

G-8

AST 学習を用いた行動支援シミュレータ

Action Support Simulator using AST Learning

尾高 啓司 酒居敬一 阿江 忠

広島大学大学院 工学研究科

1 はじめに

コンピュータ上で、人間が行う直感的な思考やあいまいな判断を再現することは現在の人工知能の分野におけるひとつの目標である。例えば人間が行動する場合の思考を再現することは、周りの状況や人間の状態に大きく左右され、一般的に知識獲得や知識表現が難しいが、多方面への応用が期待され、さまざまな方法が提案されている。我々はこのような問題の解決法の一つとして AST 学習法を提案した [1]。AST 学習法は学習により問題解決のアルゴリズムを獲得する学習法であり、問題とする環境が複雑に変化する場合に実時間で対応し未知の状況に対しても柔軟に対応することが出来る。本研究では、既存の AST 学習法を評価するためのシミュレータを利用し、まずこれを行動を支援するという目的でシミュレータを変更し、さらに、ゴール到達率の低下に対して AST 学習法の利用方法に着目し、先読み、そして追従性の考慮といった改良を行いその実験と評価を行った。

2 AST 学習

AST (Abstract State Transition) 学習法は阿江らによって脳型アーキテクチャの学習アルゴリズムとして提唱された [2]。AST 学習はパターン認識した結果、得られる (ベクトル) 記号の集合をメタ記号として扱い、メタ記号の時系列データから状態の遷移を学習する学習法である。

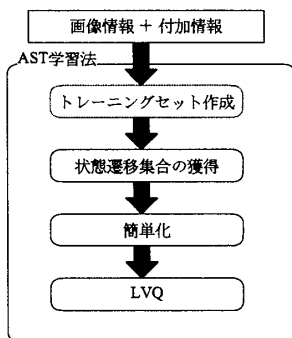


図 1: AST 学習の流れ

具体的にはメタ記号を画像情報と付加情報の集合として獲得し、その記号列 (トレーニングセット) から学

Action Support Simulator using AST Learning
Keiji Otaka, Keiichi Sakai, Tadashi Ae
Graduate school of Engineering, Hiroshima University

習する。状態の遷移は画像情報を同値類でまとめたものをオートマツ的に学習し、状態の遷移先の決定は付加情報を学習ベクトル量子化 (LVQ) の手法を用いて行う。AST 学習法はその特徴から、環境が動的に変化し学習サンプルに無い未知の状況に遭遇しても、柔軟に対応することができる。

3 行動支援シミュレータ

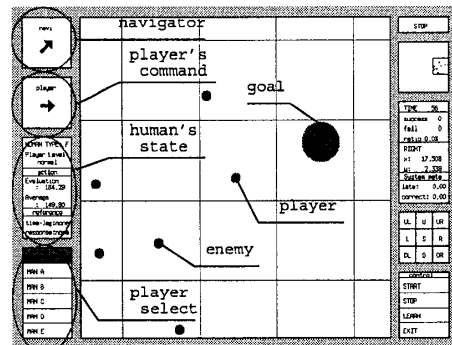


図 2: 行動支援シミュレータの外観

AST 学習の結果、システムは動的な環境において、未知の入力があつたとしても柔軟に、そしてリアルタイムに対応することができるものとなった。そこで本研究では既存の AST シミュレータを、人間に対して指示を出すことが可能な行動支援を目標としたシミュレータへと変更し、その機能を試す実験を行った。ただこの実験は上手いかず、プレイヤーのゴール到達率の低下を招いた。これは学習の結果である指示とそれに対する人間の反応の遅れによる軌跡の違いによるものであることが明らかであった。そこで以下の機能をシミュレータに実装した。

3.1 先読み

現在の状況から学習結果を用いて予想される行動に従つたとして行動した場合を予測し、予測の結果、危険に陥る (敵との距離が近付き過ぎる) と判断した場合、移動可能な方向で危険にならない方向へ暫時的に命令を変更できる機能を実装した。この機能では危険と判断する値によって性能が変化するという問題 [3] があるが、本実験では適当な関数においてヒューリスティックに見つけたよい値を固定的に与えている。

3.2 追従性の考慮

先読みが AST 学習システムの出力に対して変更を行うのに対し、入力されるデータに変更を加えて解決

を行うのが追従性の考慮機能である。プレイヤーの指示に対する遅延を計測し、それを利用して遅れる分だけ先の状態を予測する。予測は単純に現在の状況を反応が遅れる時間分だけ続けた場合にプレイヤーと敵がどの位置に移動しているかを考え、そして予測した状態に対して学習結果を適応し、その結果にあたる指示を現在の時刻で出すといった機能になっている。

