

神経回路網特論

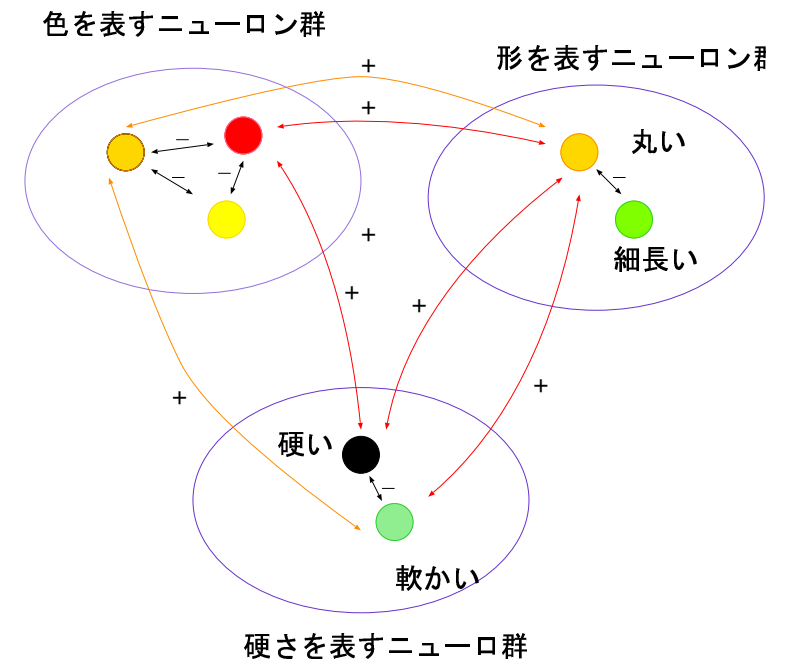
第7回 連想記憶モデル その(2)

伊達 章

2006年6月1日

講義のテーマ

- 全体を通してのテーマ：
情報表現と計算
- 前2回のテーマ：
短期記憶の表現：ニューロンの反応選択性
 - ・ 長期記憶の短期表現：スパース表現
- 連想記憶モデル
 - ・ 想起のダイナミックス



講義のテーマ

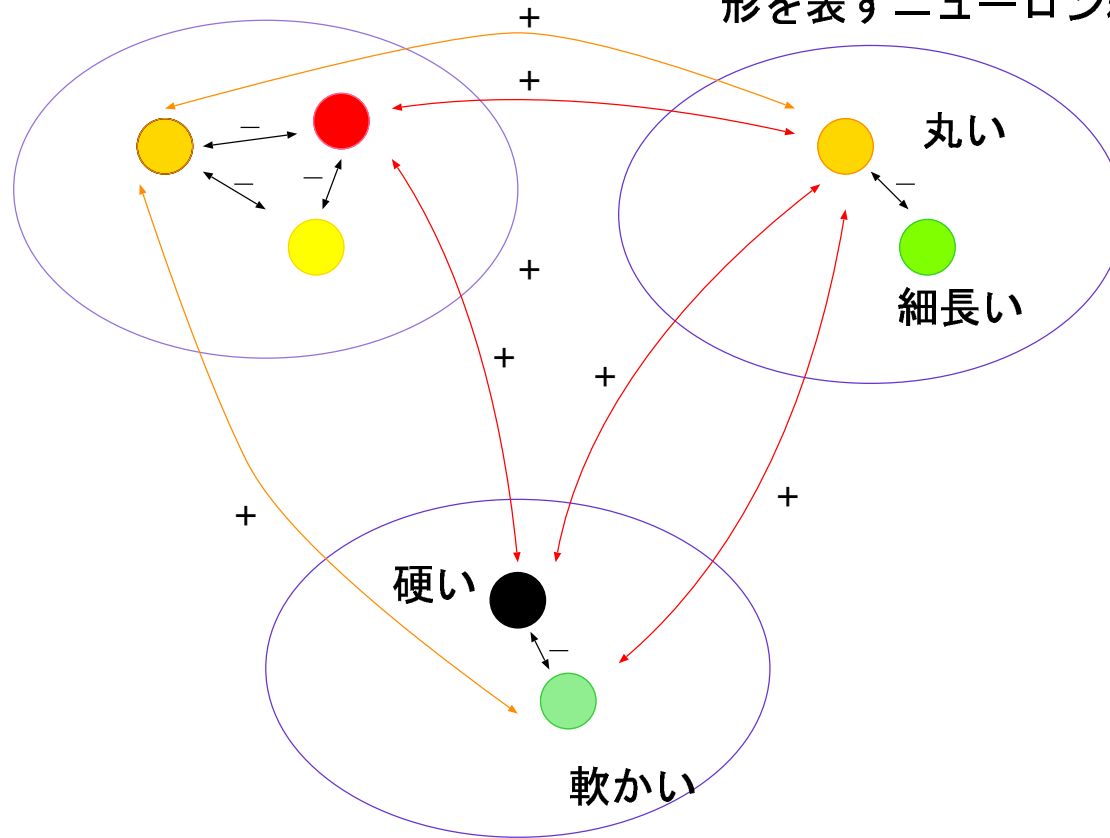
- 全体を通してのテーマ
 - ・ 情報表現と計算
- 今回のテーマ： 記憶の数理モデル
 - ・ 夢の役割： にせ記憶 逆学習によりなくす
 - ・ いくつかの複数の事項を同時に短期記憶するモデル：振動子モデル

