

クラシファイアーシステムによる 機械学習に関する実験的考察

○川上 敬 (道女子短大) 嘉数侑昇 (北大工)

1. 緒言

近年、自動化技術の進歩に伴い、自ら最適な解を導く事の出来るような高性能を持った機械学習システムの実現が必要とされている。この機械学習の実現手法として人間の脳をまねたニューラルネットワークが広く知られているが、その一方で、例 (インスタンス) からの強力な学習手法としてのクラシファイアーシステム [Holland, Holyoak, Nisbett, and Thagard, 86] が提案されている。クラシファイアーシステムとは、任意の環境において、何らかの実行 (戦略) を誘導する為のプロダクション・ルールを学習するような機械学習システムである。そこでここでは、このクラシファイアーシステムを2次元平面上のロボットナビゲーション問題に適用し、いくつかの数値実験を行い、その学習性能について考察する。ここで扱うロボットナビゲーション問題とは工場内を移動するロボットが障害物を回避しながら目的地までの最短経路を学習する問題である。

2. ロボットナビゲーション問題

ここで対象とするロボットナビゲーション問題は、単純な2次元迷路として扱う事が出来、この迷路上でロボットが目的地までの最短経路を学習する。ここで、ロボットは次の3項目について知覚可能であるとする。

- 1) 現在の位置 (座標)
- 2) 現在地の周りにある障害物
- 3) 現在地が目的地である事

又、ロボットは、指示された方向 (東西南北のいずれか) に1座標単位だけ移動する事が出来る。クラシファイアーシステムの学習能力を検証する為に、予め迷路に関する知識は与えないものとし、環境として与えられる迷路は、線形の探索は出来ないものとする。即ち、その経路にループを含む迷路となる。

3. クラシファイアーシステム

クラシファイアーシステムは単純かつ強力な学習能力を持つ機械学習システムで、エキスパートシステムで用いられるプロダクションルールベースと類似の特徴を持っている。一般に単純なクラシファイアーシステムの構成は図1で示されるような構成要素を持つ。システムにおいて、ある環境に対する実行 (戦略) はクラシファイアーと呼ばれるストリングルールによって表現され、このクラシファイアーの有限個の集合により全ての実行が制御される。従って、システムは、学習回数を重ねるごとにクラシファイアー群を問題に適応するように進化させる事を目標とする。

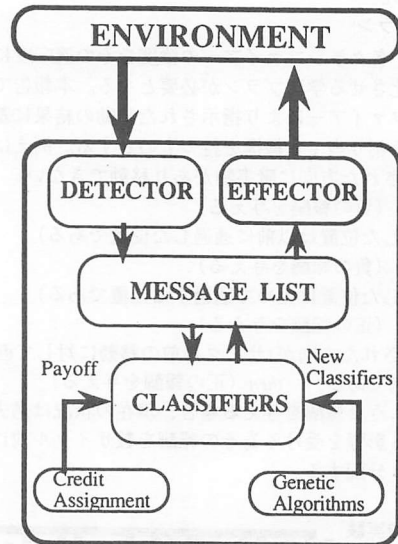


図1. クラシファイアーシステムの構成

3.1 Detector

Detectorはロボットが知覚出来る環境からの情報をクラシファイアーシステムが処理可能なようにビット列にデコードし、Message List上に転記する。

3.2 Classifierの構造

1つのクラシファイアー cf_i は、条件部分 c_i と行為部分 a_i から構成されるストリングルールである。即ち、 $cf_i = (c_i, a_i)$ 。

又、条件部と行為部はそれぞれ次のように表現される。

$$c_i = \{0, 1, \#\}^{lc}$$

$$a_i = \{0, 1\}^{la}$$

ここで、#は0,1の両者とマッチするワイルドカード記号、 lc, la はそれぞれ、予め決められた、条件部と行為部のストリング長である。ロボットナビゲーション問題に適用する場合、具体的には c_i は現在の位置座標とマッチングされ、 a_i により移動方向が指示される。又、各クラシファイアーには、環境への適応度に応じて増減される“強度”が割り当てられる。

3.3 Message List上の競争

次に、Detectorによりデコードされた環境からの情報と各クラシファイアーの条件部とのマッチングにより、条件を満たしたクラシファイアーが活性化される。ここで、活性化したクラシファイアーはそれぞれの行

為部をMessage List上に転記し、そのコードに従い次の戦略がEffectorにより翻訳され、実行される。但し、複数のクラシファイアーが活性化した場合、クラシファイアー間の競合により、その勝者を決定する。この競合は各クラシファイアーに対応する強度を基にして行われ、より高い強度を有するクラシファイアーが選ばれる。

3.4 学習プラン

ここで、各クラシファイアーの強度をその適応度に応じて変化させる学習プランが必要となる。本報告ではクラシファイアーにより指示された移動の結果に基づき報酬を割り当てる機構を持つものとする。例えば、

