

# 制限付きベイジアンネット BESOM の正則化の一手法

第96回人工知能基本問題研究会(SIG-FPAI)

2015-01-13

産業技術総合研究所

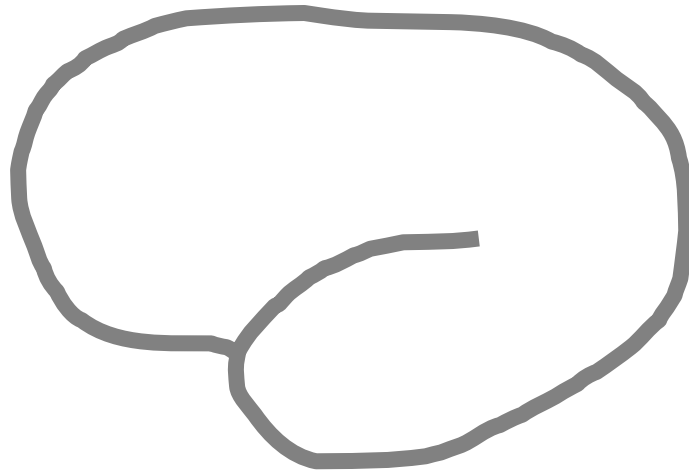
一杉裕志

# 発表の流れ

- 背景：
  - 全脳アーキテクチャの解明を目指す
- 大脳皮質ベイジアンネットモデル
- 過適合・局所解の問題の解決：正則化
- 現段階での性能評価

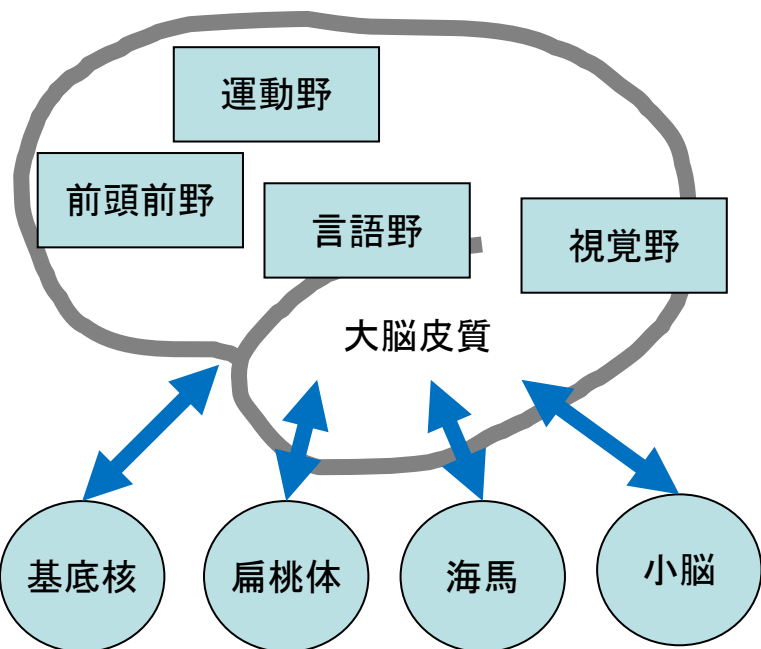
# 私の研究の目標

- 脳を模倣して「人間のような知能」を持つ機械を作る。



# 脳の各器官のモデル

脳を構成する主要要素



## 脳の各器官の機械学習装置としてのモデル

大脳皮質: SOM、ICA、ベイジアンネットワーク

大脳基底核、扁桃体: 強化学習

小脳: パーセプトロン、リキッドステートマシン

海馬: 自己連想ネットワーク

## 主な領野の情報処理装置としての役割

視覚野: deep learning

運動野: 階層型強化学習

前頭前野: 状態遷移機械?

言語野: チャートパーサ?

脳の知能に関係する主要な器官の計算論的モデルは**不完全ながら出そろってきている**。これらの器官の間の連携のモデルを考えることで、脳全体の機能の再現に挑戦すべき時期に来ている。

いまもっと不足しているのは、  
神経科学と機械学習の知識を  
持ったやる気のある研究者

# 全脳アーキテクチャ勉強会 開催実績

参加者は  
技術者と研究者が半々

- 第1回 2013年12月 **開催趣旨説明** 約100名参加
  - － 講演者：産総研 一杉裕志、東大 松尾豊、富士通研 山川宏
- 第2回 2014年1月「**大脳皮質と Deep Learning**」約250名参加
  - － 講演者：産総研 一杉裕志、筑波大学 酒井宏、PFI 得居誠也
- 第3回 4月「**海馬と SLAM**」約200名参加
  - － 講演者：はこだて未来大 佐藤直行、産総研 横塚将志、富士通研 山川宏
- 第4回 6月 **関西編** 約100名参加
  - － 企画：理研 高橋恒一、講演者：産総研 一杉裕志、東大 松尾豊、富士通研 山川宏、NICT CiNet 西本伸志、理研 泰地真弘
- 第5回 7月「**意思決定**」約200名参加
  - － 講演者：産総研 一杉裕志、奈良女子大 新出尚之、グーグルジャパン 牧野貴樹
- 第6回 7月「**統合アーキテクチャ**」約170名参加
  - － 講演者：富士ゼロックス 岡本洋、玉川大 大森隆司、NII 市瀬龍太郎
- 第7回 9月「**感情**」約200名参加
  - － 講演者：玉川大大森隆司、京大藤田和生、東京慈恵医科大渡部文子、AGI光吉俊二
- 第8回 11月10日(月)「**時系列学習**」約200名参加
  - － 講演者：山口大宮崎真、電通大山崎匡、早稲田大尾形哲也

