

# 条件付確率表を自己組織化し、確率伝播アルゴリズムを実行する 大脳皮質のモデル(The cerebral cortex model which self-organizes conditional probability table and executes belief propagation) (論文ストーリーメモ 2007-01-09)

一杉裕志

産業技術総合研究所

**Abstract**—ベイジアンネットの条件付確率表を自己組織化マップによって獲得し、近似確率伝播アルゴリズムによってベイズ推定を行う神経回路モデルである、BESOM モデルについて述べる。

近似確率伝播アルゴリズムは、いくつかの仮定のもとで導かれる。導かれた近似アルゴリズムを実行する神経回路は、大脳皮質の解剖学特徴である6層構造およびコラム構造と、多くの点でよく一致する。

このモデルの計算量および記憶量はスケーラブルであり、その観点からも、大規模情報処理装置である脳のモデルの資格を満たしている。

## I. INTRODUCTION

近年の神経科学における精力的な研究により、大脳皮質の各領野の役割や、領野間接続の解剖学的構造に関して膨大な知見が蓄積・整理されてきている。しかし、大脳皮質の主要な機能を計算機上で再現できるほどは理解は進んでいない。

大脳皮質の機能の一部を計算機上で再現させるモデルに自己組織化マップ(self-organizing map, SOM)[1]がある。SOMとは競合学習と近傍学習を特徴とする機械学習アルゴリズムである。SOMは教師なしかつオンラインで高次元の入力をクラスタリングすることができるが、これは脳の情報処理の特徴と一致する。また、一次視覚野等で見られるコラム構造は、大脳皮質が一種のSOMであることを示唆する証拠である。

ベイジアンネット[2]もまた、大脳皮質のモデルに用いられる[6]。ベイジアンネットとは、対象を、因果関係のある確率変数の非循環有向グラフで表現するモデル化手法である。ネットワーク内の一部の確率変数の観測値が与えられると、条件付確率表を使って、残りの確率変数の値をベイズ推定できる。この際、確率伝播アルゴリズムという効率的な計算方法が使われる。ベイジアンネットには、ノイズに強いパターン認識、曖昧な情報をもとにした合理的判断、リアルタイムな動作という特徴があり、これも脳の情報処理の特徴と

一致する。大脳皮質の各領野の間には上下関係があり、上位領野と下位領野の間には双方向の結合があるが、この構造はベイジアンネットと類似している。

SOMとベイジアンネットの両方に似た機構を持つ視覚野のモデルとして、Selective Attention Model (SAM)[3]がある。SAMは、競合学習をするノードを階層的に結合したもので、各ノードの認識結果を送るボトムアップ信号と、過去の経験および文脈からの予測を送るトップダウン信号という、双方向の信号の経路を持っている。SAMにより、ノイズに強い認識や対象物の切り出しなど、脳の視覚情報処理の特徴の一部が再現されており、大脳皮質の妥当なモデルと考えられる。

SAM以外にも、トップダウン信号が予測を表すとする大脳皮質のモデルがある[4][5][6]。しかし、これらのモデルは自己組織化の機構を持っていない。

従来のモデルの別の問題は、スケーラビリティである。大脳皮質は140億個のニューロンから構成される大規模な情報処理装置であり、並列処理によって現実的な速度で動作する計算アルゴリズムが用いられているはずである。しかし、従来のモデルにおいては、計算量のオーダーは考察されていない。

本論文では、SOMとベイジアンネットの両方の特徴を備えた大脳皮質のモデルとして、BESOM (Bidirectional SOM) モデルを提案する。BESOMは、双方向結合を持つ階層的SOMである。ベイジアンネットの条件付確率表を、SOMにより自己組織化し獲得するものであるとも言える。BESOMは、SAMの学習・認識の計算式を、確率伝播アルゴリズムにもとづく理論的根拠のあるものに置き換えたものであり、それにより更なるモデルの拡張や計算機上での効率的実行の足がかりを与えている。

本論文の以降の構成は次のようになっている。II章では、BESOMモデルのアーキテクチャの概観について説明する。III章、IV章では、それぞれ学習ステップと認識ステップについて詳しく説明する。IV章ではスケーラビリティについても考察する。V章ではこのモデルと

