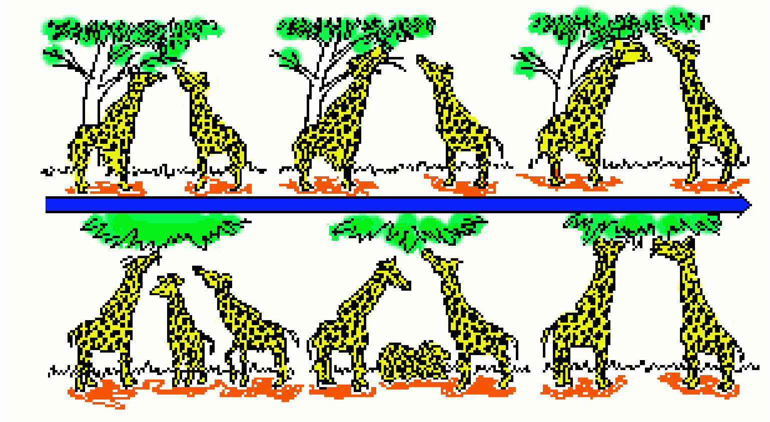
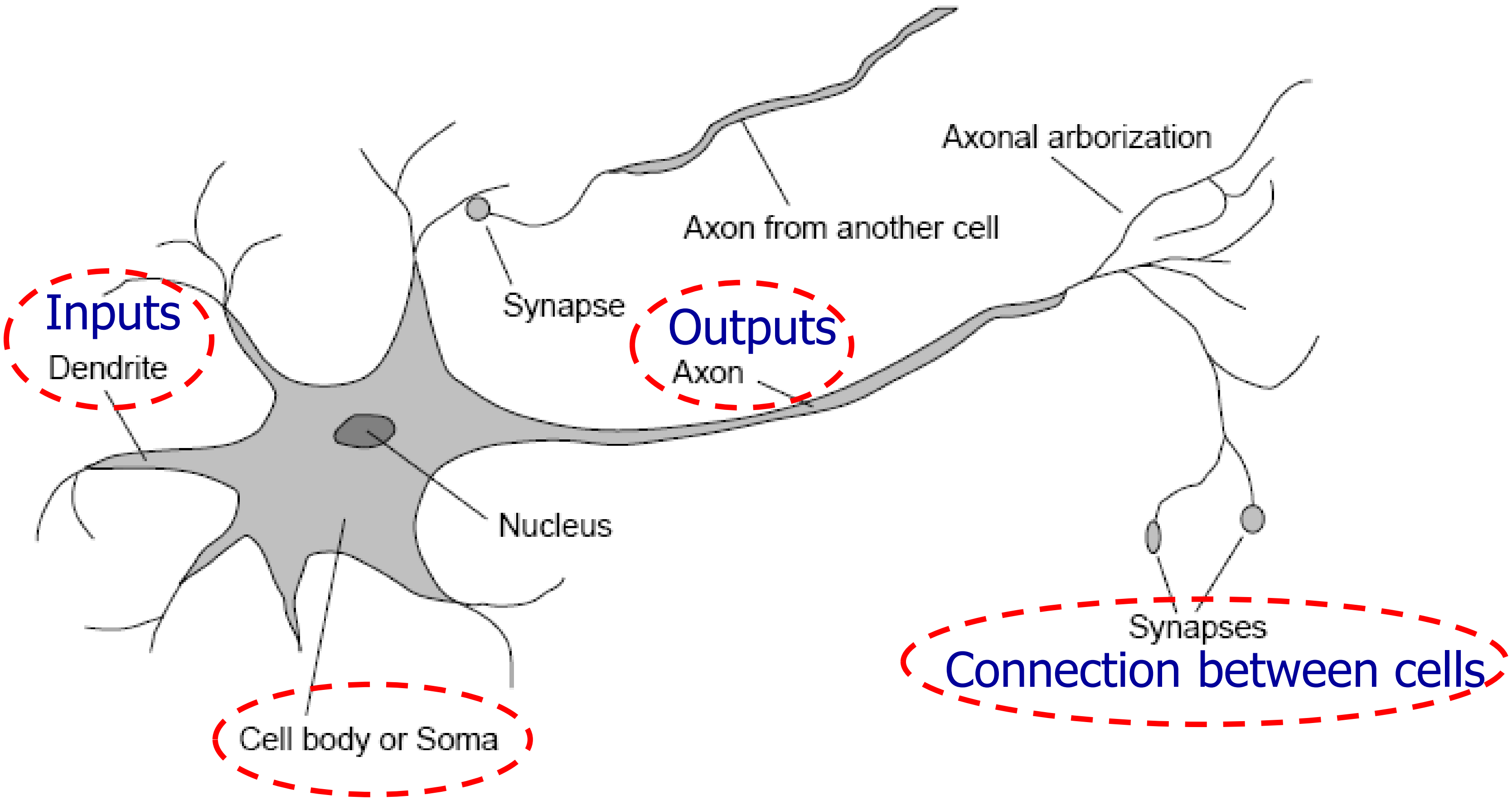
パーセプトロン