過去の講演資料はほぼすべてネットで見られます。

全脳アーキテクチャ

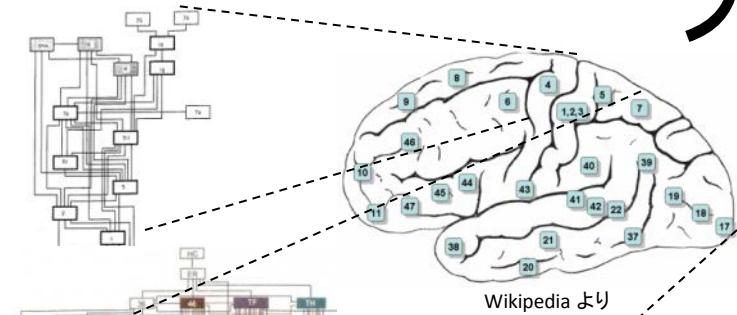
検索

# 次回 全脳アーキテクチャ勉強会

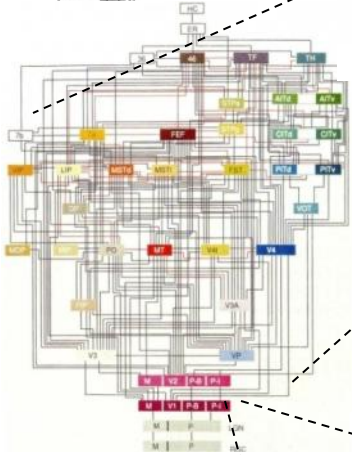
## 第9回2月4日(水)「表現学習」

- 場所: 東京 リクルートホールディングス本社(グラントウキョウサウスタワー) 定員180名
- 時間 18:00-21:00 その後懇親会
- 講演者
  - － 東京大学 酒井邦嘉
  - － 産業技術総合研究所 林隆介
  - － 立命館大学 谷口忠大
- 詳細は Facebook グループ等ではアナウンス予定
  - － 「全脳アーキテクチャ勉強会」参加者 約900人

# 大脳皮質

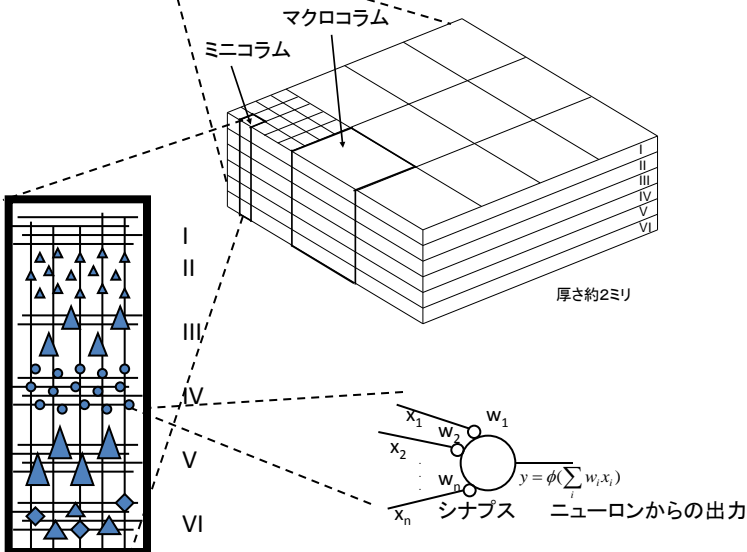


Wikipedia より



Daniel J. Felleman and David C. Van Essen  
Distributed Hierarchical Processing in the Primate  
Cerebral Cortex  
Cerebral Cortex 1991 1: 1-47

- 脳の様々な高次機能（認識、意思決定、運動制御、思考、推論、言語理解など）が、**たった50個程度**の領野のネットワークで実現されている。





# Deep Learning の源流

## Neocognitron [Fukushima 1980]

当時の視覚野の知見をもとに作られた、実際に手書き文字認識を行うモデル。  
たたみ込みNNと同じ構造を持つ。

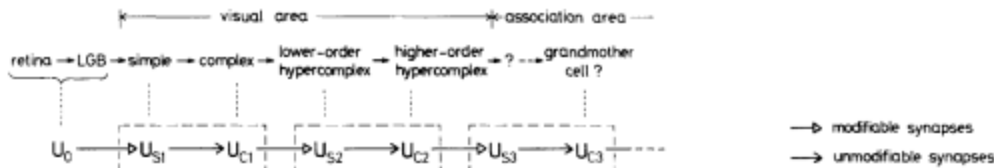


Fig. 1. Correspondence between the hierarchy model by Hubel and Wiesel, and the neural network of the neocognitron

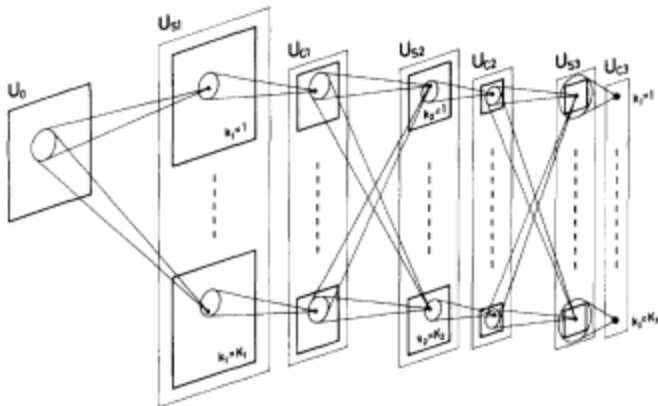


Fig. 2. Schematic diagram illustrating the interconnections between layers in the neocognitron