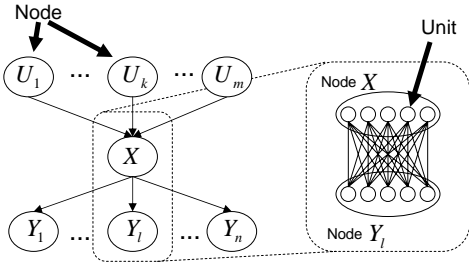


Fig. 1. BESOMのアーキテクチャ。ユニット内の構造及びユニット間のより詳細な接続についてはFig.6. 参照。

大脳皮質の解剖学的特徴との一致について述べる。最後にVI章で結論を述べる。

## II. ARCHITECTURE

BESOMは、非循環有向グラフの形に結合した「ノード」によって構成される。各ノードは、複数の「ユニット」から構成される。もし2つのノードがエッジで結ばれているなら、それぞれのノードに含まれるユニット同士は完全二部グラフで結合される(図1)。ユニット間の結合は重みを持っており、この重みは学習により変化する。(ユニットのより正確な構造についてはIV章で述べる。)ネットワーク構造は学習によって変化せず、事前知識として与えられるものと仮定する。

BESOMは、「学習ステップ」と「認識ステップ」を交互に繰り返す。

学習ステップでは、各ノードはSOMの競合層のように動作する。1つのノードは、子ノードから送られる入力ベクトルを、クラスタリングする。SOMの学習結果は、次の章で述べるように、条件付確率と見なすことができる。SOMが階層的構造を成しているので、ネットワークの上位層に行くほど、より多くの入力情報を圧縮した抽象度の高い情報を表現するようになる。

認識ステップでは、ノードのネットワークが、ベイジアンネットとして働く。各ノードは確率変数を表す。また、1つのノードの中にある各ユニットは、その確率変数を取りうる値に対応している。外界からの入力是最下端のノード(子ノードを持たないノード)の出力として与えられる。認識ステップでは、外界からの入力と、各ノードが持つ条件付確率表に従い、IV章で述べる近似確率伝播アルゴリズムを用いて、各ノードの値をMAP推定する。推定結果は、学習ステップで使われる。

BESOMを大脳皮質のモデルとして見る場合、次のよ

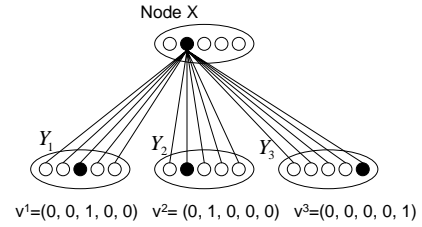


Fig. 2. 学習ステップにおけるSOMへの入力ベクトル。黒色のユニットはMAP推定結果。

うに解釈する。BESOMのノードとユニットは、大脳皮質のハイパーコラムとコラムにそれぞれ対応する。一次視覚野の場合は、各ユニットは方位選択コラムに対応する。

## III. LEARNING STEP

学習ステップでは、各ノードはSOMの競合層として働き、子ノードからの入力をクラスタリングする。

ノード \$X\$ が \$n\$ 個の子ノード \$Y\_l (l=1, \dots, n)\$ を持つとする。学習ステップでは、SOMは各子ノードの値をMAP推定した結果を入力として受け取る。ノード \$Y\_l\$ の確率変数の値 \$y\_l^l\$ の推定された事後確率を \$BEL(y\_l^l)\$ とすると、\$Y\_l\$ から入力ベクトル \$v^l\$ の要素は以下のようになる。

$$v_j^l = \begin{cases} 1 & (\text{if } BEL(y_j^l) = \max_i BEL(y_i^l)) \\ 0 & (\text{otherwise}) \end{cases} \quad (1)$$

つまり、事後確率最大のユニットに対応する要素が1で他はすべて0のベクトルである(図2)。

ノード \$X\$ の値のMAP推定結果のユニットが競合学習における勝者になる。勝者ユニットでは、参照ベクトルを入力ベクトルに近づける。ノード \$X\$ の勝者ユニット \$x\_i\$ とノード \$Y\_l\$ のユニット \$y\_j^l\$ の間の結合の重みを \$w\_{ij}^l\$ とすると、重みは次式により更新される。

$$w_{ij}^l \leftarrow (1 - \alpha)w_{ij}^l + \alpha v_j^l \quad (2)$$

ただし \$\alpha\$ は学習率である。このとき、適切な近傍関数のもとで近傍学習も行う。

ここで、近傍半径が十分小さく無視でき、また学習率 \$\alpha\$ が各ユニットの \$n\$ 回目の学習に対して \$1/n\$ になるようにスケジューリングされていると仮定すれば、重み \$w\_{ij}^l\$ は条件付確率 \$P(Y\_l = y\_j^l | X = x\_i)\$ となる。(詳細は付録参照。)