東京大学大学院
情報理工学系科学研究科
電気情報学専攻
伊庭斉志



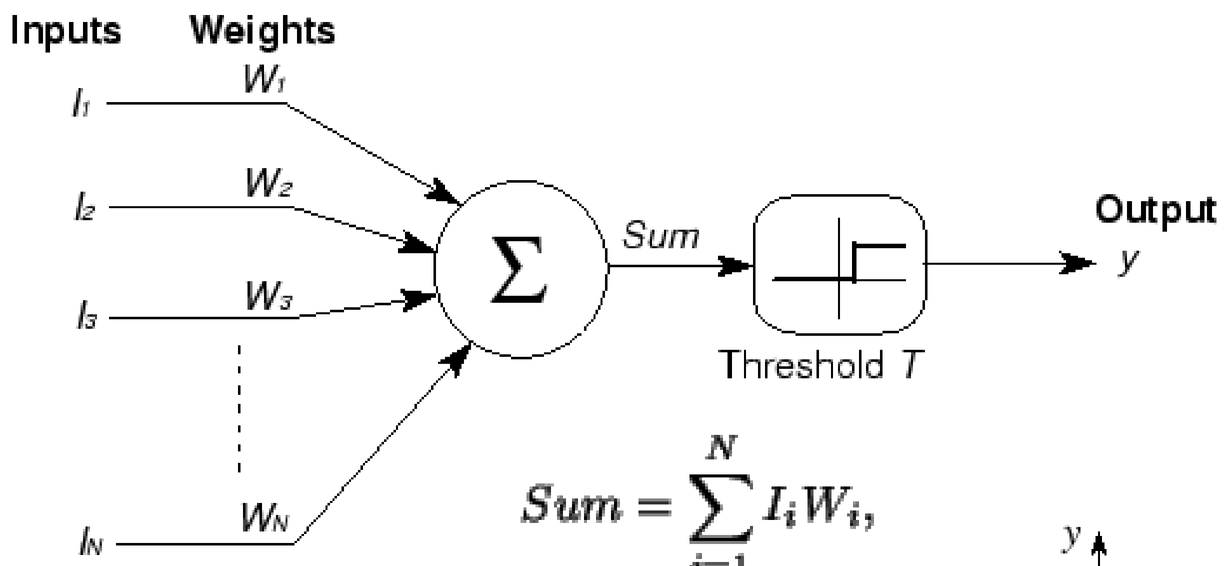
iba lab.

