と即応の間には、トレードオフがある。古典的プランニングとリアクティブプランニングという、ある意味で両極端の方法の折衷案をとって、両者を統合することが期待される。方法としては、例えば状況に応じて熟考と即応を切り換えることが考えられるが、どのような基準で切り換えるタイミングを決定すればよいか問題となる。

4. これまでの成果

前章で述べたエージェントのプランニングにおける課題を解決する研究が、これまでにいくつか行われている。本章では、それらの代表的なものについて紹介し、その研究において、前章の課題がいかに解決されているかを見ていく。また、リアクティブプランニングについては、すでに、[安部 90, 石田 90, 山田 93]の解説があるのでここでは、簡単に触れるのみとする。

4・1 リアクティブプランニング

[1] リアクティブプランニング

リアクティブプランニングでは、古典的プランニングのようなオペレータによる環境モデルの操作をいっさい行わない。よって、エージェント内部に環境モデルもオペレータも必要ない。そこで使われるのは、直接実行可能な行動からなる結論部と、センサ出力から直接判定可能な条件部で記述されるリアクティブルールである。このルールを用いたリアクティブプランニングの構成を図3に示す。観測されたセンサ情報が、複数のリアクティブルールに直接的に与えられ、それぞれのルールが並列にその条件の適用を判定する。そして、条件が満たされたルールのうち、適当なものを選択され実行される。一つのルールが実行されると、また観測を行うというループが繰り返される。これが、典型的なリアクティブプランニングである。実際には、リアクティブルールの集合の構造化や、ルール間のインタラクションの制御がなされる。なお、この種のリアクティブプランニングのシステムとして、PRS[Gorgeff 89], Pengi[Agre 87]がある。

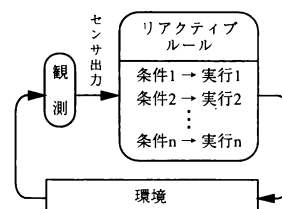


図3 リアクティブプランニング

プランニングにおいて環境モデルがどこに存在するかという視点で眺めてみると、古典的プランニングは、エージェントが環境の(決定論的)モデルを内部に持っているのに対し、リアクティブプランニングでは、環境モデルを持たずに実際に行動し、変化した環境の観測によって行動決定をしている。前者においては、環境モデルをオペレータにより変換していくことにより、エージェントが環境の変化を予測することが可能である。よって、目標状態まで達するような行為の系列を実際に行動することなしに生成することが、原理的には可能となる。しかし、再三述べているように、完全な環境モデルを持つことは現実的には不可能である。

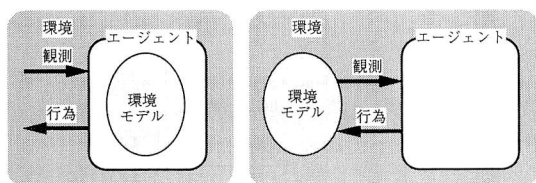
一方、リアクティブプランニングでは、エージェント内に環境モデルは存在しない。あえていえば、環境のモデルがエージェント外の環境に存在していることになる。これは、エージェントが外界から観測できる情報は、よくも悪しくも制限されており、その制限の範囲でエージェントが認識できる環境の部分が、一種の環境モデルと解釈できるからである。これらを図式すると、図4のようになる。

リアクティブプランニングでは、図4(b)のように、エージェントが環境モデルをその内部に持っていないし、そもそも環境モデルを操作するプランニングを行わないので、コストのかかるプランニングは必要ない。しかし、観測と行為の不確実性は実際に存在するわけであり、リアクティブルールのチューニングに影響を与えると考えられる。

再プランニングについては、長いプランに沿って実行をするわけではないので、おおむねリアクティブルールが、目標に近づくようにチューニングされていれば、環境の変化に対して柔軟に対応可能になる。ただし、熟考と即応の統合は、解決されていない。