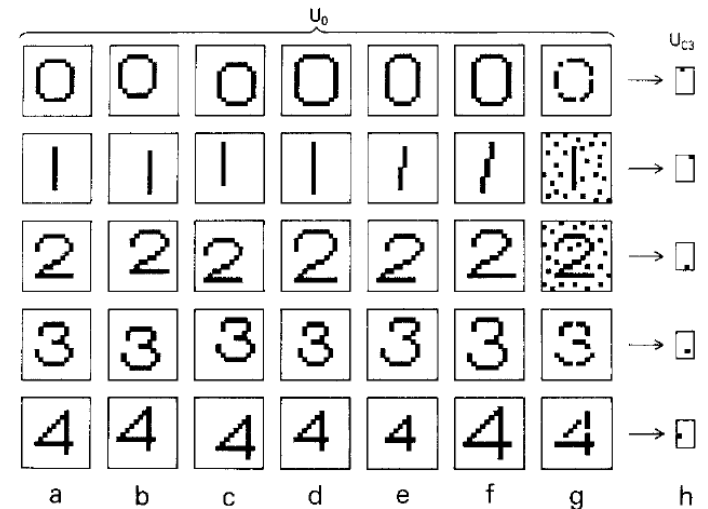
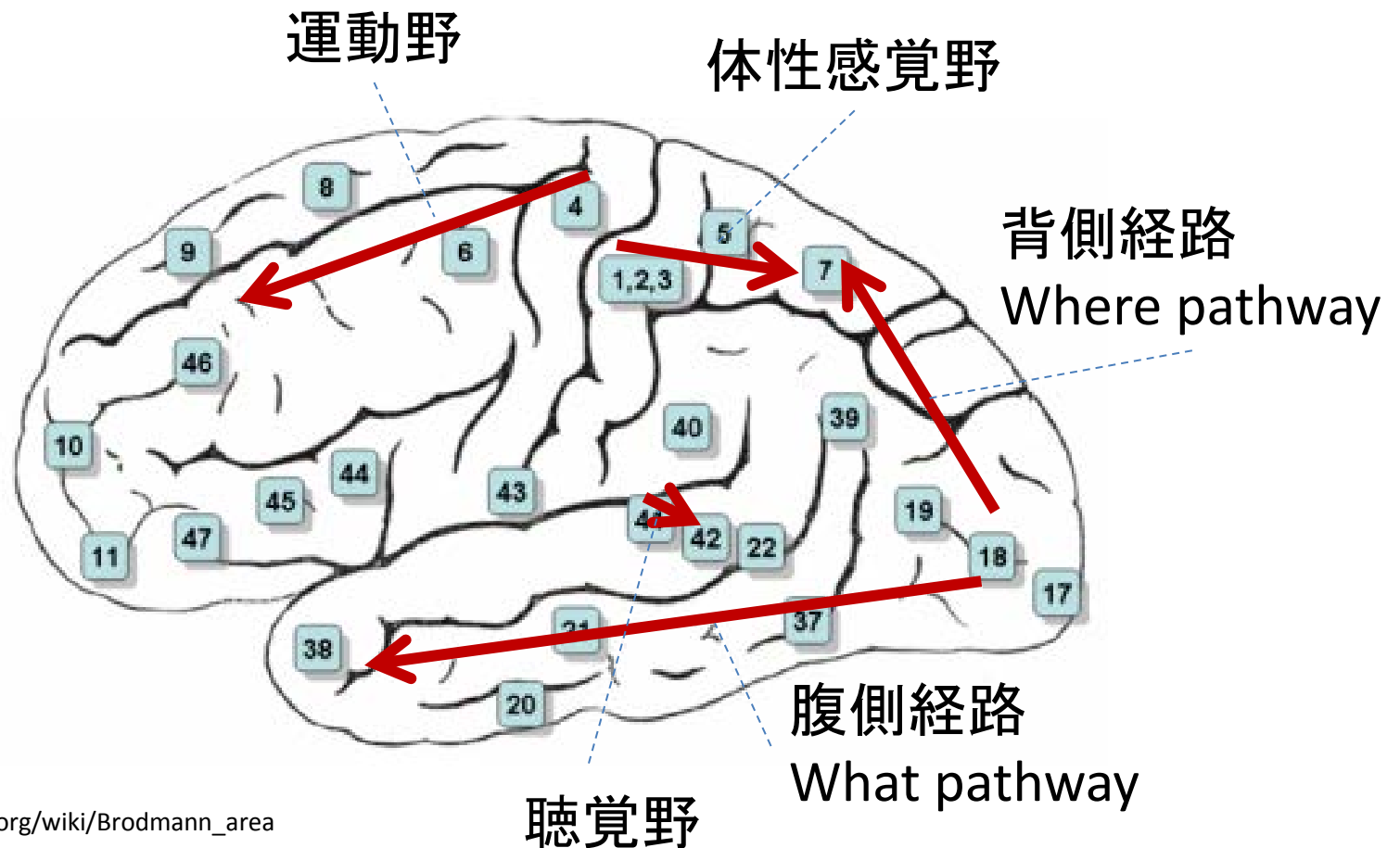


Fig. 6. Some examples of distorted stimulus patterns which the neocognitron has correctly recognized, and the response of the final layer of the network

K. Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics, 36(4): 93-202, 1980.