神経ニューロン

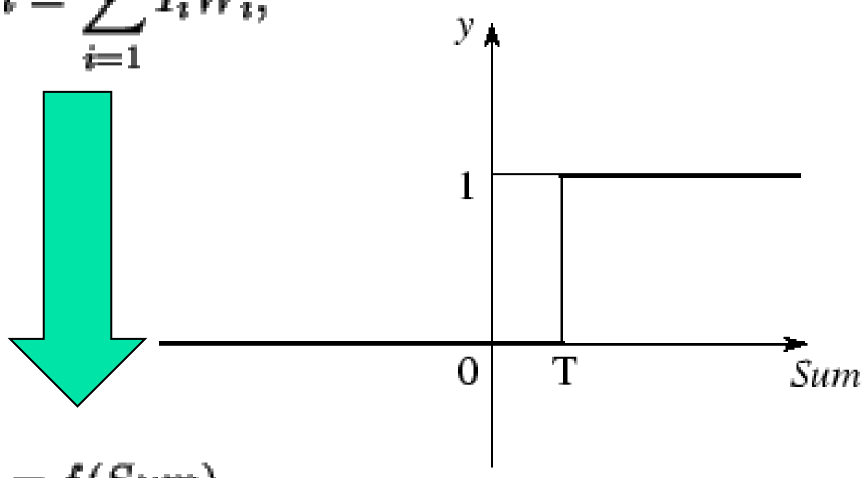


Axonal arborization
Axon from another cell
Synapse
Inputs
Dendrite
Outputs
Axon
Nucleus
Synapses
Connection between cells
Cell body or Soma