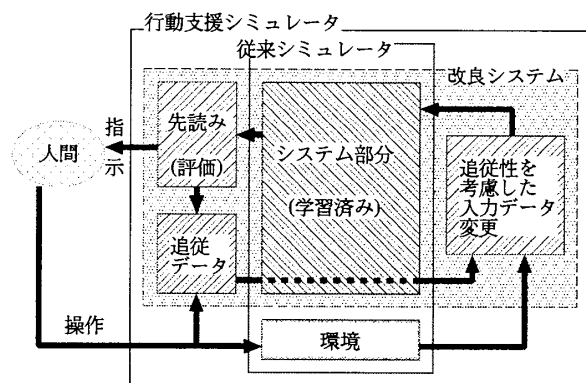


図3: 支援システムの流れ

4 実験

4.1 条件

今回の実験は以下のような条件で行った。まず本実験の目的は実装した機能(先読み・追従性考慮)によるシステムの指示の適当さの上昇をゴールに到達できる成功率を測定して確かめることである。本研究はAST学習を使ったシステムを人間が使用する場合を想定しているので、本来ならば指示の改良による成功率の増加を多くの人や場合で分けて統計的に検証を行わないといけないが、簡単に条件を同じにし、データを比較できることから今回は人間の反応を真似た疑似操作を作成し、その操作による成功率の測定と比較を行った。今回用いた疑似操作は反応の遅れと反応にかかる時間のばらつきをパラメータとしてっており、指示に対して反応が遅れる人間の反応を再現した。また比較用に改良前の状態のシステムの指示による成功率を測定した。

4.2 考察

実験の結果、先読み・追従性考慮のそれぞれについて改良前の成功率のグラフより成功率の上昇が見られた。しかし先読みにおいては、成功率の上昇は指示に対して反応の良い疑似操作に対して見られたが、逆に反応の悪い操作に対しては成功率が低下してしまった。これは人間の指示に対する反応の遅れに対してのアプローチとしては適切とは言えない結果となった。これは危険を表す関数に問題があったと考えられる。追従性を考慮した場合は全体的に成功率の低下を防いでおり有効であると言える。ただ、追従性を考慮する機能

に用いた状態の予測アルゴリズムは簡単なものであったのでさらなる改善策としては、状態系列の予測アルゴリズムの改良が上げられる。また、今回用いた先読みや追従性を考慮した処理は全体的にみるとシステム制御的であり、PID制御のようにその機能を実現することが考えられる。

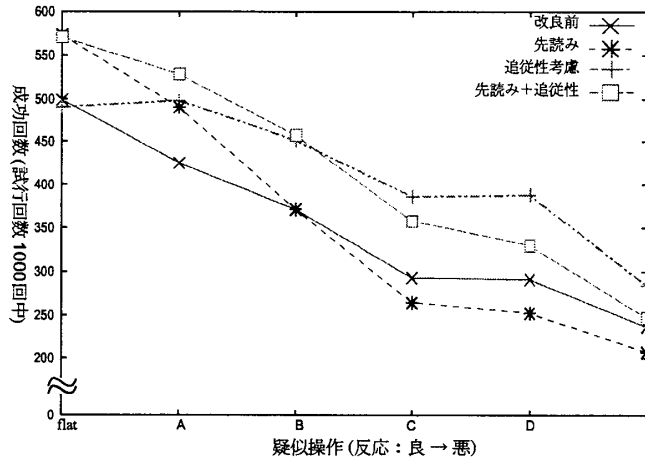


図4: 実験結果

5 おわりに

本研究ではAST学習法を用いた行動支援システムの設計を目指し、シミュレータ上で人間に対して指示を与えることとその指示を有効に利用するために先読み・追従性の考慮といった機能を付け加え実験を行った。実験の結果、予測した状態に対してAST学習法を適用した追従性を考慮する機能が有効であった。今後の課題としてシステムの実現を考えた場合、実世界からの入力とそれに応じたシステムの変更がひとつの大きな課題である。また、AST学習法そのものの改善についても考えてみると、現在のところAST学習に強化学習[3]の要素を取り込んだ学習法やマルコフモデル[4]の要素をとり入れた研究を行っており、その検証も今後の課題である。

参考文献

- [1] 小原雄樹, 酒居敬一, 阿江忠, “リアルタイム応用向きAST学習法”, 第63回情処全大, 6P-3(Sep. 2001).
- [2] 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(Mar. 1998).
- [3] Micheail G. Lagoudakis and Michael L. Litterman, “Algorithm Selection using Reinforcement Learning”, Proc. 17th International Conf. on Machine Learning 511-518(2000)
- [4] L.R.Rabiner, A tutorial on Hidden Markov Models and selected applications in speech recognition, in Proc. of the IEEE, Vol.77, No.2, Feb.1989.