概要

- 復習： 連想記憶モデルとは
 - ・ 多重分散表現
 - ・ 記憶容量
 - ・ 想起のダイナミックス

色を表すニューロン群

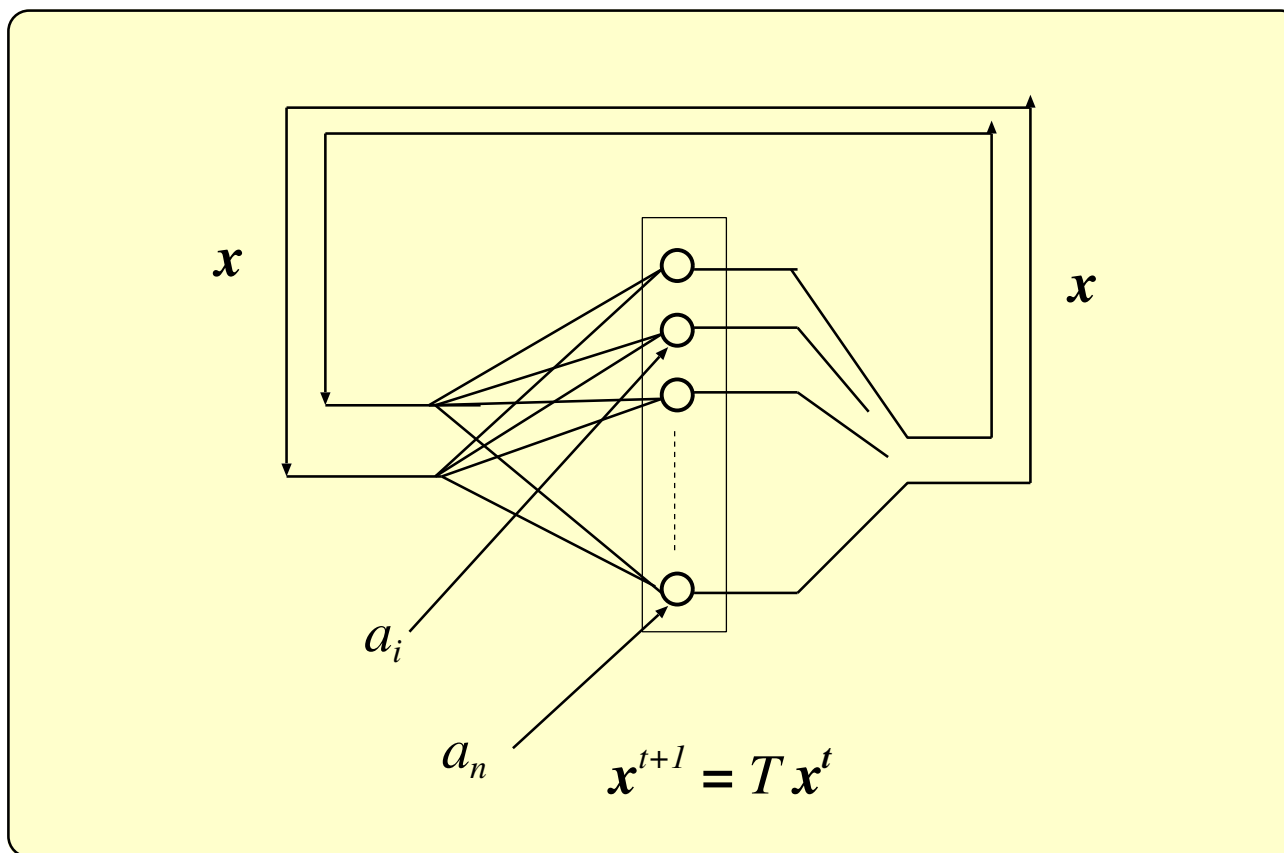
形を表すニューロン群



硬さを表すニューロン群

リンゴ, バナナ, ミカンをおぼえたニューロン群

相互結合の神経回路モデル



連想記憶モデル：自己想起型

- $\hat{x}^1 \rightarrow x^1, \hat{x}^2 \rightarrow x^2, \dots, \hat{x}^k \rightarrow x^k$ • $x^\alpha \cdot x^\beta = 0, \alpha \neq \beta, x_i \in \{-1, 1\}$

