

リアルタイム応用向き AST 学習法

6 P-3

小原雄樹 酒居敬一 阿江忠

広島大学大学院 工学研究科

1 はじめに

動的に変化する環境の中で、起こりうる全ての変化に対して必要な知識を与えてやることは不可能である。このような問題の解決には、問題処理のためのアルゴリズムを学習により獲得できるシステムの構築が望まれる。ニューラルネット学習法はパターン処理に適しており、オートマトン的な学習法は記号化された系列の学習に適している。この両者の長を融合させたハイブリッド学習法として提案された AST 学習法 [1][2] について考察する。AST 学習法では、入力パターンの認識にニューラルネット学習を用い、状態系列の学習はオートマトン的に行う。アルゴリズム獲得の学習法としての AST 学習の有用性を検証するために、コンピュータ上でシミュレーションを行った。

2 学習アルゴリズム (AST 学習)

2.1 概要

システムの目的は、複数の動いている障害物を避けて、スタート地点からゴールまで進んで行くこととする。システム上での画像情報と各々の障害物の位置・速度を入力とする。画像情報は 2 次元のビットパターン X ，速度・位置情報は 2 次元ビットパターンの付加情報 V とする。

学習は人間がシミュレータを用いて作成したスタートからゴールまでの成功例から行う。学習には状態遷移学習とコマンド学習の 2 つがある。以下で 2 つの学習を説明する。

2.2 状態遷移学習

状態遷移学習の目的は、ある状態からその次の単位時間におけるビットパターンを連想することである。学習の手続きを以下に示す。

• トレーニングセットの作成

トレーニングセットは人間が与えたスタートからゴールまでの状態軌跡とする。1 つの状態はビットパターンと付加情報から構成される。それぞれの状態間はコマンドによって連結されているとする。入力ベクトルの系列を $(X(1), V(1)) \cdots (X(k), V(k))$ と書く。

このような状態の集合をトレーニングセットとして入力する。

• 状態遷移集合の獲得

ある時間 t から $t+1$ までの状態遷移を

$$X_i(t) \xrightarrow{V_i(t)} X_i(t+1)$$

と表す。トレーニングセットから得られる状態遷移の集合を状態遷移集合と呼ぶ。

• 簡略化

リアルタイム性が要求されるシステムでは、複雑な処理の単純化を行う必要がある。状態遷移集合の要素数は通常非常に大きいため、リアルタイム処理は難しいと考えられる。そこで、2 次元ビットパターンの類似度に基づく簡略化を行い、分類する。

• LVQ 学習

単純化された状態遷移集合に対し LVQ 学習 [3] を行い、次のビットパターンを連想する。LVQ (Learning Vector Quantization) は、参照ベクトルをニューラルネットワークにおけるニューロン間のシナプス結合に対応させ学習することにより、適切な参照ベクトルを求めめる方法である。

2.3 コマンド学習

コマンド学習の目的は、ある状態からそれにふさわしい状態を連想することである。手続きは状態遷移学習と同じであるが、次の状態ではなくそのときのコマンドを学習する。

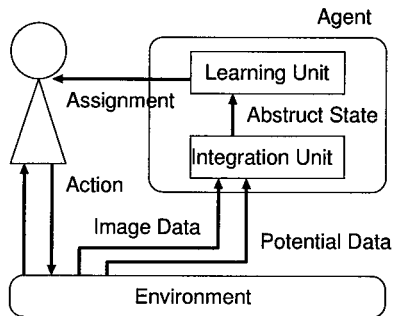


図 1: AST 学習のリアルタイム応用

3 学習結果

3.1 システムの問題点と改良

AST 学習の有効性を検証するためのシミュレータはすでに作成されていた [1]。しかし、シミュレータはゴールが線で定義されていたため、学習後のプレイヤーがゴールと並行して進む失敗例が多々見られ、ゴールを点に変更することで、問題点を解決した。

また、ゴールまでの距離によって、人間の行動も違いが現れると考え、ゴールまでの達成度を導入した。学習の際には達成度も含めて学習し、学習結果に反映させる。ここで、達成度の値によっては、プレイヤーはゴールへ向かうのみ、または障害物を避けるのみといった、極端な行動を取るようになる。達成度の値の決定は十分見当する必要があが、シミュレーションではヒューリスティックに決定している。

3.2 実行結果

付加情報 V は 10 次元のベクトルとし、各要素にはプレイヤー及び障害物の速度・位置情報を入れた。表 1 に示すパラメータを用いてシミュレーションを行った。ここで、サンプルデータ数とは

表 1: パラメータ

サンプルデータ数	30
状態数	635
学習に用いた状態数	575
縮約	$8 \times 8 \rightarrow 2 \times 2$
縮約後の状態数	16

人間が与えた成功例の数、学習に用いた状態数とは、サンプルデータからスタートとゴールの状態を除いた状態数である。このようなパラメータを用いて 300 回実行した結果を表 2 に示す。

表 2: 改良の結果

	改良前	改良後
成功回数	106	127
成功率	35.3%	42.3%

両者共に成功率が低いのは、縮約によって学習効率が下がったためだと考えられる。

4 おわりに

本稿ではアルゴリズム獲得のための学習法として AST 学習法を述べ、その有効性を検証した。

今後の課題としては、エージェントが同じ過ちを繰り返さないようにするため、追加学習の必要性がある。また、縮約は学習結果に大きく影響を与えることが予想されるので、縮約を変えての検証をする必要がある。

参考文献

- [1] 阿江忠, 荒木宏行, 日笠雄一郎, “メタ記号列の学習について —パターンと記号の統合へのメモリベース・モデルからのアプローチ—”, 信学技報, PRMU97-256, pp.39-46, 1998.
- [2] T.Ae, H.Araki, K.Sakai, “Learning Algorithm for Structured Brain Computer”, Proceedings CISST'98, edited by Hamid R.Arabnia, pp.17-24, Las Vegas, 1998.
- [3] T.Kohonen, “The Self-Organizing Map”, Proc.IEEE, Vol.78, No.9, pp.1464-1480, 1990.