〔2〕 サブサンクションアーキテクチャ

Brooksは非同期にタスクを遂行する行動(behavior)により分割されたアーキテクチャを提案した(図5)[Brooks 86]。このアーキテクチャでは、上位レベルが下位レベルの行動を抑制することにより制御がなされる。各レベルのタスクは、センサ出力を直接受け取



(a) 古典的プランニング (b) リアクティブプランニング

図4 環境モデルから見たリアクティブプランニング

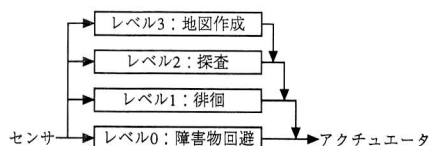


図5 サブサンクションアーキテクチャ[Brooks 86]

り並列に処理が進むため、従来の方式よりも、多重目標、多重センサ、頑健性、拡張性の点で優れているといわれている。また、下位層は前述のリアクティブプランニングと類似の働きをする。

各レベルは、有限オートマトンを接続して構成される。オートマトンは、入出力、入力バッファ、状態変数、遷移規則を持つ。また、入力にはセンサか他のオートマトンの出力に接続され、出力はアクチュエータか他のオートマトンの入力に接続される。さらに、次に示すようなゲート接続が可能である。

- (1) 抑制(inhibiter)：出力に接続され、トリガにより、その出力は一定時間削除される。
- (2) 置換(suppressor)：入力に接続され、トリガにより、その入力は一時間上位レベルの出力で置き換えられる。

構成手順としては、まず、低レベルを有限オートマトンの接続により構成する。そして、次に上位レベルを構成し、その二つのレベル間をゲート接続で結ぶ。この手順を繰り返して、サブサンクションアーキテクチャがインクリメンタルに構成できる。

サブサンクションアーキテクチャは、図3のリアクティブプランニングにおけるリアクティブルールの構造化と制御を提案したものであると解釈できる。ここでは、前述のリアクティブルールとは異なり、プランニングを上位の層に設定することが許されており、その意味で単なるリアクティブプランニングではなく、熟考と即応の統一がなされているともいえる。ただし、そのプランニングの行動が、どのようなプランニングを行うかには言及していないため、そこで古典的プランニングが行われると3章の課題が問題となる。

4・2 成功確率による熟考の制御

部分的プランニングと実行を交互に繰り返すインタリーブプランニング[McDermott 78]は、即応性と熟考性の統合を実現できる。ただし、いつプランニングを停止して、実行に切り換えるのかという問題がある。プランニングを中止することは、熟考を制御することにほかならず、熟考と即応を統合する意味でも重要なテーマであるが、このテーマに関する研究は少ない。

ここでは、筆者の研究である、プランの成功確率を

用いた熟考の制御を紹介する。熟考の制御のためには、何らかの基準が必要であるが、その基準として筆者は「動的環境において、プラン実行の成功する確率がある程度下がったときに、プランングを止めて実行を開始する」というものを用いた[山田 91]。まず、STRIPS-like な前向きな古典的プランングにより、プランが立てられる。そのプランからベイジアンネットワーク [Charniak 91, Pearl 88] が生成され、そのうえでエージェントがいる環境の変化の程度に依存した入力確率を用いて、プランの実行が完全に成功する確率が計算される。環境の観測は、プランングと並列に行われ、プランは前向きビームサーチでプランを展開していく。そして、各レベルで決定理論的に最適なプランの成功確率が、実行しきい値より下回ると、その最適なプランの実行が開始される。実行のしきい値は現在のところ、既与としている。

このインタリーブプランングの特長を次に示す。

- ・実行のしきい値を変えることによって、熟考と即応を調節できる。
- ・成功確率が、環境の変化の程度に依存する(変化が激しいとき、確率が急に下がる)ため、実行のしきい値が一定でも、環境の変化に応じて熟考が制御される。つまり、環境の変化が激しいときは熟考が抑制されて、短いプランで実行が始まり、即応性が高まる。逆に、変化が緩やかなときは、熟考を行う。