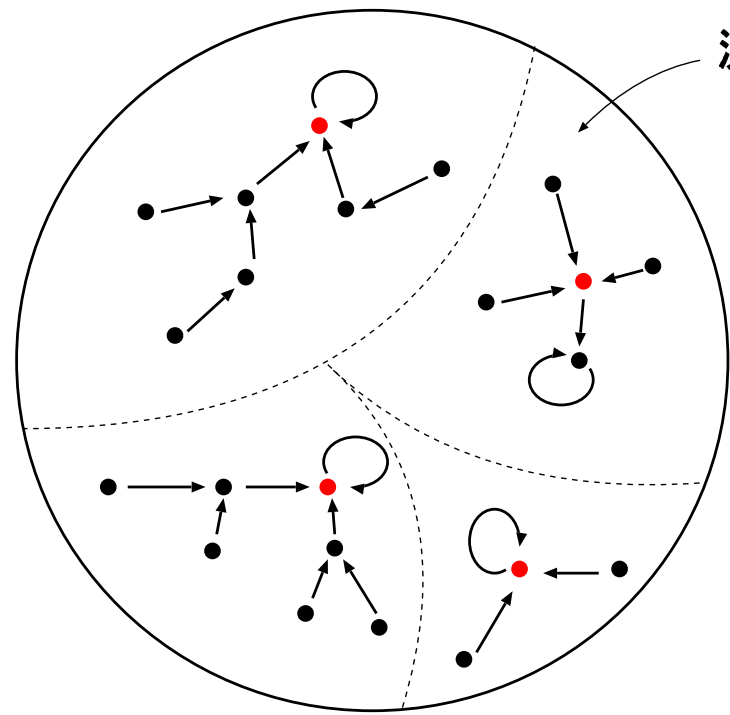
$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \boldsymbol{x}\boldsymbol{x}^\top = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} [x_1 x_2 \dots x_n] = \begin{bmatrix} x_1 x_1 & x_1 x_2 & \dots & x_1 x_n \\ x_2 x_1 & x_2 x_2 & \dots & x_2 x_n \\ & & \dots & \\ x_n x_1 & x_n x_2 & \dots & x_n x_n \end{bmatrix}$$

$$s_{ji} = \frac{1}{n} \sum_{\alpha=1}^k x_i^\alpha x_j^\alpha, \quad S = \frac{1}{n} \sum_{\alpha=1}^k \boldsymbol{x}^\alpha \boldsymbol{x}^{\alpha\top} \quad (\text{ヘブ学習, 相関学習})$$

$$\text{活動のダイナミクス: } \boldsymbol{x}' = \text{sgn}(S\boldsymbol{x}^3)$$

連想記憶モデル

$$x \Rightarrow T_w x$$



- 記憶パターン

$$x^\mu = T_w x^\mu$$

$$\mu = 1, 2, \dots, m$$

m 記憶パターン数

結合係数

$$w_{ij} = k \sum x_i^\mu x_j^\mu$$

連想記憶モデル：自己想起型

- $\mathbf{x}^\alpha \cdot \mathbf{x}^\beta = 0, \alpha \neq \beta$

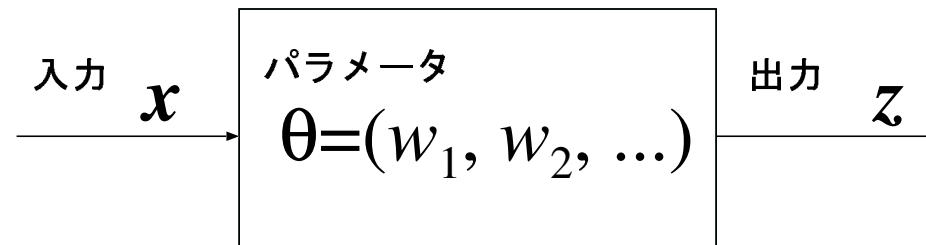
$$\text{結合係数行列： } S = \frac{1}{n} \sum_{\alpha=1}^k \mathbf{x}^\alpha \mathbf{x}^{\alpha\top}$$

$$\text{活動のダイナミクス： } \mathbf{x}' = \text{sgn}(S\mathbf{x}^3)$$

$$S\mathbf{x}^3 = \frac{1}{n} \sum_{\alpha=1}^k \mathbf{x}^\alpha \mathbf{x}^{\alpha\top} \mathbf{x}^3 = \frac{1}{n} \sum_{\alpha=1}^k \mathbf{x}^\alpha (\mathbf{x}^\alpha \cdot \mathbf{x}^3) = \frac{1}{n} \mathbf{x}^3 (\mathbf{x}^3 \cdot \mathbf{x}^3) = c\mathbf{x}^3$$

二つのダイナミクス

連想記憶モデルの見方： $z = x'$



1. ニューロン活動 x のダイナミクス (速い)
短期記憶 (長期記憶の短期的表現) に対応
2. 結合係数 w のダイナミクス (遅い いったん決めると固定)
長期記憶に対応

多重分散記憶がうまくいく仕掛け

- x^1, \dots, x^k が互いに直交

$$Sx^3 = \frac{1}{n} \sum_{\alpha=1}^k x^\alpha x^{\alpha\top} x^3 = \frac{1}{n} \sum_{\alpha=1}^k x^\alpha (x^\alpha \cdot x^3) = \frac{1}{n} x^3 (x^3 \cdot x^3) = cx^3$$