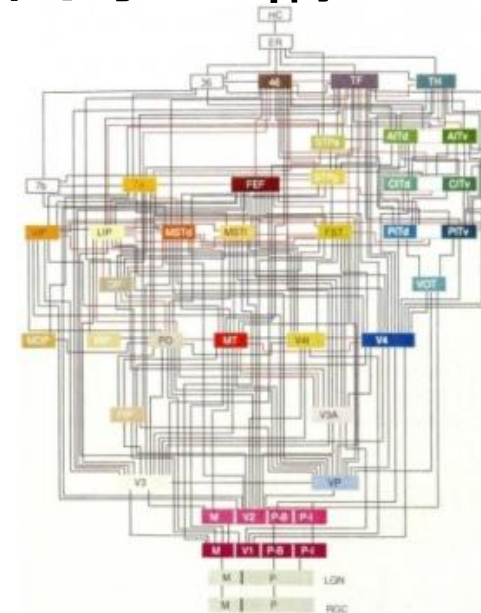
# 大脳皮質にみられる階層的構造

脳はおそらく様々な情報処理に Deep Learning を用いている。



# ベイジアンネットを使った 大脳皮質モデル

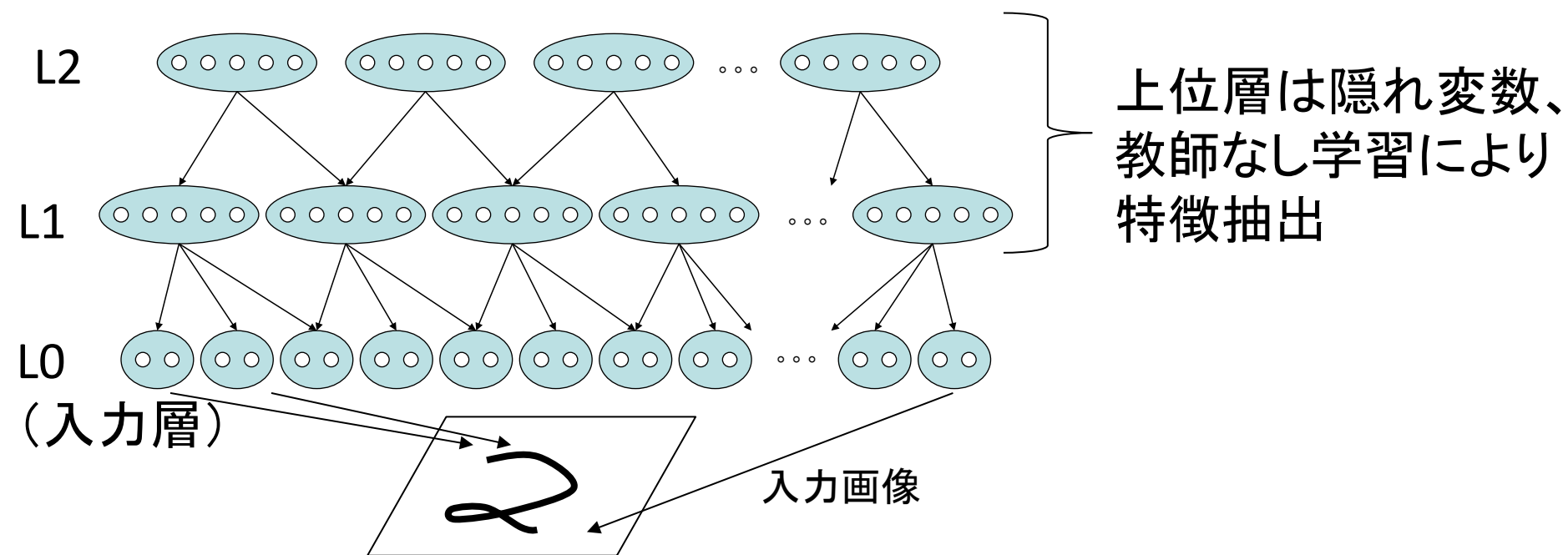
- 視覚野の機能、運動野の機能、解剖学的構造、電気生理学的現象などを説明
  - [Lee and Mumford 2003]
  - [George and Hawkins 2005]
  - [Rao 2005]
  - [Ichisugi 2007] [Ichisugi 2010] [Ichisugi 2011] [Ichisugi 2012]
  - [Rohrbein, Eggert and Korner 2008]
  - [Hosoya 2009] [Hosoya 2010] [Hosoya 2012]
  - [Litvak and Ullman 2009]
  - [Chikkerur, Serre, Tan and Poggio 2010]
  - [Hasegawa and Hagiwara 2010]
  - [Dura-Bernal, Wennekers, Denham 2012]



Daniel J. Felleman and David C. Van Essen  
Distributed Hierarchical Processing in the Primate  
Cerebral Cortex  
Cerebral Cortex 1991 1: 1-47