McCulloch-Pittsモデル



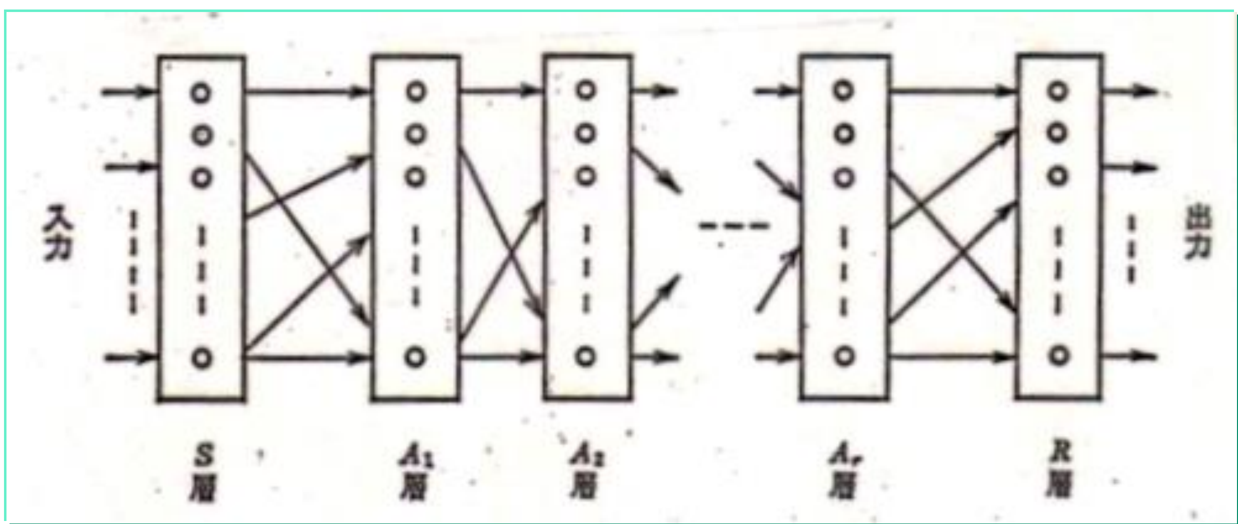
$$Sum = \sum_{i=1}^{N} I_i W_i,$$



$$y = f(Sum).$$

パーセプトロン

- Frank Rosenblatt（1961)
- 学習能力のある神経素子を構成要素とする多層の層状回路
- 脳の学習モデルとして注目される



単純パーセプトロン

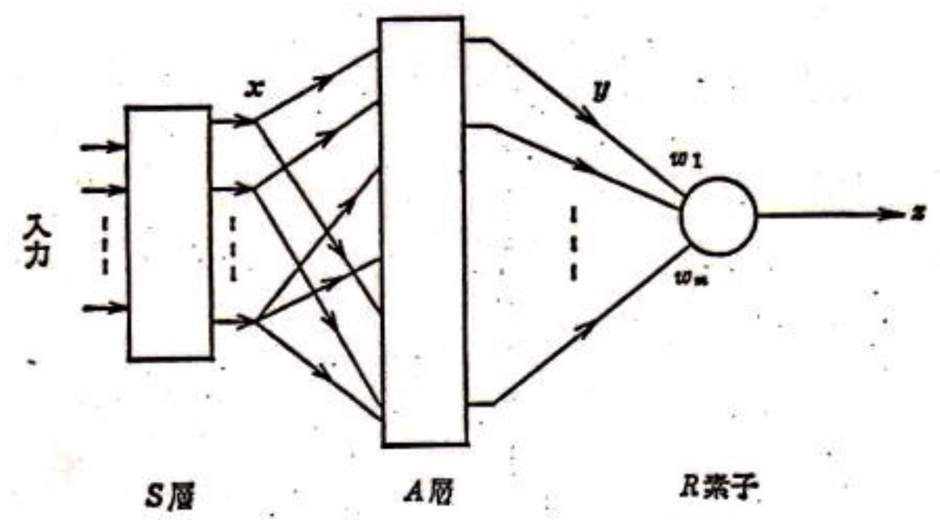
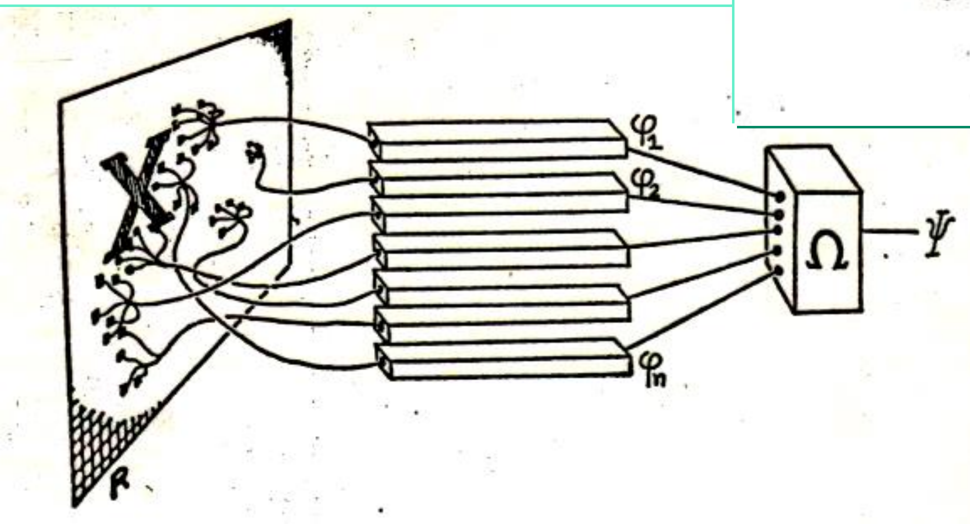
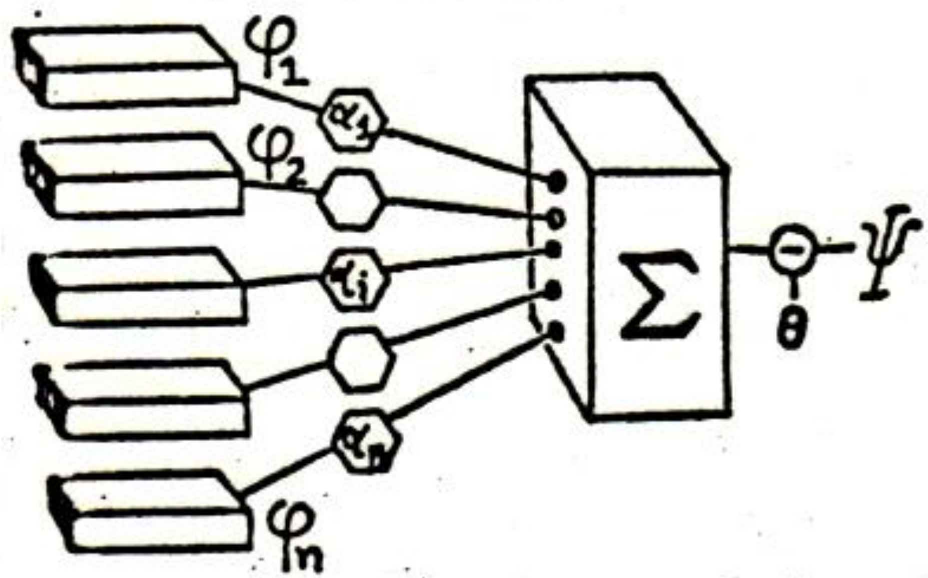