- 出力関数 $f(u)$ の強い非線型性

$$x' = \text{sgn}(Sx^3) = \text{sgn}(cx^3) = x^3$$

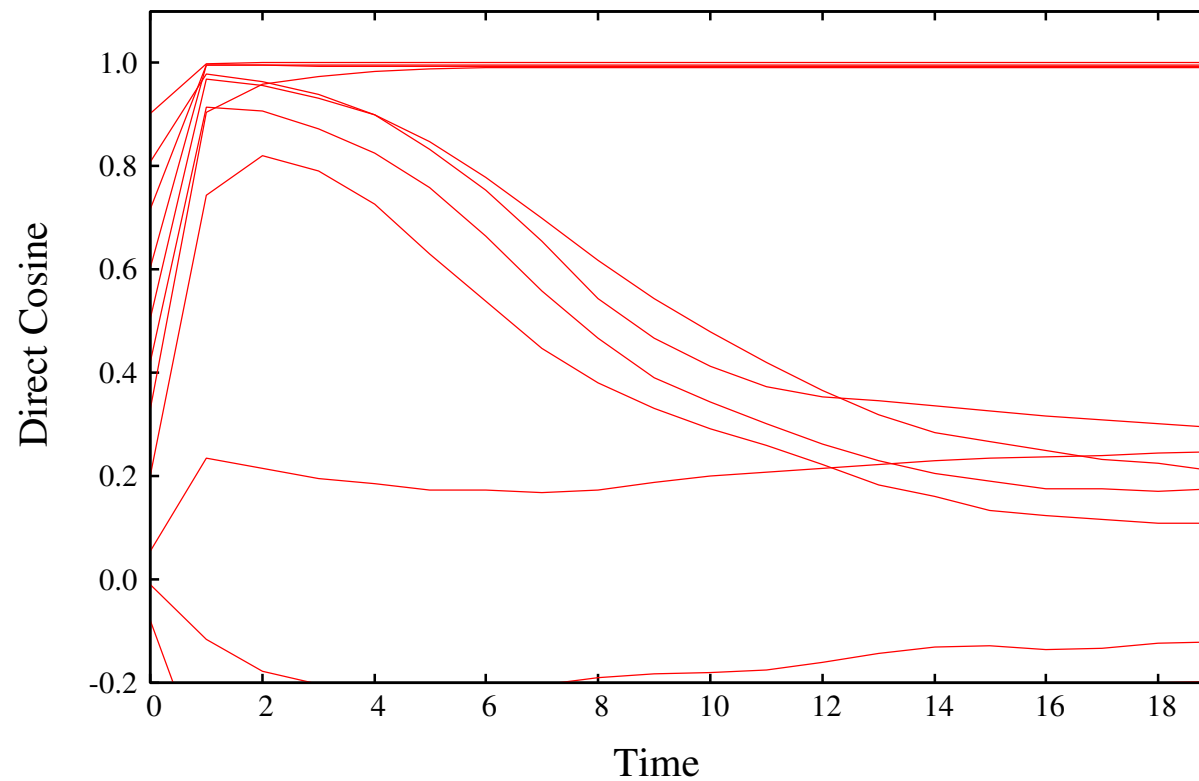
連想記憶モデルの問題

- 多重 (k 重) 表現 相互干渉 記憶容量 (k_{\max})

$$\text{結合係数行列} : S = \frac{1}{n} \sum_{\alpha=1}^k \mathbf{x}^{\alpha} \mathbf{x}^{\alpha T}$$

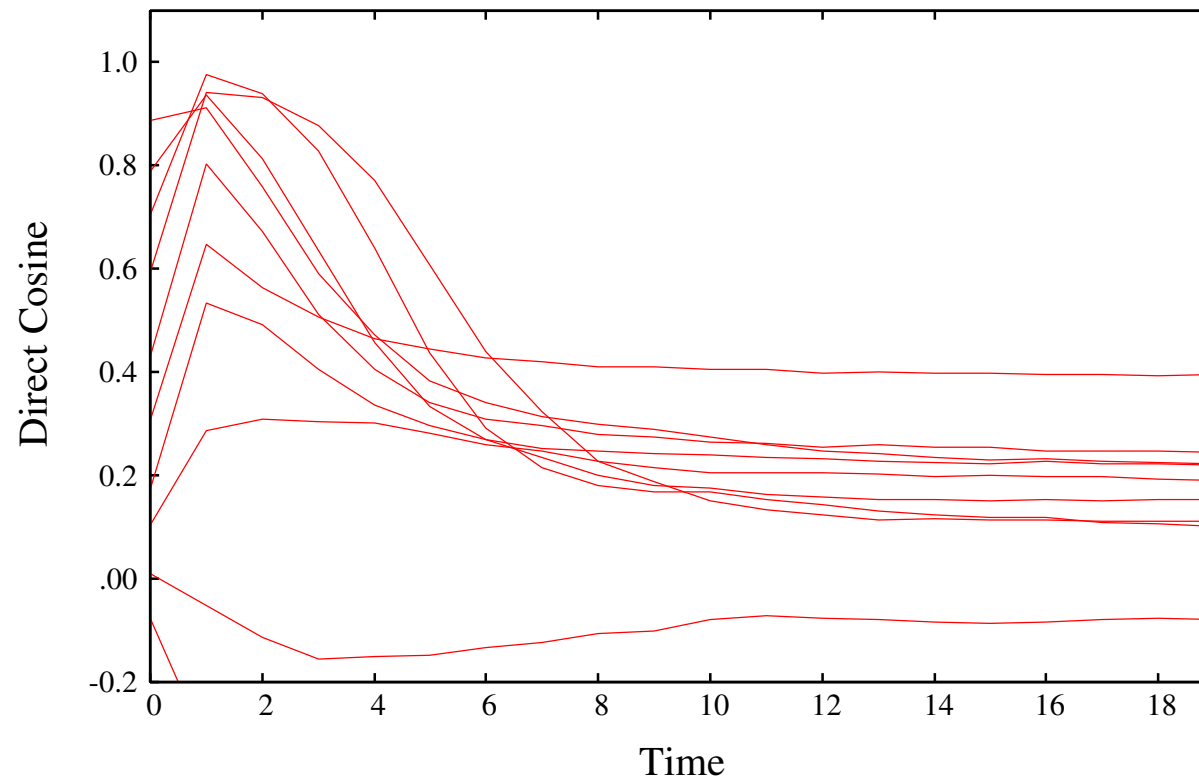
- 想起過程のダイナミクス
- 分散表現 どう符合化するか (簡単な問題ではない．来週)

