

# 連想推論における逐次学習方式の定式化とその評価\*

2N-10

- 曖昧性解消に必要な文脈情報の定量化 -

角田 達彦<sup>†</sup> 田中 英彦<sup>‡</sup>  
 東京大学 工学部

## 1 はじめに

自然言語にみられる曖昧性解消の問題は、一般の推論にもあてはまる普遍的問題である。意味、構文による論理的制約、状況依存の連想による選好、そして論理的矛盾の発生により語の多義性解消を可能にするモデル PDAI&CD を提案しているが [1]、その連想記憶部 WAVE の逐次学習方法および想起性、必要な文脈情報の定量化を行い、UNIX のコマンド連想を例題にして検討する。

## 2 連想記憶部の簡易モデル化

連想推論は複数概念(単語)を入力し、概念毎に関係する複数概念をそれぞれ想起し、活性値により尤度の高い概念を選択し、記号処理部に渡すことを目的に、相互結合型神経回路網の上に実装されている。ここでは、相関型モデル [2] との違いを明確にするため、簡略化したモデルで説明する。空間を入力ベクトル空間  $S_I$ 、選択カテゴリ空間  $S_C$ 、出力ベクトル空間  $S_O$  に分け、各空間の要素を  $I_i, C_i, O_i$  とする。理想的な連想では完全な入力ベクトルに従い、 $I_i \xrightarrow{f} C_i \xrightarrow{g} O_i$  の  $f, g$  はそれぞれ上への関数とすれば一意に出力が定まる。入力ベクトルと  $I_i$  と対応する出力ベクトル  $O_i$  が等しいかで自己想起型/相互想起型の区別がされる。不完全入力を与えられた場合は、 $f$  による、カテゴリ  $C_i$  の選択性に帰着する。各カテゴリの活性値は  $C_i = \sum_j W_{ij} I_j$  により計算され、 $C_{i^*} = \max_i [C_i]$  により最尤度のカテゴリが選択される。相関型モデルでは記憶パタンの増加に伴い、パタン間の干渉が生じ、完全なパタンが想起されないことがある(疑似極小解および解の変形)が、このモデルでは出力パタンの変形は生じない。また相関型モデルでは必要メモリ数はノード数  $n$  に対し  $O(n^2)$  であるが、パタン数  $m$  に対して  $O(mn)$  である。相関型モデルの記憶パタン数は約  $0.1n$  なので [2]、同じメモリ数での記憶容量は 1 桁ほど有利になる。

重みは図 2 のように、 $W_{ij} = 1/(1+k)$  により学習する。 $k$  は入力ベクトルの一つの要素に関係あるカテゴリ数(すなわち要素あたりのリンク数)である。誤差逆伝搬法のような繰り返し学習は必要なく、またパタン全体を一度に記憶させた場合と要素を部分的に記憶させた場合で同一の学習を行うため、逐次学習に向くが、今回は一回のみ記憶させた場合について解析する。

\*Formalization and Evaluation of Sequential Learning Method for Associative Inference

<sup>†</sup>日本学術振興会特別研究員

<sup>‡</sup>Tatsuhiko TSUNODA, Hidehiko TANAKA  
 Faculty of Engineering, University of Tokyo

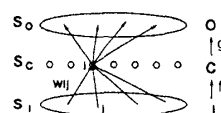


図 1: 簡易化モデル

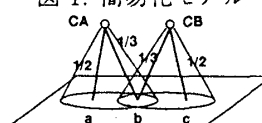


図 2: リンクの重みづけとカテゴリ選択

## 3 想起確率と必要情報数の評価

不完全パタンに対するカテゴリ選択性を解析し、曖昧性解消に必要な情報の量(必要半径と定義する)を見積もる。前述の学習方法のため、他のカテゴリを構成するパタン内にその不完全パタンが含まれている場合は、正しく分類できない。例えば図 2 の要素 b のみで CA が CB が決めることはできない。だが、a のみ、あるいは a と b が両方とも入力されると選択できる。この分類性を定量化する。要素数  $n$ 、カテゴリの関係概念要素数(リンク数)  $k$ 、パタン数(カテゴリ数)  $m$  とし、完全情報は  $k$  要素すべてを入力した場合に相当する。不完全情報は要素数  $s$  ( $0 < s < k$ ) からなる部分ベクトルである。二つのパタンが  $s$  個の要素を共有する確率は、

$$V(n, k, s) = \frac{k C_{s-1} n-k C_{k-s-1}}{n C_k}$$

となる。これを  $m$  パタンの場合に拡張するとき、一つ以上のパタンが  $s$  個の要素を共有し、他のパタンは  $0 \sim s-1$  個までの要素の共有が許される点を考慮し、必要半径  $s$  に対応した確率を求める ( $P(n, k, 0, m) = 0$ )。

$$P(n, k, s, m) \tag{1}$$

$$= \left( \sum_{r=1}^s V(n, k, r) \right)^{m-1} - P(n, k, s-1, m) \tag{2}$$

$$= (p_1 - p_2) \left( \sum_{q=0}^{m-2} p_1^q p_2^{m-2-q} \right) \tag{3}$$

$$= V(n, k, s) \left( \sum_{q=0}^{m-2} p_1^q p_2^{m-2-q} \right) \tag{4}$$

$$\left( p_1 = \sum_{r=1}^s V(n, k, r), \quad p_2 = \sum_{r=1}^{s-1} V(n, k, r) \right)$$