図 9.10 単純パーセプトロン



単純パーセプトロン



「 $\sum_{\varphi\in\Phi}\alpha_\varphi\varphi(X)>\theta$ のとき，かつそのときに限り $\psi(X)=1$ 」

$$\alpha_{\varphi_1}\varphi_1(X)+\cdots\cdots+\alpha_{\varphi_n}\varphi_n(X)>\theta$$

論理関数の集合を $\Phi=\{\varphi_1,\varphi_2,\cdots\cdots,\varphi_n\}$ とする．

単純パーセプトロンの初期の成果

- 収束性定理
  - 線形分離可能性を仮定
  - Nilsson(1965), Minsky & Papert(1969)
- 小脳のパーセプトロン説
  - Marr & Albus

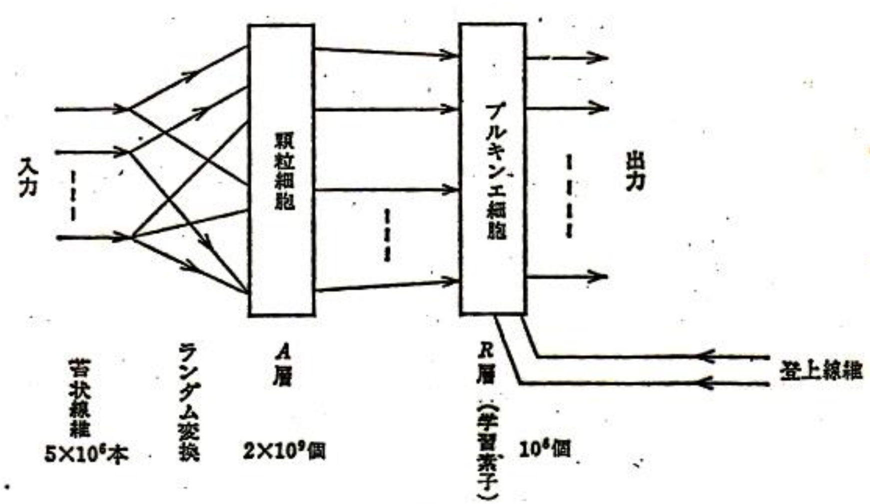


図 9.11 小脳パーセプトロン

線形分離可能性

- 線形モデルによる2値分類

$$y(\mathbf{x}) = f\left(\mathbf{w}^T \phi(\mathbf{x})\right)$$

$$f(a) = \begin{cases} +1, & a \geq 0 \\ -1, & a < 0 \end{cases}$$

$\mathbf{x}$ ： サンプル

$\phi(\mathbf{x})$ ： 特徴ベクトル

$\mathbf{w}$ ： 重みベクトル

バイアス要素　$\phi_0(\mathbf{x}) = 1$

重みベクトルと特徴ベクトルの内積をとって、それがゼロ以上であれば +1（正例と判定）、そうでなければ -1（負例と判定）を返す関数

パーセプトロン学習アルゴリズム

1. 重みベクトルを要素0で初期化
2. 学習データからランダムにサンプルを選択
3. 分類が間違っていたら以下の式で重みベクトルを更新
   - 正例の場合　$\mathbf{w} \leftarrow \mathbf{w} + \phi(\mathbf{x})$
   - 負例の場合　$\mathbf{w} \leftarrow \mathbf{w} - \phi(\mathbf{x})$
4. すべてのサンプルを正しく分類できるまで 2 に戻り繰り返す