想起過程のダイナミクス



想起に失敗する場合でも一度は記憶パターンに近づく

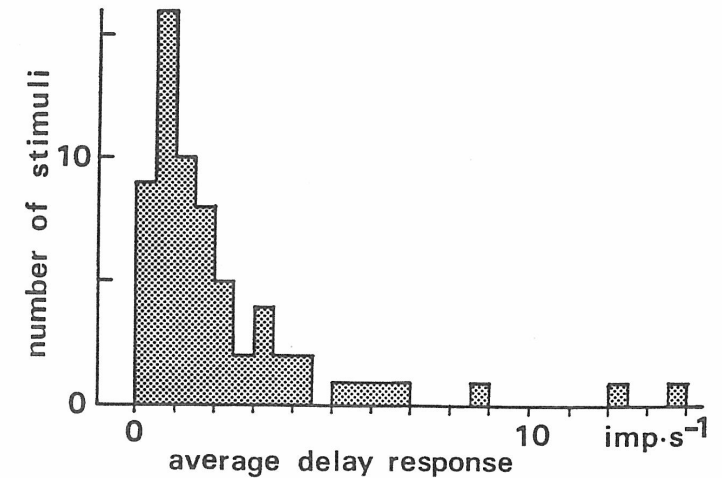
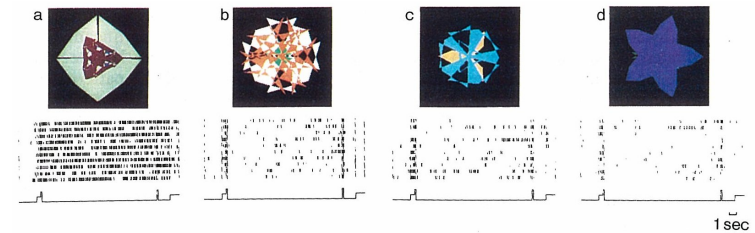
想起過程のダイナミクス



