

予測と反省に基づく時系列の暗記学習

畝見 達夫

長岡技術科学大学 工学部 計画・経営系
unemi@voscc.nagaokaut.ac.jp

比較的粒度の小さな離散時系列を入出力データとする一種の学習メカニズムを提案し、簡単な試験例題への応用を通して解説する。各ステップの入力データおよび出力データは、その時間順序に従った連想リンクによって交互に結合され記憶される。また、過去の記憶と現在の入力を比較し、類似度の高い記憶を想起表と呼ばれる記憶領域に蓄える。この想起表中のデータを元に予想を行ない、価値の高い入力データに行き着く経験の列をプランとして採用する。想起および予想の過程では、リハーサルに基づく記憶の強化が行なわれ、重要な記憶が生き残る。

A Rote Learning Mechanism for Time-sequence Based on Prediction and Reflection

Tatsuo Unemi

Department of Planning and Management Science,
Nagaoka University of Technology
unemi@voscc.nagaokaut.ac.jp

1603-1 Kamitomioka-machi, Nagaoka, Niigata 940-21, JAPAN

This paper proposes a learning mechanism for which input/output data are formed in discrete time-sequence. Each of input and output data is memorized in a node connected in the order of time. The learner stores parts of its experience which have a relatively high correlation with current input into a working memory named *recalling table*. Referring this table, it makes prediction to employ a sequence of action as a plan. During the process of recalling and prediction, a *rehearsal-based strengthening* of memory is done to let important data survive in the memory.

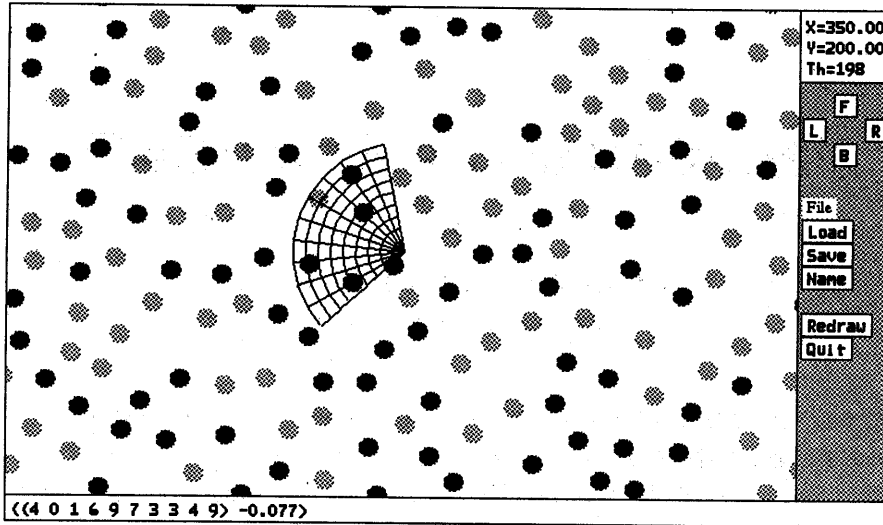


図 1: 試験例題の環境世界と虫の視野

1 はじめに

筆者は、人間の学習モデルを目指して行なった研究 [1] をもとに、そこから得られた記憶に基づく学習メカニズムの一般化を試み [2]、さらに、初期の研究で用いたものよりもいくぶん複雑な例題への応用を通し、一般化された学習メカニズムの精緻化を試み [3]、単純な環境適応問題への応用実験を行なった [4]。

ここで提案される学習メカニズムは“暗記学習”の枠組みを基本とする。暗記学習は初期の人工知能において研究された手法であり、Samuel のチェッカ・プレイヤー [5] がその代表として知られている。暗記学習では、経験をそのまま記憶し、類似の状況が再び現われたときにその記憶を参考に行動を決定するものであり、決定過程は“記憶に基づく推論” [6] と見なすことができる。

ここでは、筆者が以前に提案した学習メカニズムに“計画”を導入するための枠組みについて報告する。この拡張は、同時に、予測と反省によるフィードバック機構を組み込むことによって、より確実な学習効果、すなわち、システムの目標達成能力の向上をもたらすものである。