#### IV. RECOGNITION STEP

認識ステップでは、近似確率伝播アルゴリズムを実行し、各ノードの確率変数の値を推定する。本章では、確率伝播アルゴリズムを神経回路モデルに採用する際の問題点とその解決方針について述べた後、導かれた近似アルゴリズムの性質について考察する。

##### A. 条件付確率表の近似

確率伝播アルゴリズム[2] (図3) には、条件付確率表  $P(X | U_1, \dots, U_m)$  のサイズが親ノードの数  $m$  に対し指数的に増大してしまうという問題がある。

ここで我々は、同じ階層に属する個々のノードはそれぞれ異なる特徴の組を学習するので、それぞれの認識結果はお互いにほぼ独立であると仮定する。また、脳のスパース性から、あるノードの状態から予測される入力特徴の数は少ないと考える。この場合、ノード  $X$  の親ノードのうち少なくとも1つが特徴  $x$  を予測するならば、特徴  $x$  が存在すると予測してもよいだろう。この考察が正しいならば、ノード  $X$  の条件付確率表は、SOMが獲得した条件付確率の和で近似できる。

$$P(X | U_1, \dots, U_m) \approx \sum_{i=1}^m P(X | U_i) \quad (3)$$

こう仮定すれば、noisy-OR model [2] と似た状況になるので、確率伝播アルゴリズムがシンプルになり、計算量も大幅に削減できると期待できる。

式(3)が実際に成り立つかどうかは、生理学的実験もしくは計算機シミュレーションによって検証する必要がある。本論文では、我々は(3)が成り立つものと仮定して、近似確率伝播アルゴリズムを導出する。

##### B. メッセージ送信相手からの情報

図3のアルゴリズムにおける  $\pi_{Y_l}(x)$  と  $\lambda_X(u_k)$  の計算式では、メッセージを送信する相手 (受信者ノード) から来た情報を排除している。これが、アルゴリズムを複雑にする一因になっている。もしネットワークが木構造ならば、文献[6]で用いられているように、アルゴリズムは簡単になる。しかし、大脳皮質においては一般に、1つの領野は複数の上位領野への結合を持っており、ネットワークが木構造であると仮定することはできない。

本論文では逆に、1つのノードは十分多くの親ノードと子ノードを持つと仮定する。そうすると、通常あるノードの確率変数の値を支持する証拠は複数のノードから得られるから、メッセージ送信先から来た情報を排除せずを含めてしまっても、推定結果に大きな影響はないであろう。

このように考え、「メッセージ送信相手から来た情報を含める」という近似を行うことにする。

$$BEL(x) = \alpha \lambda(x) \pi(x)$$

$$\pi(x) = \sum_{u_1, \dots, u_m} P(x | u_1, \dots, u_m) \prod_k \pi_X(u_k)$$

$$\lambda(x) = \prod_l \lambda_{Y_l}(x)$$

$$\pi_{Y_l}(x) = \pi(x) \prod_{j \neq l} \lambda_{Y_j}(x)$$

$$\lambda_X(u_k) = \sum_x \lambda(x) \sum_{u_1, \dots, u_m / u_k} P(x | u_1, \dots, u_m) \prod_{i \neq k} \pi_X(u_i)$$

Fig. 3. 確率伝播アルゴリズム。

##### C. 近似確率伝播アルゴリズム

前節までに述べた2種類の近似に加え、親ノードからのメッセージが正規化されていると仮定する。

$$\sum_{u_k} \pi_X(u_k) = 1 \quad (4)$$

すると、確率伝播アルゴリズムは図4のように近似できる。(導出の詳細については付録参照。)