記憶パターンを埋め込みすぎた例

サルを使った実験との関係

- 複数の事項を短期記憶する過程
- 長期記憶の短期表現
 - ・ 連想記憶モデル
分散パターンを記憶
 - ・ サルの実験
一つのニューロンの興奮パターン

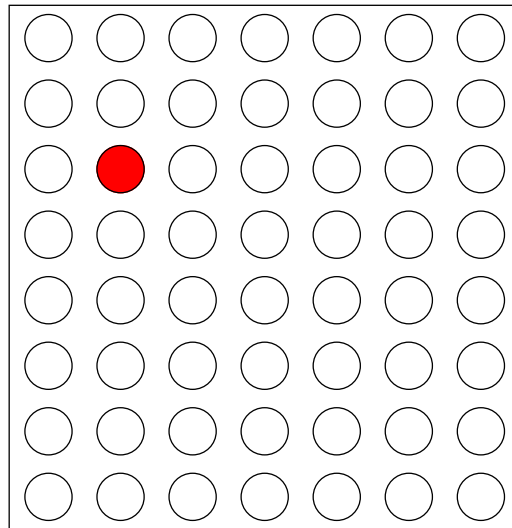


情報表現の仕方

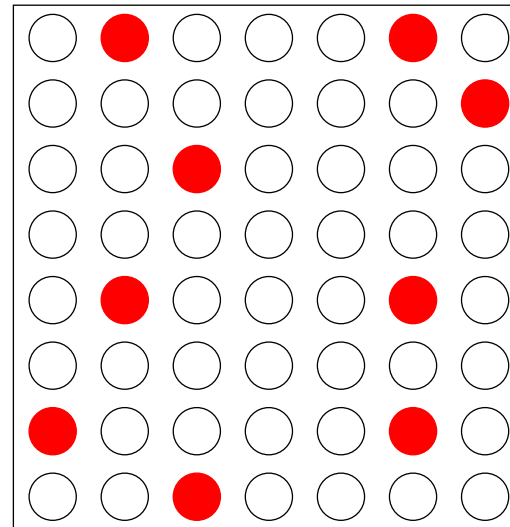


入力

局在表現



分散表現

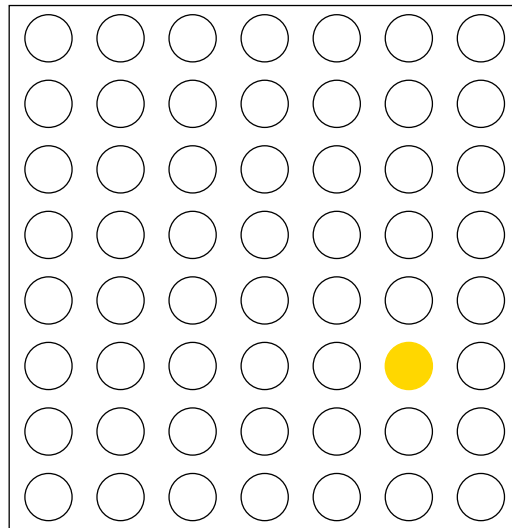


情報表現の仕方

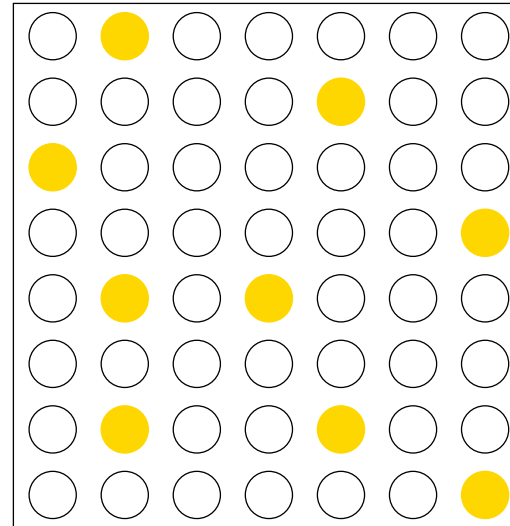


入力

局在表現



分散表現

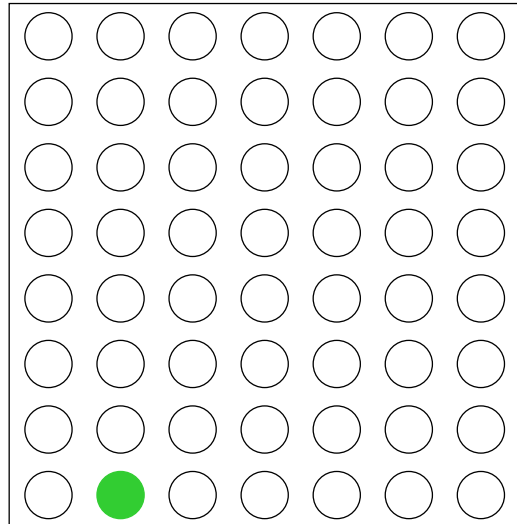


情報表現の仕方

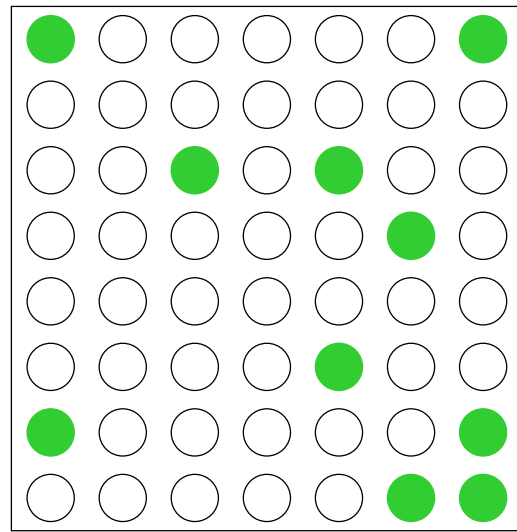


入力

局在表現



分散表現



論文紹介

- 夢の効用 : Hopfield et al. (1983)
- 複数の事項を短期記憶する過程 : Wang et al (1990)