学習例　OR の学習

- 学習データ

$$\phi(\mathbf{x}_1)=\begin{pmatrix}1\\0\\0\end{pmatrix}$$

$$t_1=-1$$

負例

$$\phi(\mathbf{x}_2)=\begin{pmatrix}1\\0\\1\end{pmatrix}$$

$$t_2=1$$

正例

$$\phi(\mathbf{x}_3)=\begin{pmatrix}1\\1\\0\end{pmatrix}$$

$$t_3=1$$

正例

$$\phi(\mathbf{x}_4)=\begin{pmatrix}1\\1\\1\end{pmatrix}$$

$$t_4=1$$

正例

第2座標と第3座標のORが第１座標である

分離平面

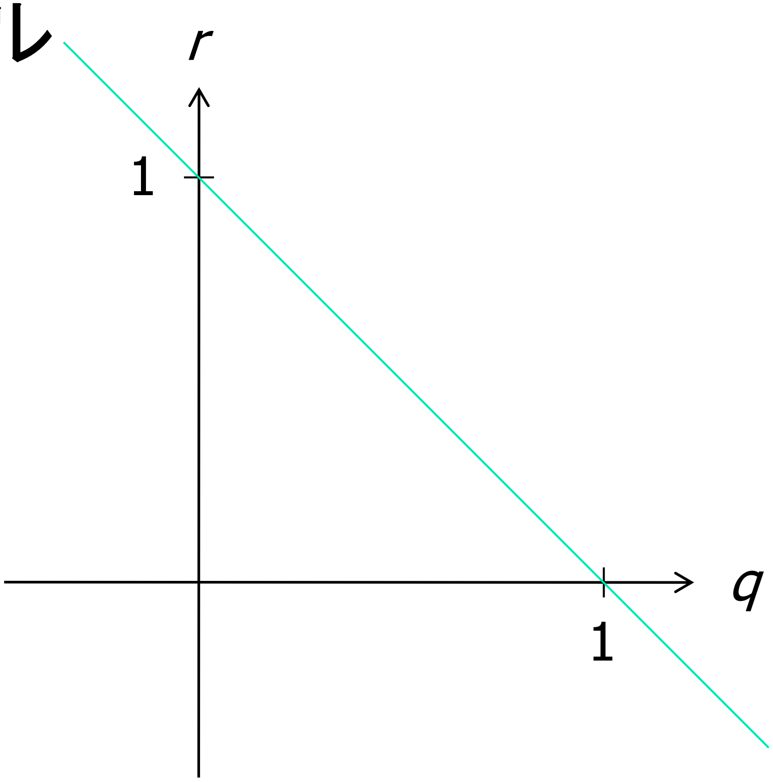
- 最終的な重みベクトル

$$\mathbf{w} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

分離平面 $\mathbf{w}^T \phi(\mathbf{x}) = 0$

入力（特徴ベクトルの2番目、3番目の要素）をそれぞれ *q*, *r* とすると

$$-1 + q + r = 0$$



なぜうまく学習できるのか？

- 正例の分類に失敗した場合

$$y(\mathbf{x}) = f\left(\mathbf{w}^T \phi(\mathbf{x})\right)$$

この部分の値が小さすぎた

$\mathbf{w} \leftarrow \mathbf{w} + \phi(\mathbf{x})$ として重みベクトルを更新すると

$$\left(\mathbf{w} + \phi(\mathbf{x})\right)^T \phi(\mathbf{x}) = \underline{\mathbf{w}^T \phi(\mathbf{x})} + \underline{|\phi(\mathbf{x})|^2}$$