近似確率伝播アルゴリズムは、各子ノードから  $\lambda(y_l)$  と  $Z_{Y_l}$  の値を、各親ノードから  $BEL(u_k)$  の値を受け取り、自分自身の  $\lambda(x)$ ,  $BEL(x)$ ,  $Z_X$  の値を計算する。普通のloopyな確率伝播アルゴリズムと同様、計算を反復し、各変数の値を収束させる。

$\lambda(y_l)$  は主にボトムアップ情報だけを用いた子ノードの認識結果、 $\kappa_{u_k}(x)$  は親ノードからの情報に基づく予測、 $BEL(x)$  はベイズ推定された事後確率を意味している。

$Z_X$  は  $BEL(x)$  の正規化に使う値だが、同時にノード  $X$  における予測と観測の一致度を表している。 $\lambda_{Y_l}(x)$  の計算式に  $Z_X$  の項が現れている点が興味深い。予測と観測が一致するノードからの入力は親ノードにおける認識への影響が弱くなり、一致しないノードからの入力は影響が強くなる効果がある。(これは仮定(3)に起因する。)

なお、各変数の値は非常に簡単な計算式だけからなり、神経回路によって十分に実現可能である。特に  $\lambda_{Y_l}(x)$  および  $\kappa_{u_k}(x)$  が行う内積計算はニューロンによる実行に適している。また、条件付確率表は、これらのニューロンに接続するシナプスによる、ヘブ則学習で獲得できる。

$$\lambda_{Y_l}^{t+1}(x) = Z_{Y_l}^t + \sum_{y_l} \lambda^t(y_l) P(y_l | x)$$

$$\lambda^{t+1}(x) = \prod_{l=1}^n \lambda_{Y_l}^t(x)$$

$$\kappa_{U_k}^{t+1}(x) = \sum_{u_k} P(x | u_k) BEL^t(u_k)$$

$$\pi^{t+1}(x) = \sum_{k=1}^m \kappa_{U_k}^{t+1}(x)$$

$$p^{t+1}(x) = \lambda^{t+1}(x) \pi^{t+1}(x)$$

$$Z_X^{t+1} = \sum_x p^{t+1}(x)$$

$$BEL^{t+1}(x) = p^{t+1}(x) / Z_X^{t+1}$$

Fig. 4. 導出された近似確率伝播アルゴリズム。

#### D. スケーラビリティ

学習時の入力ベクトルはスパースなので条件付確率表もスパースになる。入力ベクトルの非0の要素の数は定数個（子ノードの数）であり、各子ノード内のユニットの数  $s$  には依存しない。それを学習する条件付確率表もおそらく同様である。この性質を利用することにより、記憶量・計算量のオーダーを下げるができる。神経回路においては、重みが0であるシナプスを持たないことに相当し、これにより、必要なシナプス数の爆発が抑えられる。

1つのユニットが持つ条件付確率表（参照ベクトル）の表現に必要な記憶量は、本来は  $O(s)$  であるが、スパース性を利用すれば  $O(1)$  となる。

$\lambda_{Y_l}(x)$ 、 $\kappa_{u_k}(x)$  の内積計算にかかる時間も本来  $O(s)$  だが、スパース性を利用すれば  $O(1)$  となる。従って、ノード内のすべての内積計算にかかる時間は並列計算の場合  $O(1)$  である。他の変数の計算量も並列計算の場合  $O(\log s)$  以下である。

以上のように、本論文の近似アルゴリズムはスケラブルである。従って、記憶量・計算量の観点からは大脳皮質の情報処理アルゴリズムのモデルとしての資格を満たしていると言える。

### V. 神経科学的知見との対応

#### A. 大脳皮質の解剖学的特徴

大脳皮質は6層構造をしている。大脳皮質の領野間は双方向結合を持っており、その結合の仕方には以下に述べる規則性があることが知られている[7]。ボトムアップの結合は主にIII層からIV層に向かう。V層から