以下では、まず問題設定を明らかにするために、以前から用いてきた試験例題の概要について述べ、ついで、新たに提案する学習メカニズムについて解説する。

2 試験例題の概観

問題設定は“2次元平面の世界を餌を求めて歩き回る虫”である。世界は連続かつ有限なトーラス状の2次元平面で、局所的には2次元ユークリッド平面と見なせるものである。世界には一定の大きさをもつ円形の“餌”と“糞”が適当な秩序の基に配置されている。学習システムの主体である“虫”は、“視覚”、“嗅覚”、“触覚(味覚)”の3つの感覚入力と、“歩幅”、“方向変換量”の2つの運動出力をもつ。虫の目的は、“糞に触らず餌を食べること”である。すなわち、触覚(味覚)以外のデータには何の先験的価値も与えられず、虫は経験を通して目的を達成する方法を学習しなければならない。図1に環境世界の表示例を示す。

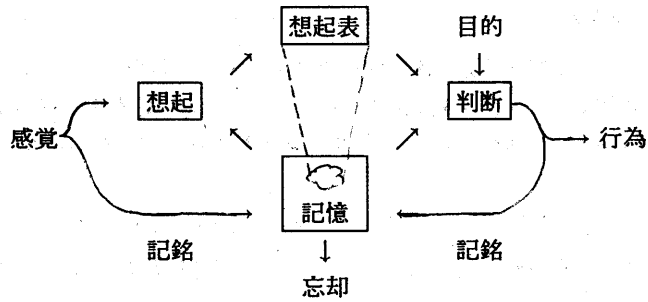


図 2: 時系列に基づく学習モデル

視覚は非負整数を要素とする固定長ベクトルである。要素の個数は視細胞の数と考えてよい。要素の値は視覚に捉えられた物体から虫までの距離である。視野は図 1 のように虫を要とする扇形の領域を角度方向と半径方向に分割したものである。

嗅覚は、餌の匂いを 1、糞の匂いを -1 とし、虫から一定距離の範囲内にある全ての物体についてそれぞれの匂いを距離で割った値の総和である。

触覚(味覚)は餌を食べたときは 1、糞に触ったときは -1、何も無ければ 0 とする。

3 学習メカニズム

以下に学習モデルの形式的表現、および試験例題における具体化の例を示す。

3.1 学習モデルの概観

図 2 に本モデルの構造を示す。想起、記憶、判断という 3 つの機能要素から構成される。それぞれの詳細については次の節で述べる。複数の結合されたヒューリスティクスに基づいて記憶要素と連想の強さを調整することにより、自らの経験を通して世界の法則を獲得し、結果として目的にそった行動決定を行なうようになる。

3.2 記憶の基本構造

過去の経験の記憶 M は印象度 p を伴う記憶要素からなる列 m の集合である。記憶要素は、感覚入力 S と運動出力 A に分けられ、 m の中では、これらが 1 つずつ交互に並べられる。すなわち、

$$M \subset \{s_1 a_1 s_2 a_2 \dots s_n a_n \mid s_i \in S, a_i \in A\} \quad (1)$$

ただし、 M に含まれる S の要素の数、および A の要素の数の最大値はあらかじめ制限されている。また、印象度はすべて $[0, 1]$ に含まれる実数である。

また、思考の文脈(注視点)を表現するために想起表 R を利用する。 R は印象度と M 中の記憶要素の組を要素とする集合で、その要素の個数の最大値があらかじめ制限されている。

以下、各機能要素における具体的計算メカニズムについて解説する。

3.3 感覚データの入力

外界からの信号を適当な時間間隔で適当な型のデータとして入力する。このとき、主にデータの変化に依存した初期印象度を各データに割り付ける。データ変化が大きい程、印象度が大きくなるように、時刻 t における 1 つの入力データ d_t に対する初期印象度の値 p_t は次のような式で計算する。

$$p_t = \alpha \cdot p_{t-1} + (1 - \alpha) \cdot (1 - \Phi(d_{et}, d_t)) \quad (2)$$

ここで、 α は $\alpha \in [0, 1)$ なる定数、 p_{t-1} は前ステップにおける初期印象度、 d_{et} は感覚データの予測値、 $\Phi(x, y)$ は $\Phi(x, y) \in [0, 1]$ なる x と y との類似度である。 α が小さいほど、変化に対して敏感になる。 d_{et} および $\Phi(x, y)$ の計算法はデータ型に依存して定義されるべきものである。次にデータがスカラー量の場合における d_{et} および $\Phi(x, y)$ の計算式の例を示す。