また、タイルワールドで実験が行われ、タイルの配置や得点の分布が偏っている場合にインタリーブの効果ができること、つまり、熟考性と即応性の統合が有効であることが報告されている[磯田 93]。

この研究は、熟考と即応を統一しているといえる。さらに、プランの成功確率を用いることにより、不確実性を扱っているが、プラン実行が成功した場合の世界しか考えていないので、その成功確率の計算精度は高くない。また、再プランングに対する解答は、与えられない。

4・3 確率的プランング

観測と行為の不確実性を積極的に扱うプランングが確率的プランング(probabilistic planning) [Blythe 94, Goldman 94, Kushmerick 94] であり、1993年頃から研究が始まった新しいプランングの研究分野である。これまでのいくつかの研究例があるが、その本質的な枠組みはほぼ同じなのでそれをまとめてみよう。

まず、入出力を以下に示す。

(1) 入力

- (a) 初期状態の確率分布
- (b) 確率的オペレータの集合
- (c) 目標状態
- (d) 確率しきい値

(2) 出力

確率しきい値以上の確率で目標を達成できるプラン。

今のところ、確率的プランングの研究は、記号により表現された環境モデルを用いた古典的プランングに確率を導入したものとなっている。(1)入力のうち(c)以外は、古典的プランングの枠組みにはないものなので、説明していく。

- (a) 初期状態の確率分布：考え得る初期状態を確率変数とした確率分布である。つまり、初期状態自身が観測の誤差などにより不確実性を含んでいることを意味し、例えば二つのブロックの状態が、 $on(A, B)$ なのか $on(B, A)$ なのか確定しないときに、その確率分布を $\{on(A, B):0.7\}, \{on(B, A):0.3\}$ というように割り当てる。
- (b) 確率的オペレータ：オペレータの実行による環境変化(STRIPSなら追加リスト中のリテラル)に対し、確率分布を割り当てたオペレータ。
- (d) 確率しきい値：生成されたプランの実行により目標が達成される確率が満たすべきしきい値。ユーザにより与えられる。

確率的プランングのタスクは、上記の入力から、確率しきい値以上の確率で目標状態を達成するプランを生成することである。ここで注目すべきは、確率的プランングは、古典的プランングのように解にたどり着くようなプランを生成するというよりも、部分プランでもいいから既与のしきい値以上の確率で目標状態を達成するプランを探索することである。プランングの手続きは次に示すように、しきい値以上に確実なプランが得られるまで、プラン評価とプラン修正を繰り返す。

(3) 手続き

- (a) プラン評価：現在のプランを実行した場合に、目標状態が成り立つ確率(成功確率)を計算する。その確率が、確率しきい値以上なら、現在のプランを出力して終了。そうでなければ、(b)へ。
- (b) プラン修正：現在のプランをその成功確率が増えるように修正して、新しいプランを作る。そして、(a)へ。

プラン評価の方法としては、プランをベイジアンネットワークで表現して計算する[Blythe 94]、因果関

系のネットワーク上で計算する[Kushmerick 94]などがある。また、それぞれの表現での計算を効率化する技術もいろいろ考案されている。しかし、確率を導入することにより、問題空間の状態数が爆発的に増加することと、もともとベイジアンネットワーク上での確率計算の厳密解法は指数オーダーのため、このプラン評価は非常にコストがかかる。

また、プラン修正の方法もさまざまである。基本となる古典的プランナとして、健全かつ完全な最小干渉 (least commitment) プランナ[McAllester 91]を用いる研究[Kushmerick 94]や、機械学習のほうで一般的に使われる PRODIGY[Minton 87]を用いた研究[Blythe 94]などがある。さらに、修正の戦略も、[Kushmerick 94]では、従来手法に近い後ろ向き探索が行われるが、Blytheの研究[Blythe 94]では、エージェントが直接関与しない外界の出来事(external event)をプランに取り込んでいくところに特徴がある。