後述の UNIX コマンドでの実験の結果に合わせ、全要素数  $n=191$ 、カテゴリ毎の関係概念数  $k=3$ 、パタン数  $m=144$  を中心にし、必要半径、関係概念数、パタン数

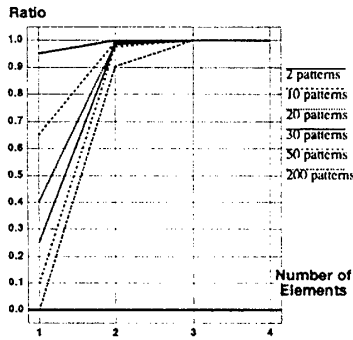


図 3:  $n=191, k=3$  の部分情報要素数と選択確率

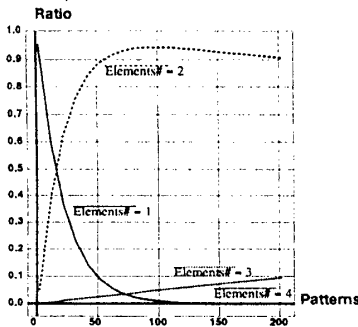


図 4:  $n=191, k=3$  の確率分布のパターン数依存性

をパラメータとして変化させたときのシミュレーション結果を図 3、4、5、6 に示す。

図 3 は  $n=191, k=3$  のもと、記録パターン数を 2 から 200 まで変化させた場合の部分パターンの要素数と想起される確率を示している。要素数が 1 のみの場合は記録パターン数が多くなると想起されにくくなるが、2 要素が入力されると、概して想起される確率が非常に高くなることわかる。図 4 は  $n=191, k=3$  での、各必要半径の割合のパターン数依存の推移を示す。必要半径 3 のグラフはパターン数とともになだらかに上昇し、必要半径 4 (即ち全要素呈示で想起不可) はほぼ 0 である。図 5 は  $n=191, k=3$  の必要半径の期待値の記録パターン数依存性を示す。数パターンの記録では要素数は 1 個示すとかなり絞れるが、50 パターンを超えると、2 個以上の要素を示さないと絞れる確率がほとんどないことわかる。だが記録パターン数の増加に対する必要半径の期待値は 2.1 付近に飽和していくことがわかる。図 6 は  $n=191$  での  $k$  を 1 から 7 まで変化させたときの必要半径の期待値を、記録パターン数毎に示している。 $k$  依存性は非常に小さく、このことはカテゴリ毎の関係概念要素数は多い方が必要な部分情報数の割合が小さくなる点で有利であることを示す。

## 4 UNIX コマンドの実験

具体的実験例として、UNIX のコマンド連想を評価し、比較する。UNIX のコマンド連想は、Workstation 上の `/usr/bin` 中の 144 コマンドに対し、コマンド毎に特徴的キーワードを連想した結果に基づく。例えば、`pwd`

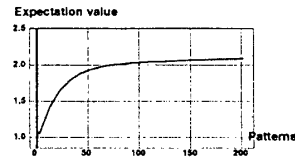


図 5:  $n=191, k=3$  の必要半径のパターン数依存性

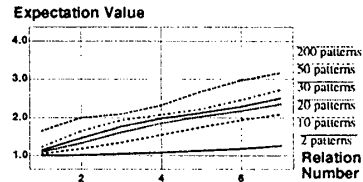


図 6:  $n=191$  の必要半径の関係概念数依存性

に対して `show`、`current`、`working`、`directory` が対応している。それらのキーワード (要素) から逆にコマンド (カテゴリ) を想起することを目的にし、部分的キーワードを与えて想起される確率を割出す。その結果を表に示す。部分的キーワードセットの全組合せ事例数 62,634 での確率と必要半径の期待値を示す。

必要半径	1	2	3	4	5	6	期待値
確率	0.40	0.41	0.18	0.02	0.00	0	1.82

全要素数は 191、カテゴリ数 (記録パターン数) は 144 である。カテゴリ毎の関係概念要素数のばらつきは以下の通りであり、平均要素数は 3.0 である。前章のシミュレーションではこの値を用いた。

関係概念要素数	1	2	3	4	5	6	7
カテゴリ数	10	42	41	37	12	1	1

実験上の必要半径の期待値 1.82 と予想値 (2.06) の違いは、以下の表のように、コマンド特有の概念と特有でない概念のばらつきが激しいためと思われる。

関係カテゴリ数	1	2	3	4	5	6
要素数	116	30	15	8	8	4
関係カテゴリ数	7	8	12	13	19	29
要素数	4	4	2	1	1	1

このように必要半径は 1.82 であるので、2 語のキーワードでかなり絞り込めることがわかる。逐次に情報が呈示される場合には、文脈として 1 語分以上の情報を保持すれば良いことがわかる。

## 5 おわりに

曖昧性解消のための情報量を見積もる方法を、連想記憶の想起性に基いて定式化し、理論値とともに、UNIX コマンド連想を例題に実験値を求め、検証した。今後は一般的な自然言語の語義の曖昧性解消に対して必要半径を求め、文脈情報量の数値化を行う予定である。

## 参考文献

- [1] 角田達彦, 白石知之, 田中英彦. PDAI & CD に基づく意味の学習および文脈依存の多義性解消. 電子情報通信学会技術研究報告, Vol. DE93-1, pp. 1-8, 1993.
- [2] S. Amari and K. Maginu. Statistical neurodynamics of associative memory. *Neural Networks*, Vol. 1-1, pp. 63-73, 1988.