大脳皮質は、Deep Learningと同じ構造をもった  
巨大なベイジアンネットらしい。

# Deep Learningと同じ構造をもった ベイジアンネットの一例



楕円は確率変数  
白丸は確率変数がとり得る値

# 「ベイジアンネットによるDeep Learning」が 工学的にも有望と考える理由

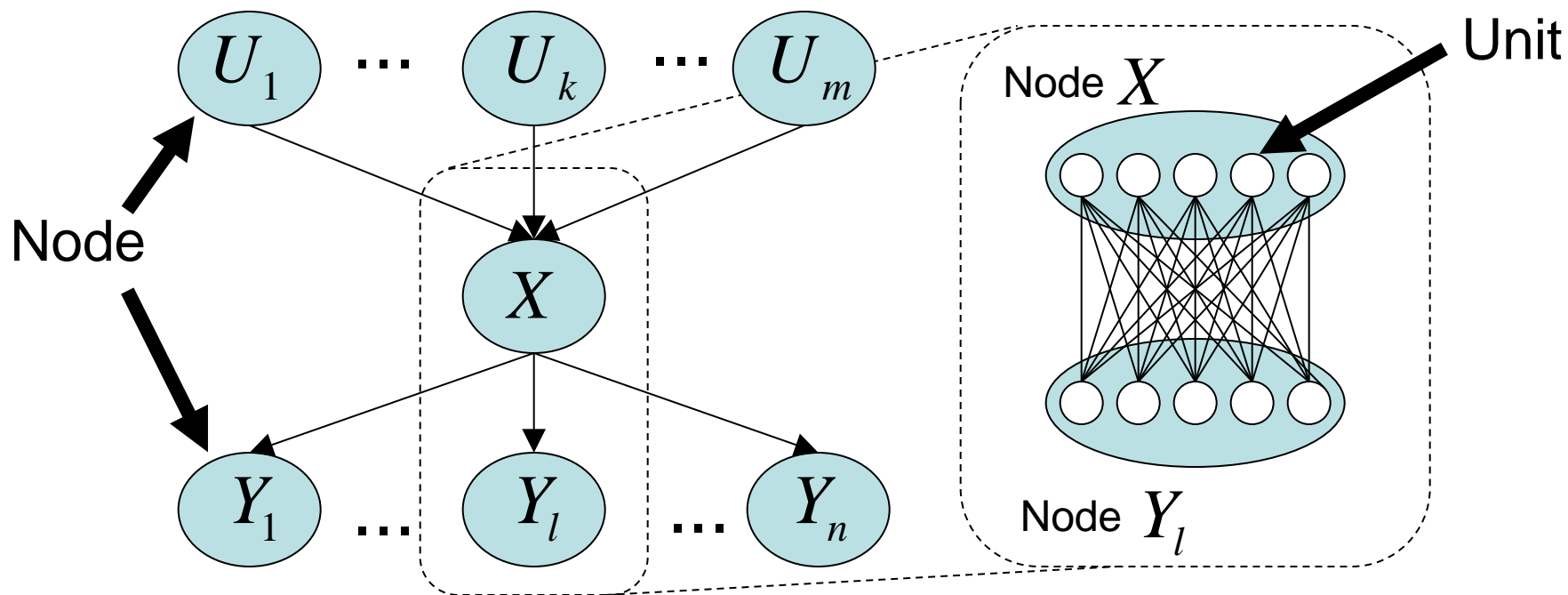
- 我々が提案する BESOM は、ノード数に対し線形時間で1入力を認識・学習。
  - 脳と同様、並列処理によりリアルタイム動作が可能。
  - 将来的には100万ノード規模。
- Feed forward NN より高機能。
  - Explaining away effect がある。
  - 時間的・空間的文脈の情報が自然に扱える。
  - 半教師あり学習、欠損データの学習が自然に扱える。
- 生成モデルを素直に表現可能。
  - 様々な事前知識を作り込みやすい。
    - 神経科学的知見からも作り込むべき事前知識のヒントが得られる。(ネットワークトポロジやハイパパラメタ)

# BESOM (Bidirectional SOM) とは？

- 大脳皮質ベイジアンネットモデルの1つ。
- ベイジアンネットと Deep Learning と SOM と ICA を組み合わせた機械学習アルゴリズム。
  - **ただし未完成な技術。**
- 目指す目標：
  - 計算量に関するスケーラビリティ
  - 精度のスケーラビリティ(次元の呪いの克服)
  - 機械学習理論的妥当性
  - 神経科学的妥当性

BESOM, 自己組織化マップ, ベイジアンネットワーク, 大脳皮質の構成要素の間の対応

BESOM	自己組織化マップ	ベイジアンネットワーク	大脳皮質
ノード	競合層	ノード (確率変数)	マクロコラム
ユニット	入力ベクトルの要素, 競合層のユニット	確率変数を取りうる値	ミニコラム
親ノード	入力層から見た競合層	親ノード (原因)	上位領野
子ノード	競合層から見た入力層	子ノード (結果)	下位領野
ユニットの出力	入力との類似度	事後確率	コラムの発火
結合の重み	参照ベクトルの要素	条件付確率	シナプスの重み



# BESOM Ver.3.0 の機構

- 制限付きの条件付確率表のモデル

$$P(x|u_1, \dots, u_m) = \frac{1}{m} \sum_{k=1}^m P(x|u_k)$$

- ノード数に対しほぼ線形時間で動作する  
認識・学習アルゴリズム [一杉、高橋 JSAI 2014]