エージェントのプランニングとして、確率的プランニングを見てみると、不確実性を積極的に扱っている点は

評価である。しかし、動的環境においては、プランニングが高速に行われることが必須であることから、現状ではその計算効率の改善が望まれる。また、再プランニング、熟考と即応のトレードオフについては、現状では対処していない。しかし、今後エージェントのプランニングの分野において確率を導入する試みに対し、理論的基盤を与え得るだろう。

5. ま と め

エージェントのプランニングを、動的環境におけるプランニングとして捉え、そこでの課題と現状での解決案について解説した。古典的プランニングでは、エージェントのプランニングとして、限界があることを説明し、解決案であるリアクティブプランニングなどの手法について述べた。最後に、プランニングに限らず、エージェント(あるいは、マルチエージェント)という視点を持つことにより、人工知能と周辺領域をも巻き込んだ新しい研究分野が開けることを切に期待する。

◇ 参 考 文 献 ◇

- [安部 90] 安部憲広: プランニング, 人工知能学会誌, Vol. 5, No. 6, pp. 737-747 (1990).
- [Agre 87] Agre, P. E. and Chapman, D.: Pengi: A Implementation of a Theory of Activity, *AAAI-87*, pp. 268-272 (1987).
- [Blythe 94] Blythe, J.: Planning with External Events, *UAI-94*, pp. 94-101 (1994).
- [Brooks 96] Brooks, R. A.: A Robust Layered Control System for a Mobile Robot, *IEEE Robotics and Automation*, Vol. 2, No. 1, pp. 14-23 (1986).
- [Charniak 91] Charniak, E.: Bayesian Networks without Tears, *AI Magazine*, Vol. 12, No. 4, pp. 50-63 (1991).
- [Fikes 71] Fikes, R. E. and Nilsson, N. J.: STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving, *Artif. Intell.*, Vol. 2, pp. 189-208 (1971).
- [Goldman 94] Goldman, R. P. and Boddy, M. S.: Epsilon-Safe Planning, *UAI-94* (1994).
- [Gorgeff 89] Gorgeff, M. P. and Ingrand, F. F.: Decision-Making in an Embedded Reasoning System, *IJCAI-89*, pp. 972-978 (1989).
- [石田 90] 石田 亨: 知識表現と動的世界—最近のプランニング研究から—, 人工知能学会誌, Vol. 5, No. 2, pp. 146-153 (1990).
- [磯田 93] 磯田佳徳, 山田誠二, 豊田順一: インタリーブ・プランニングとその実験的評価, 情処学会人工知能研資, 93-AI-86-7 (1993).
- [Kushmerick 94] Kushmerick, N., Hanks, S. and Weld, D.: An Algorithm for Probabilistic Least-Commitment Planning, *AAAI-94*, pp. 1073-1078 (1994).
- [McAllester 91] McAllester, D. and Rosenblitt, D.: Systematic non-linear planning, *AAAI-91*, pp. 634-639 (1991).
- [McDermott 78] McDermott, D.: Planning and Action, *Cognitive Science*, Vol. 2, pp. 71-110 (1978).
- [Minton 87] Minton, S. and Carbonell, G.: Strategies for Learning Search Control Rules: An Explanation-based Approach, *IJCAI-87*, pp. 228-235 (1987).
- [Pearl 88] Pearl, J.: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann (1988).
- [山田 91] 山田誠二: インタリーブによるリアクティブ・プランニング, 情処学会人工知能研資, 91-AI-79-8 (1991).
- [山田 93] 山田誠二: リアクティブプランニング, 人工知能学会誌, Vol. 8, No. 6, pp. 729-735 (1993).

「著者紹介」は、前掲 (Vol. 10, No. 1, p. 88) 参照。