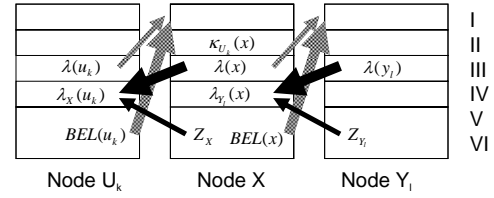


Fig. 5. 近似確率伝播アルゴリズムと6層構造の対応。

IV層への結合がある場合もある。トップダウンの結合は主にV・VI層からI層に向かう。III層からI層への結合も少しある。

また、コラム内の解剖学的構造から、コラム内ではIV層に入力された情報はII・III層を經由しV層から出力されると考えられる[8]。

この2つの知見を合わせると、コラム内における情報処理の途中結果であるIII層の情報が上位領野に送られ、最終結果のV層の情報は下位領野に戻されるという、とても奇妙な構造をしていることになる。この構造の機能的意味については知られていない。

#### B. 近似アルゴリズムとの対応

近似確率伝播アルゴリズムに現れる7つの変数のうち、ノード間の通信に関わる5つの変数を、領野間の結合規則に当てはめてみると、図5のようなになる。ただし、内積を計算する  $\kappa_{u_k}(x)$  はI層ではなく星状細胞の多いII層とした。 $Z_x$  はIII層の可能性もあるが、 $BEL(x)$  との関連の深さからV・VI層とした。

図に示されているように、近似アルゴリズムは領野間の結合の規則と素直に対応が付き、偶然とは考えられない。（ただし、III層からI層へのトップダウンの結合については、このモデルでは説明できない。）

7種類の変数を、解剖学的知見と一致するようにコラム内に配置してみたものが図6である。前述の、IV層、II・III層、V層の順の情報の流れは、変数  $\lambda_{Y_l}(x)$ 、 $\lambda(x)$ 、 $p(x)$ 、 $BEL(x)$  の順の情報の流れに対応している。（解剖学的知見によるとさらにV層、VI層、IV層の順で情報が流れる[8]が、その経路は図6には現れていない。）

またこの図からは、1)コラム内では垂直方向だけでなく、ほとんどの情報処理が行われること、2)I層およびIV層

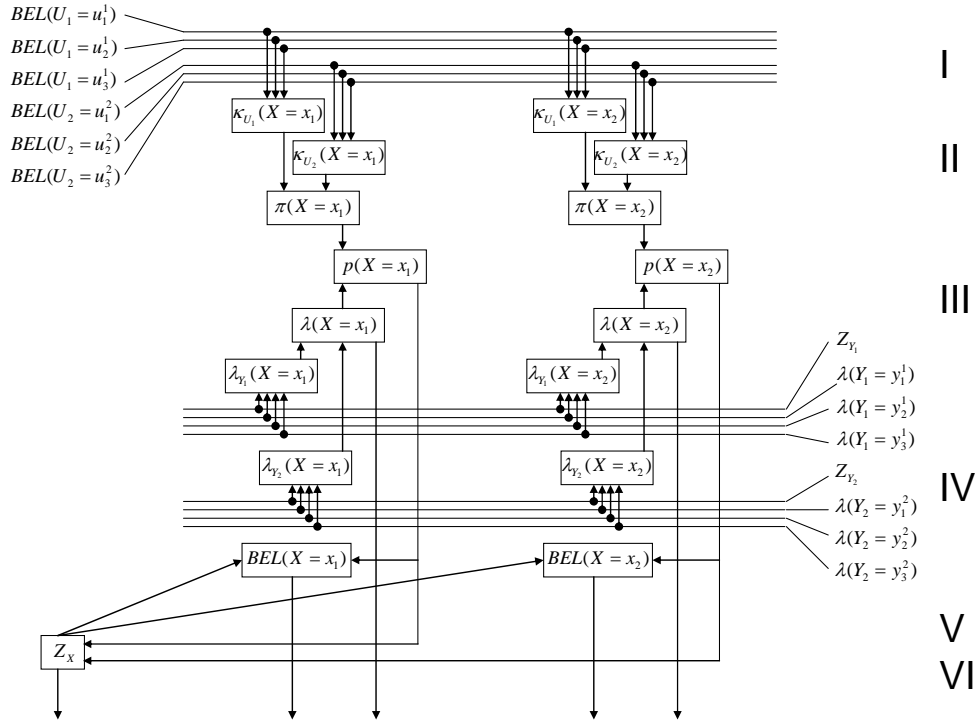


Fig. 6. ユニット内の構造。ノードXの2つのユニット  $X=x_1, X=x_2$  内の変数間の情報の流れを図示した。ノードXは2つの親ノード  $U_1, U_2$  および2つの子ノード  $Y_1, Y_2$  より情報を受け取る。図の中のそれぞれの変数は、入力に対する足し算、掛け算、割り算のどれかを行う。ユニット内の構造は大脳皮質のコラム内の構造と多くの点で類似している（本文参照）。

