

認知心理学：第5講

表象：人間の扱うデータ形式

黒沢 学(kurosawa@cck.dendai.ac.jp)

前回の復習

- 注意
 - 入力された情報のうち意味あるものを選択するシステム
 - フィルタモデル/特徴統合理論

最初の質問

- ある部屋にはギリシャ語の辞書がおいてあり、中にギリシャ語の文を持っていくと、複数の人がそこにあるルールブック通りにそれを別な文字に置き換え、日本語の文として出力してくれる。しかし、それらの人はみなギリシャ語の意味が分かって作業しているのではない。では、「この人たちはギリシャ語を知っていると言える」のか?(中国語の部屋)

最初の質問

- 今からいう文を聞いてください。あとでそれに関して質問します。

表象とは

- 表象(representation: 表現/表示)
 - あるものをかわりに表しているする表現・表記など
 - 「人間の中で情報がどのような形をしているか」
 - 遠刺激→近刺激→心的表象
 - 人間が活動するためには近刺激を意味づける必要がある(ビデオカメラはアクセルを踏まない)

表象の性質

- 複数の表象がある
 - 表象された形式による分類
 - 命題的・構文論的な性質
 - アナログ的・イメージ的な性質
 - 分散表現(ネットワーク表現)
 - 表象されるモダリティによる分類
 - 音声的表象
 - 図的表象

命題表象

- 命題
 - 真偽が判断できる意味の最小単位
 - 内部構造を持つ(述語と引数)
- 命題表象が行われていることを支持する証拠
 - 長期的に保持されている情報は表面的な変化をしても命題の変化はない
 - 命題単位で再生が行われ、読みの時間もかかる
- 命題表現の方法
 - ノード・リンクによるネットワーク表現
 - E.g. Collins & Quillian の階層的な意味表現ネットワーク

イメージ表象

- イメージ表象: 具体的・視覚的・連続的な表象, 心的イメージ
- イメージ表象の存在を支持する証拠
 - イメージを伴うことで記憶成績が上がる; 場所法など
 - Pavioの2重符号化説
 - メンタルローテーション・メンタルサーチ
- イメージ論争
 - イメージは表象にとって本質的か? 情報処理上の付帯物? → アナロジー: テキストファイル・ビットマップファイル・オブジェクト化されたファイル
 - 図であっても「命題的に」計算されている証拠
 - イメージすれば視覚と同じ処理が行われる証拠

比喩: ファイル形式

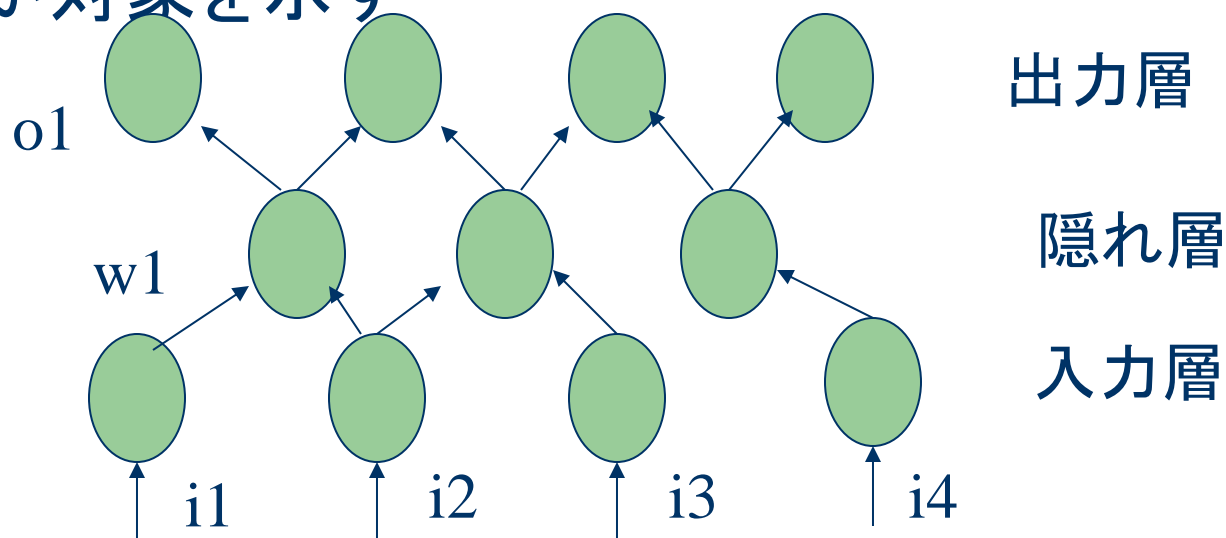
- テキストファイル
- バイナリファイル
 - 画像
 - ラスタ形式(photoshop...)
 - ベクタ形式(illustrator...)
 - 音声
 - ...