$$d_{et} = \beta \cdot d_{et-1} + (1 - \beta) \cdot d_{t-1} \quad (3)$$

$$\Phi(x, y) = 1 - \frac{|x - y|}{d_{max} - d_{min}} \quad (4)$$

ここで、 β は $\beta \in [0, 1)$ なる定数であり、データ値が取り得る区間 $[d_{min}, d_{max}]$ が既知であるものとする。 $\beta = 0$ のときは $d_{et} = d_{t-1}$ つまり直前のデータ値が予測値となる。 β が 1 に近づくに従って予測値に対するデータ値の影響が小さくなり予測値の時間変化は緩やかなものとなる。 $\Phi(x, y)$ の値は x と y が最も似ているつまり $x = y$ のとき 1、 x と y が最も似ていないとき 0 となる適当な関数であればよい。試験例題において、嗅覚および触覚については (4) 式を用いた。

データがスカラー量を要素とする固定長ベクトルの場合、ベクトルの各要素について上の式を応用できる。 d_{et} をベクトルとし、その n 番目の要素を d_{et}^i とすると、各要素の予測値および類似度は次のようになる。

$$d_{et}^i = \beta \cdot d_{et-1}^i + (1 - \beta) \cdot d_{t-1}^i \quad (5)$$

$$\Phi(x, y) = \frac{1}{n} \sum_{i=1}^n \left(1 - \frac{|x_i - y_i|}{d_{max} - d_{min}} \right) \quad (6)$$

ここで、 $x = (x_1, x_2, \dots, x_n)$ 、 $y = (y_1, y_2, \dots, y_n)$ である。また、ベクトルの各要素にもそれぞれに対応する印象度を与える。時刻 t における i 番目の要素 d_t^i に対する印象度 p_t^i は空間方向の変化も反映させて次のように計算する。

$$p_t^i = \lambda \cdot \{ \alpha \cdot p_{t-1}^i + (1 - \alpha) \cdot (1 - \Phi(d_{et}^i, d_t^i)) \} + (1 - \lambda) \cdot \{ \alpha \cdot p_{t-1}^i + (1 - \alpha) \cdot (1 - \Phi_{st}^i) \} \quad (7)$$

ここで、

$$\Phi_{st}^i = \frac{\Phi(d_{t-1}^{i-1}, d_t^i) + \Phi(d_{t-1}^{i+1}, d_t^i)}{2}$$

λ は $\lambda \in [0, 1]$ なる定数である。また、 $d_t^0 = d_t^1$ 、 $d_t^{n+1} = d_t^n$ として計算する。試験例題における視覚については、これらベクトルに関する式を用いている。各入力データはそれぞれの初期印象度と組にされ、想起および記録過程に送られる。

3.4 記銘と忘却

入力された感覚データ, および出力された運動データについて記銘が試みられる. 現時点の入出力データを記憶要素の形式に変換し, 直前に記銘された記憶要素の後ろに連結する.

暗記学習では経験を経る毎に新たな記憶領域を次々と消費し続けることになる. 生体の場合は生れつき一生分の記憶領域が用意されている可能性もあるが, 現状で利用可能な計算機上で実現する場合には, メモリおよび計算コストの点から, このような無制限な記憶の消費は避けるべきである. ここでは対処法として, あらかじめ要素数の最大値, すなわち記憶容量を制限しておき, 印象度の低い方から制限を上回った分の個数だけ削除するという方法を採用する. これにより, 印象度の低い部分は忘却されることとなる.

3.5 類似性に基づく想起

各入出力データについて現時点の現実のものとの記憶データ中のもとの類似度を測り, 類似性が高く印象度の高い記憶データを想起表に載せる. 但し, 既に想起表に載っているデータについては, その想起表中の印象度を増加させるに留める. C 中のデータ d_c とその初期印象度 p_c および記憶中の記憶要素 $\langle d_n, p_n \rangle$ に対して, R 上の印象度 r_t は次のように計算する.

$$r_t = \begin{cases} r + (1-r) \cdot p_c \cdot p_n \cdot \Phi(d_c, d_n) & \text{if } \langle \langle d_n, p_n \rangle, r \rangle \in R \\ p_c \cdot p_n \cdot \Phi(d_c, d_n) & \text{otherwise} \end{cases} \quad (8)$$

ベクトルの場合には, 類似度を各要素の印象度が反映するよう (6) 式を次のように拡張する.

$$\Phi(x, y) = \frac{1}{n} \sum_{i=1}^n \left(1 - \sqrt{p_x^i \cdot p_y^i} \cdot \frac{|d_x^i - d_y^i|}{d_{max} - d_{min}} \right) \quad (9)$$