- 正則化

- 近傍学習

- 勝率ペナルティ

- 側抑制ペナルティ

- エッジ選択

} 本発表



# 学習の目的

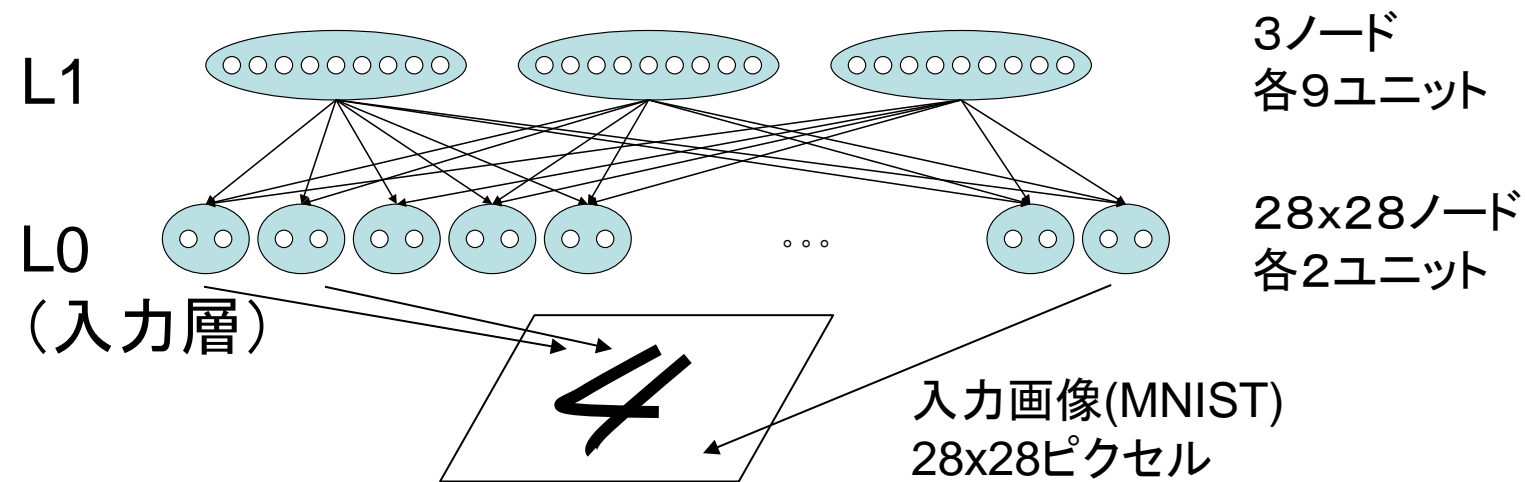
- 学習の目的:
  - 時刻  $t$  における入力を  $\mathbf{i}(t)$  と、入力が独立同分布 (i.i.d) から生成されると仮定して、パラメタ  $\theta$  を EM アルゴリズムで MAP 推定

$$\begin{aligned}\theta^* &= \arg \max_{\theta} \left[ \prod_{i=1}^t P(\mathbf{i}(i) | \theta) \right] P(\theta) \\ &= \arg \max_{\theta} \left[ \prod_{i=1}^t \sum_{\mathbf{h}} P(\mathbf{h}, \mathbf{i}(i) | \theta) \right] P(\theta)\end{aligned}$$

ベイジアンネットは確率伝搬アルゴリズムの収束結果を使って高速にEMアルゴリズムを実行可能。

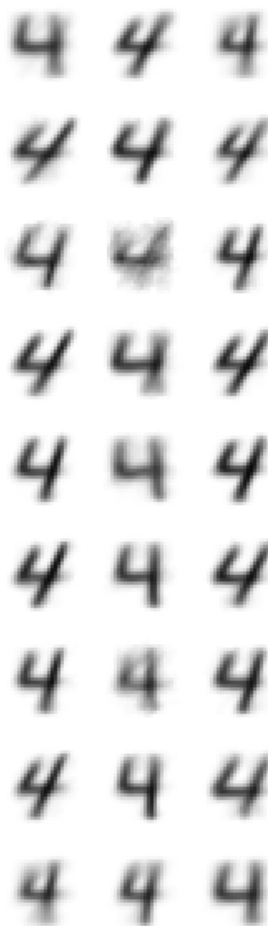
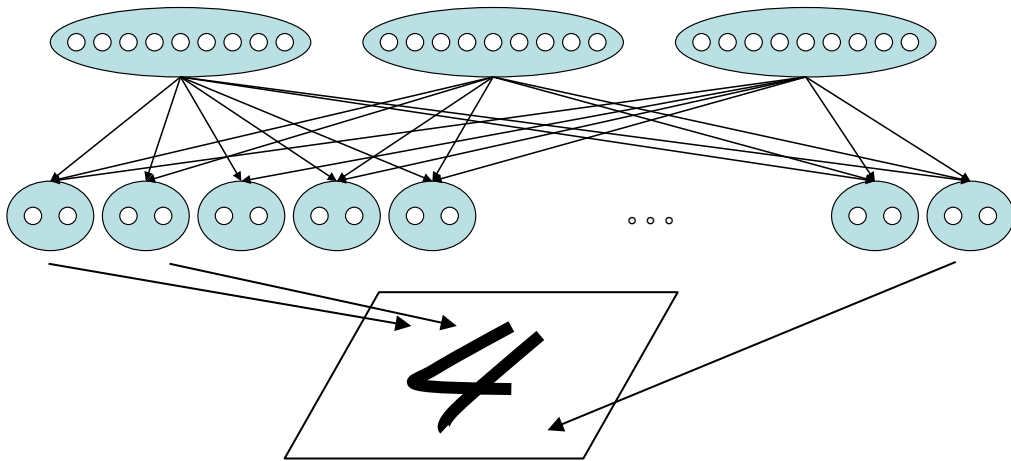
# 2層 BESOM

## 手書き数字の教師なし学習



1種類の文字のみ入力。

# 手書き数字 教師なし学習の 学習結果



× 使われない  
ユニットが多い

× 各ノードが  
似た応答になり  
がち

**局所解**

認識アルゴリズムは山登り法、MPE を真の値とみなしてパラメタをオンライン学習、近傍学習あり。

3ノード x 9ユニットの学習結果  
(各ユニットが応答する画像の平均)

# 勝率ペナルティ

- 各隠れ変数の周辺事前分布が一様分布から離れるほど、大きなペナルティを与える。

$$P^{WinRate}(\theta) = \prod_{X \in \mathbf{X}} e^{-C^{WinRate} D_{KL}(Q(X) || P(X; \theta))} \quad (8)$$