- 1)if (指示された方向に障害物があり移動できない)
 then (負の報酬を与える)
- 2)if (移動した位置は以前に通過した位置である)
 then (負の報酬を与える)
- 3)if (移動した位置は始めて通過する位置である)
 then (正の報酬を与える)
- 4)if (指示された方向が1サイクル前の移動に対して直進方向である) then (正の報酬を与える)

又、このような報酬を与える場合、現在の状況は過去の戦略にも影響を受ける為その報酬を数サイクル前にさかのぼり分配する。

4. 計算機実験

上述の構成に基づき計算機上にクラシファイアーシステムを構築し、以下に示すロボットナビゲーション問題への適用実験を行った。環境として与えたナビゲーション空間を図2に示す。

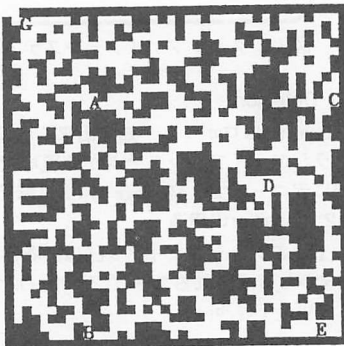


図2. ナビゲーションタスク

図中、目的地は

'G'で示され、5つの出発点を'A'~'E'で示す。本実験においてクラシファイアーの初期集団は乱数により発生させ、その集団数は2000とした。但し、クラシファイアー中の'#'の個数はストリング長の10%以下に制限し、GAによる新しいクラシファイアーの生成は行わない。又、報酬の分配については、報酬の総額Rに対して、2サイクル前までさかのぼり割り当てる。即ち、

$$S(\text{Wcf}(t)) = S(\text{Wcf}(t)) + R * 0.6$$

$$S(\text{Wcf}(t-1)) = S(\text{Wcf}(t-1)) + R * 0.3$$

$$S(\text{Wcf}(t-2)) = S(\text{Wcf}(t-2)) + R * 0.1$$

ここで、 $S(\text{Wcf}(t))$ は時刻tでの競合に勝ったクラシファイアーの現在の強度である。

これら5種のタスクそれぞれの学習結果を図3に示す。

図中、横軸は学習回数、縦軸は学習性能を示し、次式により計算された値である。

$$\text{Performance} = \frac{N_{\text{path}_{\min}}}{N_{\text{path}_i}} * 100$$

ここで $N_{\text{path}_{\min}}$ は最短経路における移動回数, N_{path_i} は学習回数iの時の移動回数である。

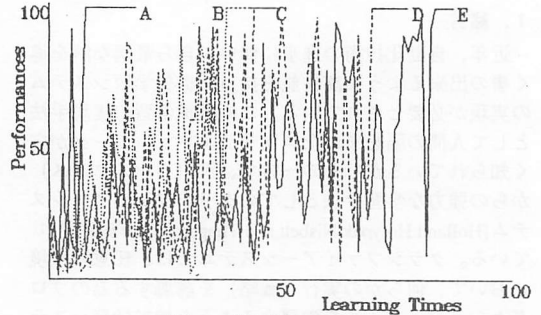


図3. 学習結果(A~E)

5. 考察

実験結果をもとに以下の事が推測される。

- 1)ナビゲーションの解は学習回数を重ねるうちに集束する(図3)。これは次第にクラシファイアー間の強度に大きな差が生じ、有効なクラシファイアーの列がただ1つ存在する事を示している。
- 2)解の収束以前では、解が振動しているが、これはナビゲーション空間上で未知の領域が減少するように広範囲の探索を行っていると考えられる。
- 3)簡単な問題(例えばタスクA)程、解の集束が速い。これは知識が少ない期間にも偶然良い解を得られる可能性が高く、それにより有効なクラシファイアーの列を発見しているものと考えられる。
- 4)本報告で用いた単純な学習プランに対しても上記の結果が得られた事から、クラシファイアーシステムの強力な学習能力が確認された。

6. 結言

クラシファイアーシステムの機械学習システムとしての学習能力をロボットナビゲーション問題を対象として実験により検証し、その高性能を確認した。又、実験にもとづき考察を加え、いくつかの性質を示した。

参考文献

- [Goldberg,89] D.E.Goldberg, :Genetic Algorithms in Search, Optimization & Machine Learning, Addison-Wesley, pp.1-25, 1989
- [Holland,75] J.H.Holland, :Adaptation in Natural and Artificial Systems, University of Michigan Press, pp.20-120, 1975.
- [Holland,Holyoak,Nisbett,and Thagard,86] J. H. Holland, K.J. Holyoak, R.E. Nisbett, and P.R.Thagard, : Induction, The MIT Press, pp.102-150, 1986.
- [Wilson and Goldberg,89] S.W. Wilson and D.E. Goldberg, : A Critical Review of Classifier Systems, in Proc. of 3rd Int. Con. on Genetic Algorithms, Laurence Erlbaum Associates, pp.244-255,1989.
- [Zhou and Grefenstette,89] H.H. Zhou and J.J. Grefenstette, :Learning by Analogy in Genetic Classifier Systems, in Proc. of 3rd Int. Conf. on Genetic Algorithms, Laurence Erlbaum Associates, pp.291-297,1989.