ここで, $\langle d_x^i, p_x^i \rangle$ は, x の i 番目の要素, $\langle d_y^i, p_y^i \rangle$ は, y の i 番目の要素である. (8) 式は想起されたデータ 1 つ 1 つについて逐次的に処理するときの r_t の計算手続きを示しているが, 複数の相異なる根拠が同一の要素を支持するときには, 結果的にそれらの複数の印象度を合成することになる. 2 つの印象度 x, y を合成する関数 \mathcal{F}_2 は (8) 式から次のように定義される.

$$\mathcal{F}_2(x, y) \stackrel{\text{def}}{=} x + y - xy$$

これを任意個の印象度からなる集合 $X = \{x_1, x_2, \dots, x_n\}$ を引数にとる関数に拡張すると, 次のような X の要素に関する基本対称式の整式となる.

$$\begin{aligned} \mathcal{F}(X) &\stackrel{\text{def}}{=} \mathcal{F}_2(x_1, \mathcal{F}_2(x_2, \dots, \mathcal{F}_2(x_{n-1}, x_n) \dots)) \\ &= x_1 + x_2 + \dots + x_n \\ &\quad - x_1 x_2 - x_1 x_3 - \dots - x_1 x_n - x_2 x_3 - \dots - x_{n-1} x_n \\ &\quad + x_1 x_2 x_3 + \dots + x_{n-2} x_{n-1} x_n \\ &\quad \vdots \\ &\quad + (-1)^{n-1} x_1 x_2 \dots x_n \\ &= \sum_{i=1}^n \left\{ (-1)^{i-1} \sum_{S \subseteq X, |S|=i} \left(\prod_{x \in S} x \right) \right\} \end{aligned} \quad (10)$$

この関数は、引数の要素について対称であるため、その値は計算順序によらない。また、引数の全ての要素の値が $(0, 1)$ の範囲にあれば、値も $(0, 1)$ に入り、かつ値はどの引数要素よりも大きくなる。

R に含まれる要素の個数が制限されているため、上の式で計算した r_t を含む 2 項組 $\langle (d_n, p_n), r_t \rangle$ および変更を受けなかった R の要素の中から印象度の高い方から N_R 個が新たな R の要素となる。

逐次実行型の計算機では、全ての記憶データについて類似度を計算していたのでは記憶要素数に比例した計算時間を要することになる。小さな計算コストで類似性の高いものを優先的に検索する手続きが用意できれば問題はないが、一般にはそのような手続きは存在しない。但し、データ間に全順序関係が成り立つ場合は、順序に従った何段階かの分類木を作っておく方法がある [2]。

3.6 連想に基づく想起表の更新

想起表中のデータから更に記憶データ中のリンクを辿って連想されるデータで想起表の内容を置き換える。このとき R 中の新たな印象度は元の値とリンクの印象度との相乗平均を取る。想起の場合と同様に、ある記憶要素に対応する R 中の要素は 1 つのみとし、同一の記憶要素に対して複数の連想がある場合は、(10) 式で定義された関数を用いて合成した印象度を与える。これらの要素の内、印象度の高い方から決められた個数の要素だけを残し、それ以外は想起表から削除する。

3.7 記憶の強化と減衰

M および R 中の全ての印象度を各ステップ毎に減衰させる。ただし、 R の要素に含まれる記憶要素の印象度を強化する。具体的には次のように計算する。

$$p_t = \begin{cases} p_{t-1} + r \cdot (1 - p_{t-1}) & \text{if } \langle (d, p_{t-1}), r \rangle \in R \\ \eta \cdot p_{t-1} & \text{otherwise} \end{cases} \quad (11)$$

ここで、 η は $\eta \in (0, 1)$ なる定数である。この減衰と強化によって、より有用と思われるデータを記憶に残すことができる。

R に含まれる記憶要素とは、すなわち過去の経験の中で、現在の状況に似ていると思われるものに対応する。このような記憶要素に対する強化は、後述する予測に伴う強化と共に、謂わば“リハーサルに基づく記憶の強化”と名付けるべきものである。

3.8 予測に基づく計画

行動は“計画”に従って決定される。計画とは M に含まれる要素 m の部分列であり、システムにとって価値のある、すなわち、評価関数の値が正であるような感覚データを右端の要素としてもつ。ある時刻には計画は 1 つしかなく、それに含まれる運動命令データを実際の出力として選択する。システムは学習によって、置かれた状況にに対し、目的達成の意味で相応の記憶要素列を計画として採用するようになっていけるかどうか、学習性能を決定する重要な鍵となる。