には水平線維が見られること、3)II層およびIV層には星状細胞と呼ばれる細かい細胞が多数あること、といった解剖学的知見との一致を見ることができる。

### C. 同じ階層の領野間の接続の問題

大脳皮質では、上下関係のない、同じ階層に属する領野間にも接続がある場合があるが、それは BESOM モデルでは説明できない。選択的注意など、確率伝播アルゴリズムとは違う目的の通信路である可能性が考えられる。

## VI. CONCLUSION

双方向に結合するSOMをベイジアンネットワークとみなして近似確率伝播アルゴリズムを導き出し、それが大脳皮質の6層構造およびコラム構造とよく対応し、またスケラブルであることを示した。このモデルは、皮質内のニューロン発火の意味を説明しており、それは生理学的実験により検証可能である。

本モデルをシミュレーションし大規模実験を行うためには、同じ階層に属するノードどうしを独立にする機構を明らかにしなければならない。我々は現在この問題に取り組んでいる。この問題が解決すれば、条件付確率表の近似式(3)の妥当性を検証するシミュレーションが可能になるだろう。

本論文で述べた BESOM の基本機能に加え、SOMへの実質的な入力次元を減らす各種の選択的注意の機構、短期記憶の機構、時系列学習を可能にする機構、などのモデルの拡張を行う必要があると考えている。

BESOMモデルは非常に高い表現力を持っている。我々は、運動前野と補足運動野による時系列運動の獲得の機構や、前頭前野背外側部による状態遷移表の獲得と行動計画の機構についても、BESOMモデルで表現することに取り組み始めている。

今後は、より多くの文献を調査しモデルを拡張するとともに、脳の主要な機能を計算機上で再現させることを目指す。

## APPENDIX

学習ステップで獲得されるユニット間の重みが、条件付確率とみなせることは、以下のように示すことができる。ノードXのユニット  $x_i$  がMAP推定結果になった回数を  $n$ 、Xの子ノードYからのn回目の入力ベクトルの要素を  $v_j(n) \in \{0,1\}$ 、その学習結果を  $w_{ij}(n)$  とする。また、n回のうちYのユニット  $y_j$  もMAP推定結果で

あった回数を  $m(n) = \sum_{i=1}^n v_j(i)$  とする。学習率を

$\alpha = 1/n$  とする。また、 $w_{ij}(1) = v_j(1) = m(1)$  とする。

このとき  $w_{ij}(n)$  ( $n > 1$ ) の値は次のようになる。

$$\begin{aligned} w_{ij}(n) &= (1 - \alpha)w_{ij}(n-1) + \alpha v_j(n) \\ &= ((n-1)w_{ij}(n-1) + v_j(n))/n \\ &= ((n-1)(m(n-1)/(n-1)) + v_j(n))/n \\ &= (m(n-1) + v_j(n))/n \\ &= m(n)/n \end{aligned}$$

この値はユニット  $x_i$  が MAP 推定結果になった回数のうち子ノード  $y_j$  も MAP 推定結果であった回数の比率である。MAP 推定結果が正しいと仮定するならば、すなわちこの値は条件付確率  $P(y_j | x_i)$  である。

認識ステップで用いられる近似確率伝播アルゴリズムは以下のように導出する。

まず、 $\pi(x)$  は(3)を用いて次のように近似できる。

$$\begin{aligned} \pi(x) &= \sum_{u_1, \dots, u_m} P(x | u_1, \dots, u_m) \prod_i \pi_X(u_i) \\ &\approx \sum_{u_1, \dots, u_m} \left( \sum_k P(x | u_k) \right) \prod_i \pi_X(u_i) \\ &= \sum_{u_1, \dots, u_m} \sum_k P(x | u_k) \prod_i \pi_X(u_i) \quad (6) \\ &= \sum_k \sum_{u_k} P(x | u_k) \pi_X(u_k) \sum_{u_1, \dots, u_m / u_k} \prod_{i \neq k} \pi_X(u_i) \quad (7) \end{aligned}$$