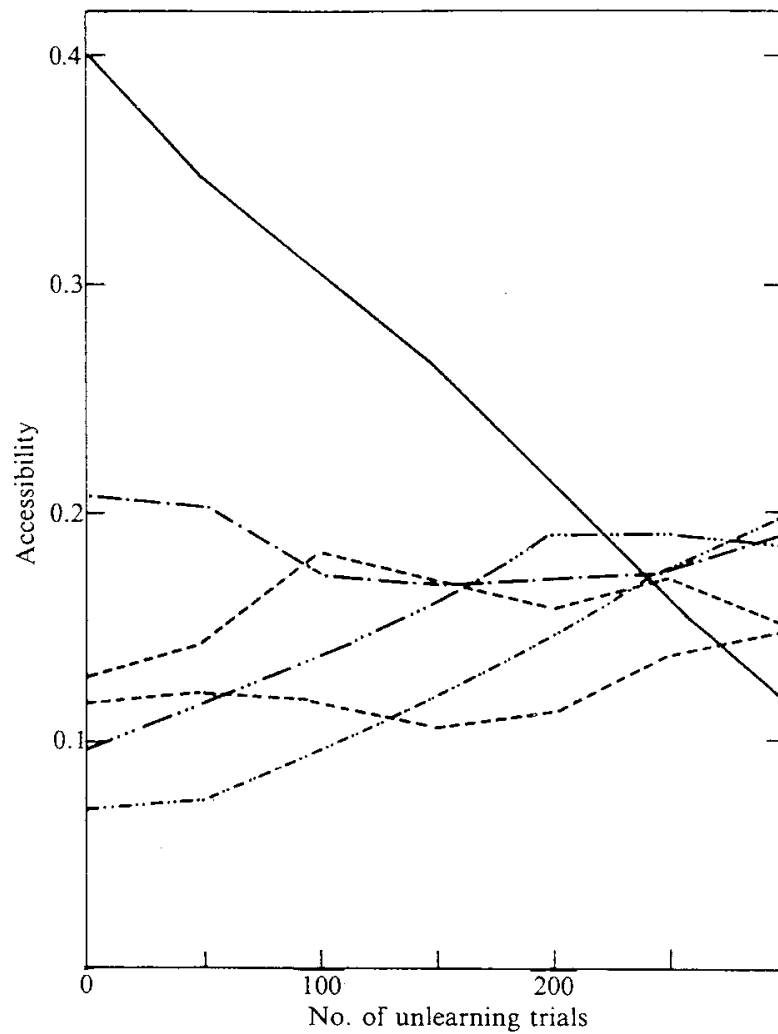


Fig. 1 The fraction of random starting states which leads to particle memories (accessibility). The five dashed lines are the five nominal memories. The solid line is the total accessibility of all spurious memories. In these trials ϵ was set at 0.01.

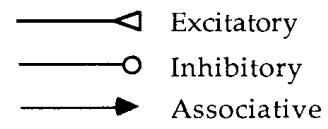
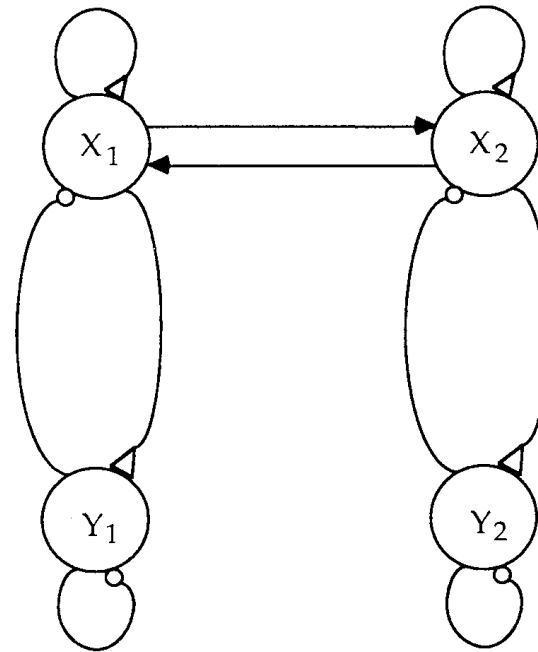
Memory 1
Memory 2
Memory 3
Spurious memory

+	+	+	+	-	-	-	-		+	+	-	+	-	+	-	-
+	+	+	+	-	-	-	-		-	-	+	-	+	-	+	+
+	+	-	-	+	+	-	-		+	-	-	+	+	-	-	+
+	+	+	+	-	-	-	-		+	-	-	+	+	-	-	+

にせ記憶

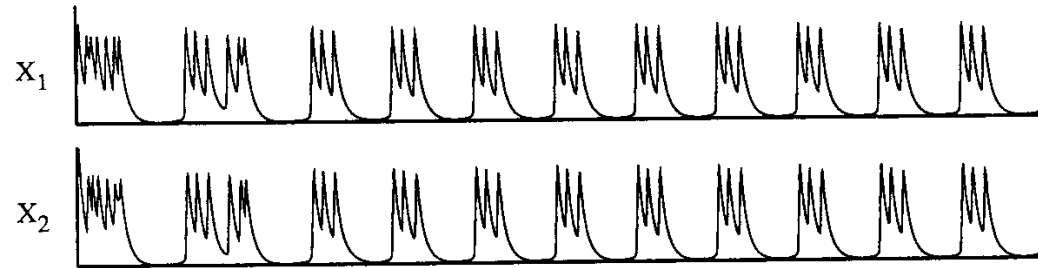
論文紹介

- 夢の効用 : Hopfield et al. (1983)
- 複数の事項を短期記憶する過程 : Wang et al (1990)