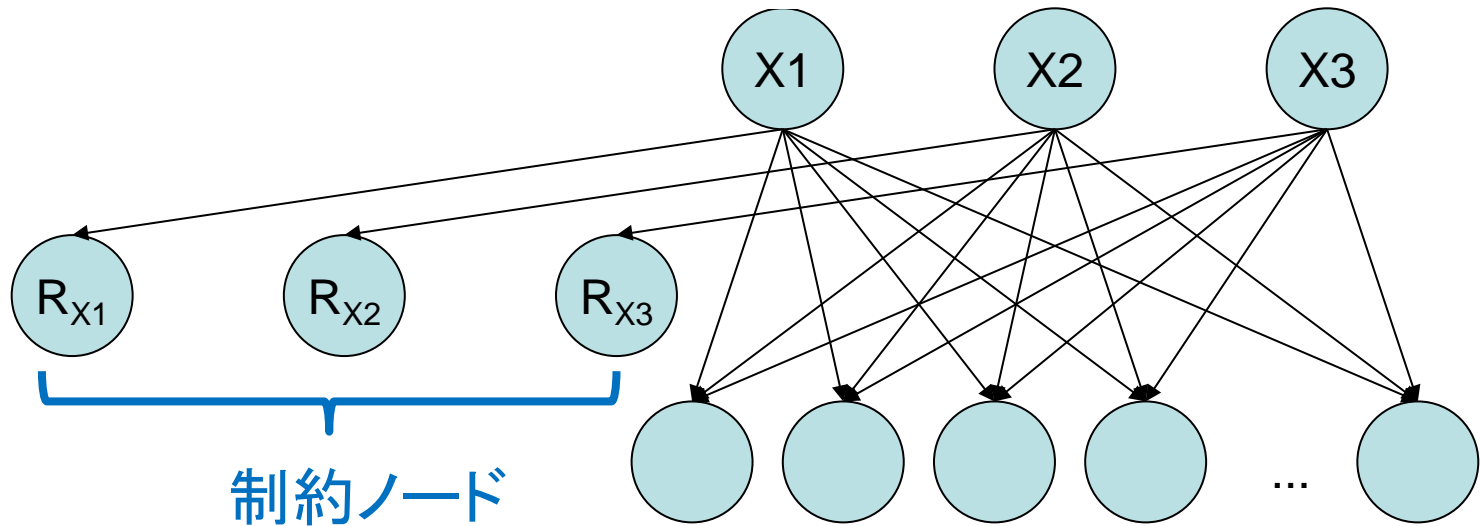
$$Q(X = x_i) = 1/s \quad (9)$$

**問題点:このような複雑な事前分布を持つとき、効率的なパラメタ学習の方法は自明ではない。**

# パラメタ学習の方法

- 勝率ペナルティを持ったベイジアンネットは、すべての隠れ変数の下に**制約ノード**を持つベイジアンネットと(近似的に)等価。この上でEM アルゴリズムを実行すればよい。

$$P(R_X = 1 | X = x; \theta) = e^{-(1/t)C^{WinRate} R(x; \theta)}, \quad R(x; \theta) = s \frac{Q(x)}{P(x; \theta)} \log \frac{Q(x)}{P(x; \theta)}$$



# 側抑制ペナルティ

- 2つのノードの間の相互情報量が大きいほど大きなペナルティを与える。
- 独立になってほしいノードのペアの集合を  $L$  として、事前分布を下記のように定義。

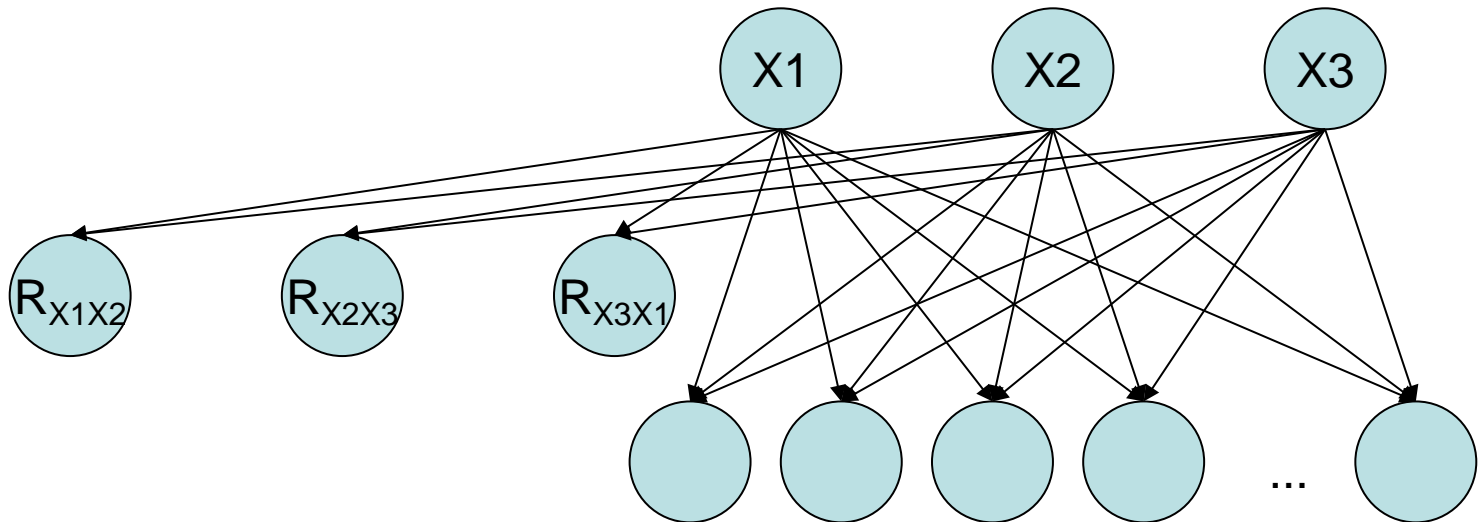
$$P^{Lateral}(\theta) = \prod_{(U,V) \in L} e^{-C^{Lateral} I(U,V;\theta)} \quad (17)$$

$$I(U, V; \theta) = \sum_u \sum_v P(u, v; \theta) \log \frac{P(u, v; \theta)}{P(u; \theta)P(v; \theta)} \quad (18)$$