もともとの値　必ず正

少なくとも同じサンプルに関しては間違いにくくなっている

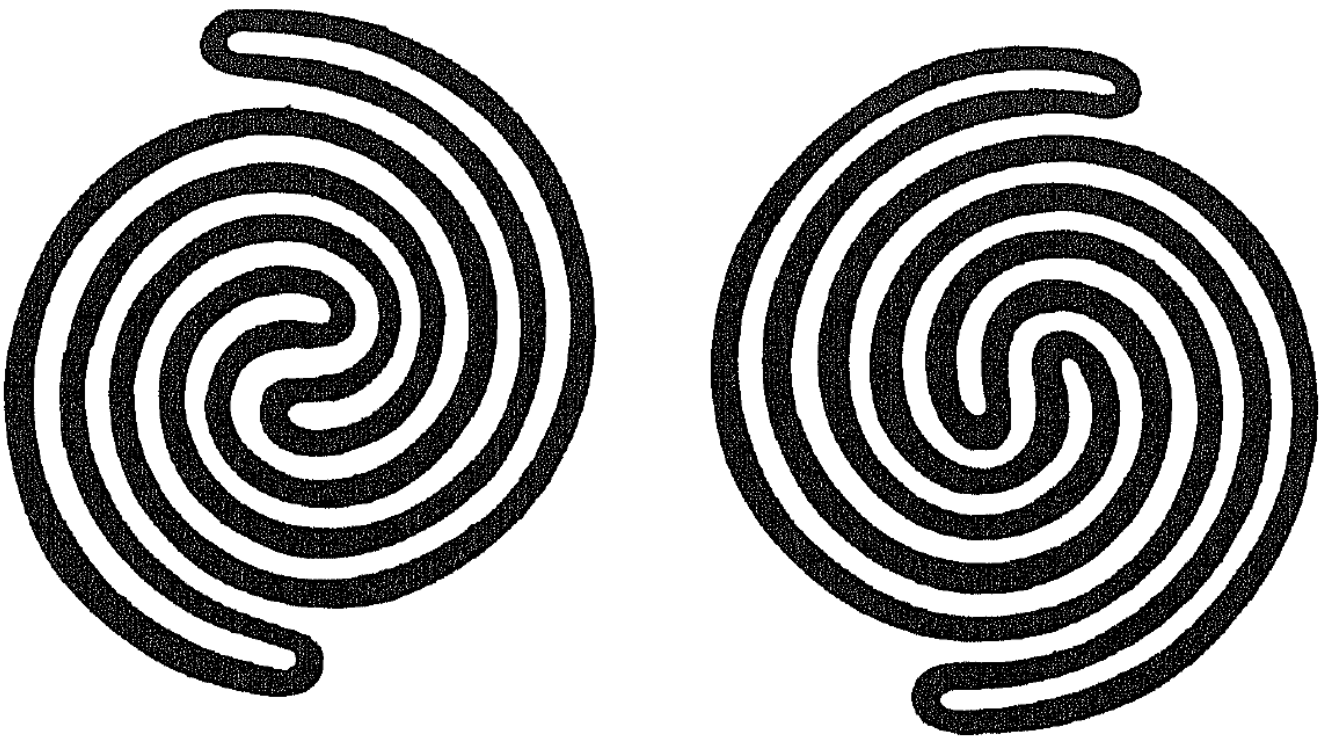
パーセプトロンの収束定理

- 正例と負例が線形分離可能であるならば、前述のアルゴリズムにより、それらを分離する超平面を有限のステップで見つけられることが保証されている。

- データが線形分離可能でないならば、パーセプトロン学習は収束しない
- 線形分離可能である場合でも、ステップ数が非常に大きくなることがある

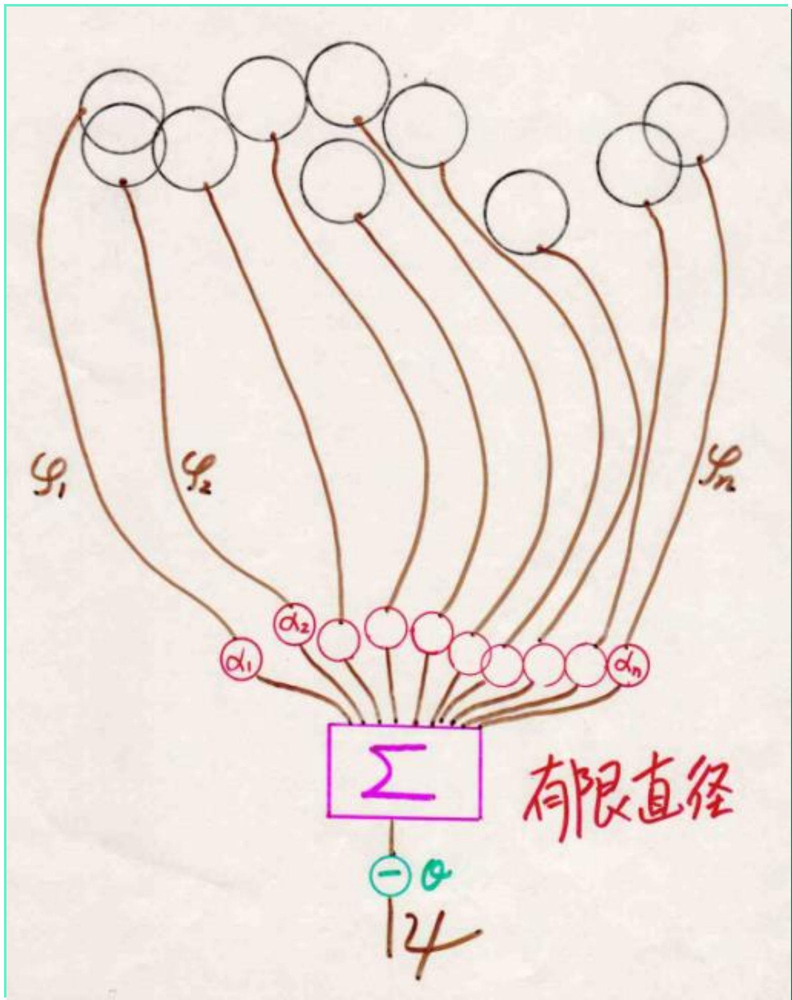
連結性の判断はできるか？

- 人は判断できる？？？



(a) (b)

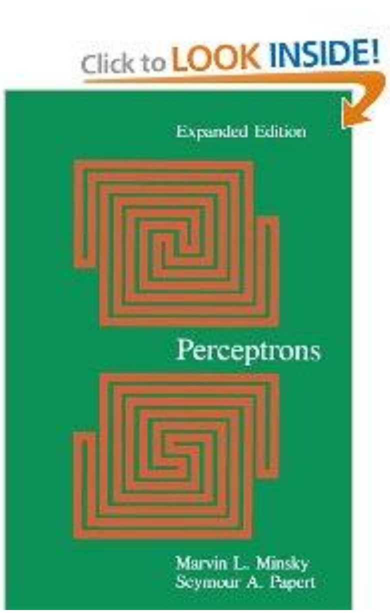
連結性の判断はできるか？



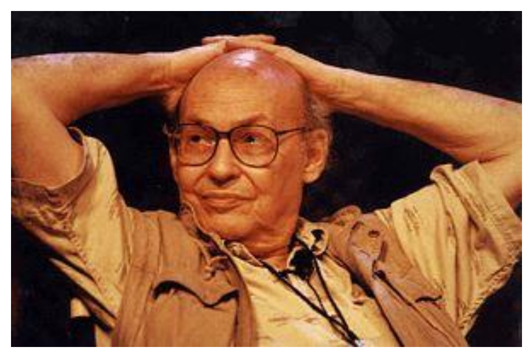
連結性が判断できるパーセプトロンがあったとする

連結性の判断はできるか？

定理 0.8：「任意の幾何学的図形が連結されているかどうかを判定する有限直径パーセプトロンは，存在しない．すなわち，有限直径のパーセプトロンは，$\phi_{\text{連結}}$ を計算することはできない．」



Marvin Minsky & Seymour Papert (1969).
Perceptrons, MIT Press, Cambridge, MA.



Marvin Minsky, 1927/8/9 -

単純パーセプトロンの限界

- 識別できる対象に限界がある
  - 線形分離可能性から来る制約
    - X-orの判別ができない
  - 有限直径パーセプトロンから来る制約
    - 強連結性の判断ができない
    - パリティ（偶奇性）の判別が不能
      - このため膨大な数の素子を必要とするので実用的でない

課題：　X-OR，Parityができないこと
を平易な言葉で説明する。

Minskyの悪魔

- この結果は非常に衝撃的であった
- 当時の数理学的･情報処理的な脳研究の楽観主義を一掃した
- AI一般に対しての過剰な懐疑論を生んだ
- Minskyからの自壊
  - K-line, society of mind
- PDPモデルなど多層ニューラルネットワークの研究が再生につながる