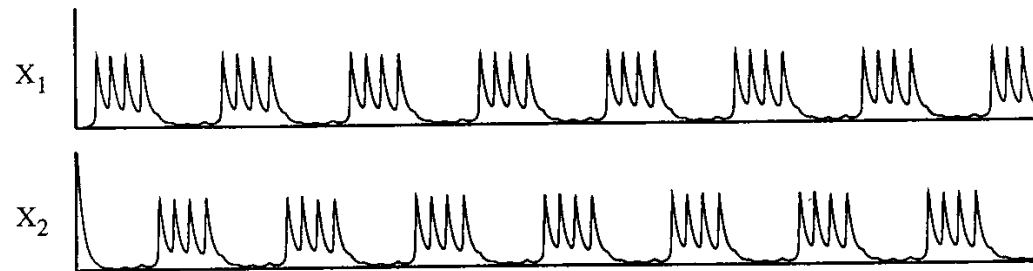


結合振動子

a



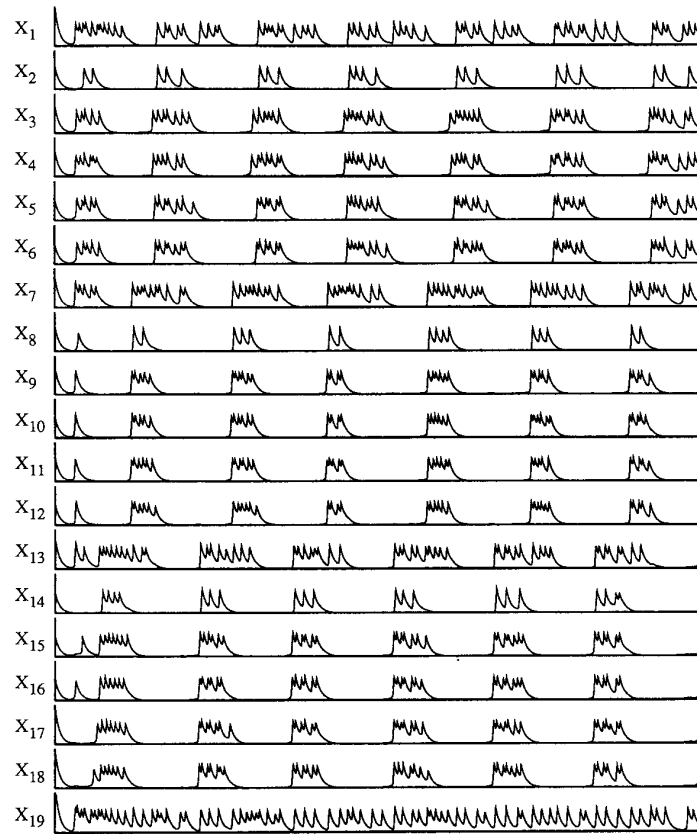
b



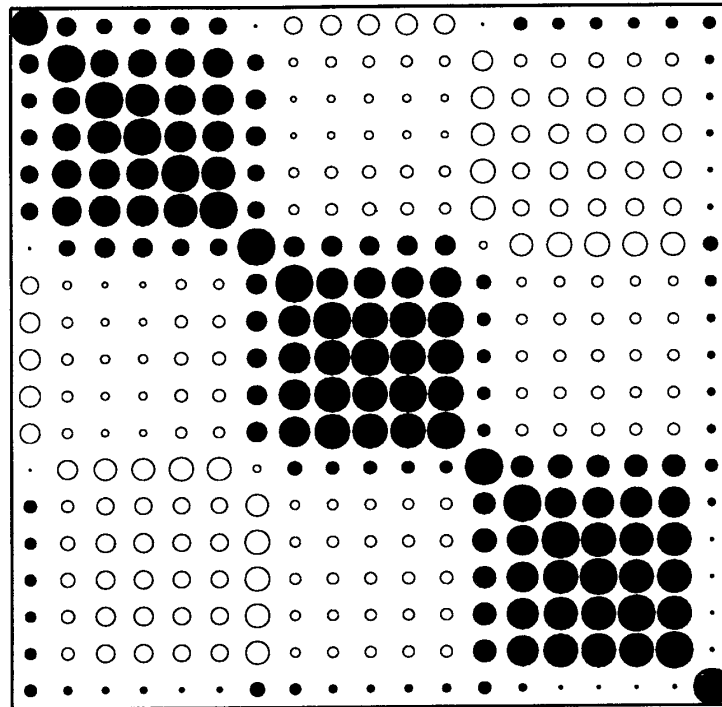
振動子による同期・非同期発火パターンの例

$$\begin{aligned}
\xi^1 &= (1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, \dots, 0) \\
\xi^2 &= (0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, \dots, 0) \\
\xi^3 &= (1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 0, \dots, 0)
\end{aligned}
\tag{3.2}$$

記憶させたパターン（8個のうちの3個）



同時に3つのパターンを想起している状態 (19/50個の素子を表示)



各素子間の発火の相関 (19/50 個の素子を表示 . が正の相関)

記憶の問題

1) 長期記憶のデータ構造

- ・いろいろな概念の表現
- ・概念間の複雑な関係の表現

2) エピソード記憶

- ・個々の事項を時空間的に組み合わせ一つのストーリーとする
- ・脳の中の仮想的な時空間

3) 整理の仕方：短期記憶として経験する事項をどう整理して長期記憶にしまうか

- ・不要な部分を捨て、必要な部分だけを抽出
- ・今までの記憶と適合する形で長期記憶を改変

これらについて気分が感じ取れるモデルは未だに無い