人工知能が持つべき知識とその性質

- 推論には背景となる知識が必要
- フレーム問題
 - 関係する知識を予め枠で囲えるかという問題
 - ロボットR1はフレーム問題に悩む
 - 不要な知識を無視するためには明示的に表象しないとならない
 - 人間は(一見)フレーム問題に悩まない
 - 不要な知識は思いつきもしない

分散表現

- ニューロンを模したネットワーク(ノード・リンク, ノード間のリンクの重みづけ)
- 対象は構文論的な性質を持たない: 活性化のパターンが対象を示す



分散表現の利点

- 表象の中には構文論的な特徴を持たないものもある
 - 意識的に行うわけではない直観的な推論
 - 筋肉運動の制御
 - 中国語の部屋-全体としてみれば高度に知的な作業をしているものがあってもその心的表現が構文論的な表象であるとは限らない
- 特定のノードが特定の対象物を表象するのではなく、複数のノードの活性化の重なり合いで多くの対象が表象される
 - 「おばあさん細胞」はあるのか？
 - 「優雅な退行」

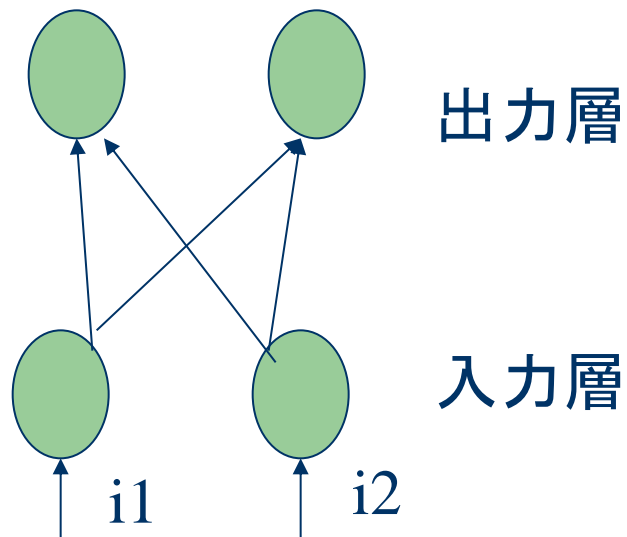
ニューラルネットの工学的応用(1)

- 家電などにおける曖昧な対象の制御
- 工学的な実装
 - 例えば, バックプロパゲーションによる重みづけの決定
 - ファジィ関数

ニューラルネットの工学的応用(2)

- パーセプトロン

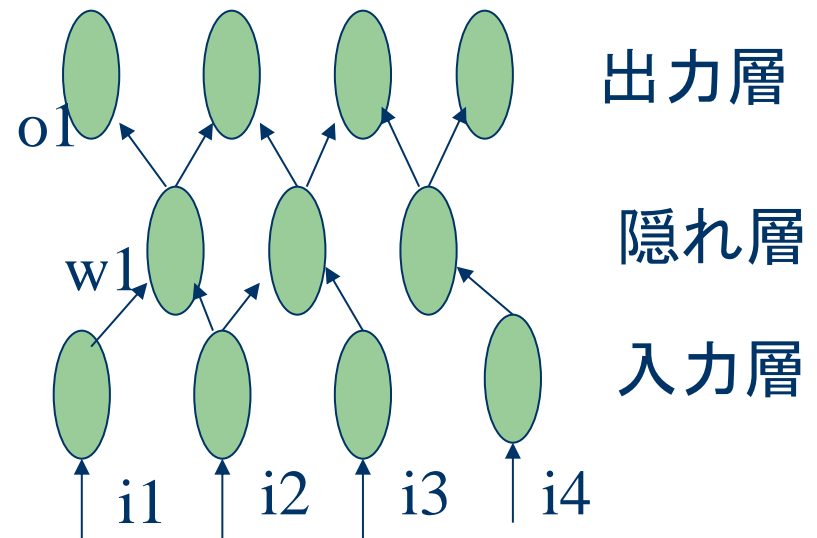
- 出力層と入力層が直接つながっている(隠れ層がない)
- 線形分離不可能な関係を表現できない



ニューラルネットの工学的応用(3)

- 隠れ層を作りネットワークに重みを学習させる

- 通常, 一つのユニットでは入出力にシグモイド関数 $f(x)=1/1+\exp(-x)$ を仮定
- リンクの重みづけは当初ランダムだが, 教師信号との差分を小さくするように変化させる($\Delta\theta = -\eta f'(x)$)
- 誤差を出力層から順次前に伝播(誤差逆伝播法)



表象：ひとまずまとめ

- 表象とは
 - あるものの表現, 特に心的に保持された情報の表現
 - さまざまな形式が取られる
 - 命題
 - イメージ
 - ノードの活性値
 - 保持するメカニズム-記憶

今日のまとめ

- 知覚/選択された情報は意味づけられ心的に表現される
- 表現のされ方にはさまざまなものがある(命題・イメージ・ネットワーク)
- ことば
 - 表象(命題的・アナログ的・分散表現)
 - 命題・述語・引数・ノード・リンク・階層的意味表現ネットワーク
 - 2重符号化説・メンタルローテーション・メンタルサーチ・イメージ論争・ニューラルネットワーク
 - 中国語の部屋

さらなる学習への手がかかり

- 休んだ者への問題
 - 日常行う作業(計算する・単語を覚える・運転する...)はどのように表象されていると仮定することで整合的に説明できるか. なるべくいろいろと例を挙げて考察せよ.
- 本とリンク
 - 考える脳・考えない脳 信原幸弘 講談社現代新書
 - 認知の微視的構造 クラーク 産業図書