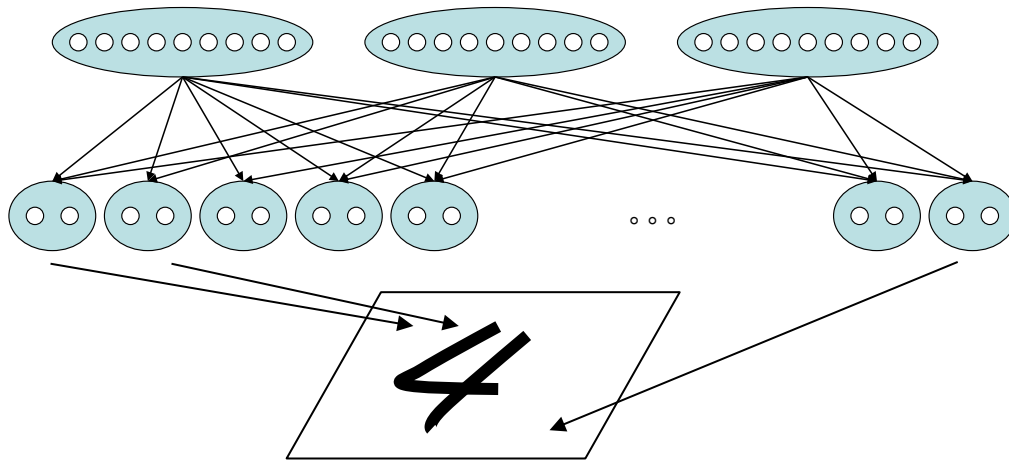
# パラメタ学習の方法

- **制約ノード**を2ノード間の共通子ノードとして持つベイジアンネットと等価。この上で EM。

$$\begin{aligned} P(R_{UV} = 1 | u, v; \theta) &= e^{-(1/t)C^{Lateral} R(u,v;\theta)} \\ &= s^2 \frac{P(u,v)}{P(u)P(v)} \log \frac{P(u,v)}{P(u)P(v)} \\ &= s^2 (P(u|v)/P(u)) \log P(u|v)/P(u) \end{aligned}$$



# 手書き数字 教師なし学習 正則化ありの場合の学習結果



傾き、横幅、横棒の位置

BESOM は非線形ICAを用いた特徴抽出器として使える。

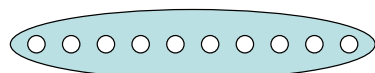


# 4層 BESOM

## 手書き数字 教師あり学習

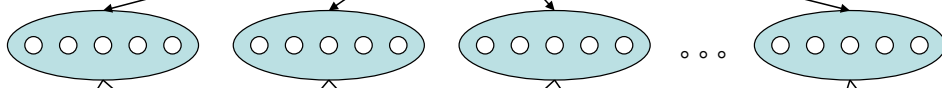
訓練時はL3に観測値を与える。  
テスト時はL3の事後確率最大の値を認識結果とする。

L3(出力層)



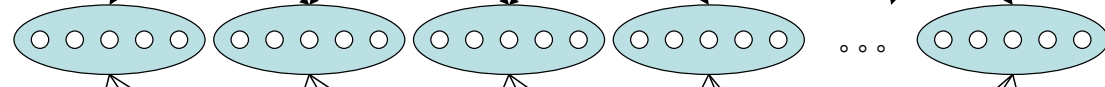
教師ノード10ユニット

L2



3x3ノード  
各30ユニット

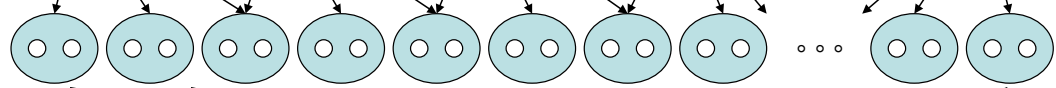
L1



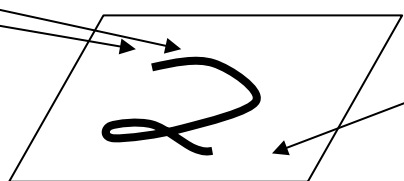
5x5ノード  
各20ユニット

L0

(入力層)



28x28ノード  
各2ユニット



入力画像(MNIST)  
28x28ピクセル

pre-training なし、近傍学習なし、認識アルゴリズムはOOBP、MPE を真の値とみなしてパラメタをオンライン学習

# 実験結果

	勝率ペナルティあり	勝率ペナルティなし
側抑制ペナルティあり	<b>84.7%</b>	84.0%
側抑制ペナルティなし	79.9%	61.3%

いずれも正則化した方が認識率が高い。

ただし MNIST の認識率としてはまだよくない。

大きな理由: 条件付確率表モデルの表現力の問題。

→ 今後の課題

# まとめ

- 神経科学的知見は、大脳皮質は Deep Learning と同じ構造のベイジアンネットであると示唆。
- ベイジアンネットを用いた Deep Learning は工学的にも有望。
- 勝率ペナルティ、側抑制ペナルティという2つの正則化によって過適合・局所解の問題を緩和。
- ただし、現在の条件付確率表のモデルでは性能は不十分。