ここで、正規化の仮定(4)から次の式が成り立つ。

$$\sum_{u_1, \dots, u_m / u_k} \prod_{i \neq k} \pi_X(u_i) = 1 \quad (8)$$

したがって、(7)は次のようになる。

$$\pi(x) = \sum_k \sum_{u_k} P(x | u_k) \pi_X(u_k)$$

$\pi_{Y_l}(x)$  は「メッセージ送信相手の情報を含める」ことで、次のように近似できる。

$$\begin{aligned} \pi_{Y_l}(x) &= \pi(x) \prod_{j \neq l} \lambda_{Y_j}(x) \\ &\approx \pi(x) \prod_j \lambda_{Y_j}(x) \end{aligned}$$

$$= \lambda(x) \pi(x)$$

$\lambda_X(u_k)$  の計算式の中に含まれる次の式は、(3)および(8)を用いて次のように近似できる。

$$\begin{aligned} &\sum_{u_1, \dots, u_m / u_k} P(x | u_1, \dots, u_m) \prod_{i \neq k} \pi_X(u_i) \\ &\approx \sum_{u_1, \dots, u_m / u_k} \left( \sum_{j \neq k} P(x | u_j) + P(x | u_k) \right) \prod_{i \neq k} \pi_X(u_i) \\ &= \sum_{u_1, \dots, u_m / u_k} \sum_{j \neq k} P(x | u_j) \prod_{i \neq k} \pi_X(u_i) \\ &\quad + P(x | u_k) \sum_{u_1, \dots, u_m / u_k} \prod_{i \neq k} \pi_X(u_i) \\ &= \sum_{u_1, \dots, u_m / u_k} \sum_{j \neq k} P(x | u_j) \prod_{i \neq k} \pi_X(u_i) + P(x | u_k) \end{aligned}$$

この式の第一項は「メッセージ送信相手の情報を含める」近似を適用すると式(6)に一致するので、次のように近似できる。

$$\begin{aligned} &\sum_{u_1, \dots, u_m / u_k} \sum_{j \neq k} P(x | u_j) \prod_{i \neq k} \pi_X(u_i) + P(x | u_k) \\ &\approx \sum_{u_1, \dots, u_m} \sum_j P(x | u_j) \prod_i \pi_X(u_i) + P(x | u_k) \\ &\approx \pi(x) + P(x | u_k) \end{aligned}$$

上記式を代入することで、 $\lambda_X(u_k)$  は次のように近似できる。

$$\begin{aligned} \lambda_X(u_k) &= \sum_x \lambda(x) \sum_{u_1, \dots, u_m / u_k} P(x | u_1, \dots, u_m) \prod_{i \neq k} \pi_X(u_i) \\ &\approx \sum_x \lambda(x) (\pi(x) + P(x | u_k)) \\ &= \sum_x \lambda(x) \pi(x) + \sum_x \lambda(x) P(x | u_k) \end{aligned}$$

以上の結果を整理すると、図4のアルゴリズムになる。

## REFERENCES

- [1] T. Kohonen, Self-Organizing Maps. Springer-Verlag, 1995.
- [2] Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann, 1988.
- [3] Neural network model for selective attention in visual-pattern recognition and associative recall, APPLIED OPTICS 26 (23): 4985-4992 Dec 1 1987.
- [4] Kawato M., Hayakawa H., Inui T.: A forward-inverse optics model of reciprocal connections between visual areas. Network: Computation in Neural Systems 4, 415-422, 1993.
- [5] R.P.N. Rao and D.H. Ballard, Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects, Nature Neuroscience, Vol.2 No.1, pp.79-87, Jan 1999.
- [6] George, D. Hawkins, J., A hierarchical Bayesian model of invariant pattern recognition in the visual cortex, in proc. of IJCNN 2005, vol. 3, pp.1812-1817, 2005.
- [7] Pandya, D.N. and Yeterian, E.H., Architecture and connections of cortical association areas. In: Peters A, Jones EG, eds. Cerebral Cortex (Vol. 4): Association and Auditory Cortices. New York: Plenum Press, 3-61, 1985.
- [8] Gilbert, C.D., Microcircuitry of the visual-cortex, Annual review of neuroscience, 6: 217-247, 1983.