計画の選択は、想起表中の記憶要素からの予想に基づいて行なう。予想とは、一時的な連想の繰り返しである。すなわち、想起表中の記憶要素から、記憶中の列を右へたどり、もし、価値の高い感覚データに出会ったならば、そこまでの部分列を計画として採用する。想起表の要素は複数あるため、計画の候補も複数存在する場合があるが、印象度、価値の大きさ、計画の長さをもとに、最適と考えられる 1 つの候補を選択する。

想起表中のどの要素からも、相応の計画が発見できなかった場合には、計画は空となる。計画が空の場合、出力すべき運動命令はランダムに選択される。初期状態すなわち何も学習していない状態では、計画は空である。あるいは、最初に試すべき行動を設定しておくことも可能である。このような先天的な計画は、生物で例えれば、本能に当ると考えられる。

また、予測の効率を良くするため、価値のある感覚データを記憶する際に、既に記憶された列を左に向って、価値を逆伝播させる。すなわち、各記憶要素に価値を表わすパラメータを付加し、列の左側に向って記憶要素を1つ経る毎に価値の絶対値を減衰させながら、値を記録する。

3.9 反省に基づく計画の撤回

採用された計画は、無条件に最後まで実行されるわけではない。計画が確実に目的の達成をもたらすとは限らないからである。計画は M の要素であるから、運動命令と共に、途中経過における感覚データも含んでいる。この感覚データを、計画の実行中に実際に得られる感覚入力と比較し、類似度の意味で大きく異なる場合には、実行中の計画を撤回し、新たな計画の選択を試みる。このときの類似度は、印象度を加味したつぎのような式により計算する。

$$\Phi'(d_p, d_c) = \sqrt{p_p \cdot p_c} \cdot \Phi(d_p, d_c) \quad (12)$$

すなわち、印象度が小さい場合には、類似性が低くても計画の撤回への影響が小さくなるようにする。

4 おわりに

ここで提案したモデルは、記憶の構造としてネットワークを基本としながらも、必ずしも神経回路網の形態、つまり閾素子と活性化伝播にこだわるつもりはない。しかし、モデルを知能のアーキテクチャとして捕え直すならば、代数的モデル表現に基づく計算機シミュレーションだけではなく、自律した演算素子を記憶要素とするネットワークによる計算モデルを考える必要があろう。

あるいは、分散協調問題解決の枠組み [8] で捉え直すことも有用と思われる。図2における感覚入力、想起、判断、行為出力といった機能単位、および記憶と想起表中の個々の要素をエージェントと考えることができる。

また、ここでは基本的には時間的な前後関係による連想のみを用いており、データは具体的レベルのみであって記憶の抽象化は一切行なわれていない。暗記学習の枠からはみ出すことにはなるが、より高度な機能を実現するにはデータの抽象化 [9] が必要となろう。

参考文献

- [1] 畝見達夫：学習機能を中心とする基礎的認知行動モデル，日本認知科学会第1回大会・発表論文集，A-2，(1984)
- [2] 畝見達夫：離散時系列の暗記学習，計測自動制御学会第8回知識工学シンポジウム，pp. 47-50，(1988)

- [3] 畝見達夫：離散時系列の暗記学習メカニズムと試験例題への応用，情報処理学会研究報告，89-AI-67-4，(1989)
- [4] 畝見達夫：系列の暗記学習による環境適応シミュレーション，計測自動制御学会 第11回知識・知能システムシンポジウム，pp. 151-156，(1990)
- [5] Cohen, P. R. and E. A. Feigenbaum (Eds.) : *The Handbook of Artificial Intelligence*, Vol. III, Chapter XIV, B.2, William Kaufmann, (1982)
- [6] Stanfill, C. and D. Waltz : *Toward Memory-Based Reasoning*, Communications of the ACM, Vol. 29, pp. 1213-1228, (1986)
- [7] 麻生英樹：ニューラルネットワーク情報処理-コネクショニズム入門、あるいは柔らかな記号に向けて-，産業図書，(1988)
- [8] 小野典彦：分散協調問題解決，淵一博監修，古川・溝口共編，知識プログラミング，共立出版，(1988)
- [9] 畝見達夫：抽象化リンクによる信号と記号の融合，日本認知科学会第5回大会・発表論文集，B-1，(1988)