

カテゴリー特異的な意味記憶の障害を階層構造と自己組織化マッピングを用いたモデルで説明する試み

浅川伸一（東京女子大学）

物品や対象の認識や呼称に選択的な障害を持つ脳損傷患者は、これら概念の認識過程にかかわる重要な手がかりを与えてくれる。とりわけある特定の対象の認識や呼称についてだけ障害のある「意味記憶のカテゴリー特異性」障害は示唆に富む。生物、果物野菜、道具などが選択的に障害されることが知られている。脳損傷後に現われるカテゴリー特異性を説明するためにはどのようなモデルが必要なのだろうか？

本稿では、これまで提案されて来た仮説が、2層の自己組織化マップ (SOM) とその間を繋ぐ Hebb 則による連合によって表現できるという単純なモデルを提案し、簡単な数値実験を行なった。

現在までに提案されて来た意味記憶に関するカテゴリー特異性を説明するモデルは3つに大別される。(1) Warrington & Shallice (1984) は、カテゴリー特異的な障害は対象の認識のためには種類の異なる情報が必要であると考えた。生物を区別するためには感覚的情報が決定的に重要であり、一方、機能的情報 (例えば、どうやってそれを使うのかといった) が非生物を区別するためには重要であると考えた。この「感覚/機能」仮説は、近年におけるカテゴリー特異的な障害の説明のためにもっとも広く用いられて来た仮説である。(2) Santos & Caramazza (2002) は、概念が「領域特殊性」によって体制化されているが故に概念毎に選択的な障害が発生すると論じている。(3) Humphreys & Forde (2001) たちの Hierarchical Interactive Theory (HIT) 仮説は、症例のパターンを説明する際に、関連したカテゴリーの事例の類似性が関与していることが強調されている。彼らは生物についてのカテゴリー選択的な障害は、これらの事例が知覚的に類似した事例を持つカテゴリーに属しているからだとする。

感覚/機能仮説に基づくにせよ、領域特殊性仮説に基づくにせよ、類似性仮説に基づくにせよ、ある特徴ベクトルによって表現された対象が互いに近傍に表象され、その損傷によってカテゴリー特異的な障害が発生すると考えれば、自己組織化写像による入力ベクトルの体制化機構を想定することで、これまでに提案されたモデルの要求を満足すると考えることができよう。なぜなら、感覚情報によって表現された概念と機能情報によって表現された概念とが異なる位置に表象されることは「感覚/

機能」仮説を実現したものであり、結果として形成される2次元布置がカテゴリー毎に体制化されているとすれば「領域固有性仮説」を実現したことになり、同時に類似性の判断に基づく類似性仮説を満たしていることになるからである。

本稿では Hinton & Shallice (1991) で用いられた文字の視覚情報と概念とを結びつけ深層失読をシミュレートするために用いられたデータを用いてモデルの検証を行った。データは室内の物品、動物、身体の部位、食物、室外の対象の5つのカテゴリー、計40事例である。各事例の意味ベクトルは例えば max-size-less-foot や has-legs などの項目からなる68次元の0,1ベクトルであり、語彙ベクトルは3文字あるいは4文字からなる単語の0,1表現からなる28次元ベクトルであった。意味、語彙それぞれに対してSOMを適用し、入力ベクトルの2次元の布置を学習させ、同時にSOMのwinner-take-all回路によって勝者となった意味SOMユニットと語彙SOMユニットの間でHebb則による学習を行わせた。

本稿で用いたアルゴリズムは以下の通りである。各入力ベクトル x に対して最もマッチしたSOMのユニット i を winner-take-all 回路

$$x \mapsto c(x) = \underset{i \in \{1, \dots, N\}}{\operatorname{argmin}} d_w(w_i, x). \quad (1)$$

を定義し、学習則

$$\Delta w_i(t) = -\alpha(t)n(i, c(x)) \frac{\partial d_w(w_i, x)}{\partial w_i} \quad (2)$$

を適用する。ここで $n(x, y)$ は近傍関数であり $n(x, y) = \exp(-|x - y|^2/\sigma^2)$ とした。 $\alpha(t)$ が学習係数であり、学習回数 t の単調減少関数である。ここでは $\alpha(t) = \eta(1 - t/t_{\max})$ とした。 η は定数である。 α の存在によってSOMの収束が保証される (Kohonen, 1997)。

2つのSOMの勝者ユニット間を結ぶHebb則においては活性値を $[0, 1]$ に制限するために入力 x に対する応答関数として $y = 1 - \exp(-\beta x)$ を用いた。従って通常のHebb則にはなく、2つのSOMの勝者同士を結ぶ学習の際には

$$\delta w = \beta \exp(-\beta x) \quad (3)$$

という学習則を用い、それ以外のときには指数関数に

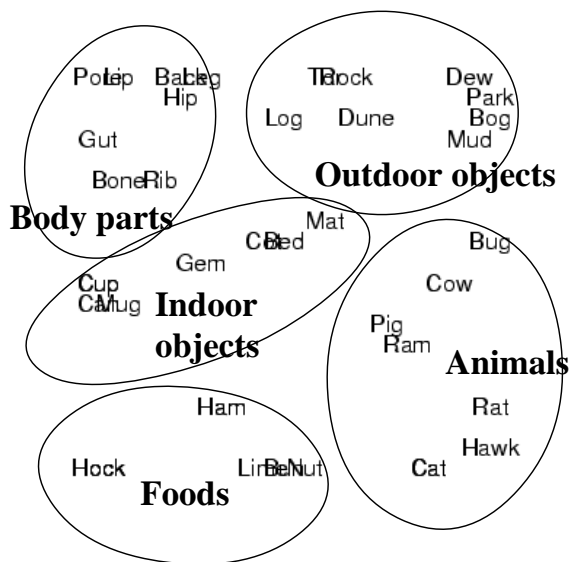


図 1: 意味 SOM の結果

従って忘却させることにした。

$$\delta w = \gamma w \quad (4)$$

忘却に指数関数を用いたのは、SOM の学習途中に勝者が変化した場合、過去の情報を速やかに消去し、新たな結合を学習しなければ意味 SOM と語彙 SOM との間に一対一対応の学習が成立しにくくなるからである。β が大きければ新たに勝者となった SOM のユニット間の学習が速やかに成立し、γ が大きければ忘却しにくくなることを意味する。これは意味 SOM と語彙 SOM の学習時に勝者ユニットが一貫している場合には問題を生じないが、SOM の学習にともなって勝者ユニットに移動や変化が生じた場合、意味と語彙との間に混乱が生じることを意味する。

20 行 20 列の SOM に Hinton & Shallice (1991) の 5 つの 카테고리 40 事例を学習させた意味 SOM の結果を図 1 に示した。図から 5 つの カテゴリ が分離されて表象されていることが分かる。このことから SOM が カテゴリ 特異性を説明するモデルとして有効であることを示唆していると考えられる。すなわち、図 1 において動物を表す領域に損傷が生じると動物概念に特異的な意味記憶の障害が生じるものとみなすことができよう。

一方、Hebb 則による意味 SOM と語彙 SOM との結合の Hinton グラフの一例を図 2 に示した。この結果は、2 つの SOM の学習と提案した Hebb 則の変形による学習則によって意味と語彙との間に一対一結合が成立しうることを示したものであり、本稿で提案したモデルの妥当性を示すものと言える。しかし、図 2 右からもわかるとおり、完全に一対一対応が取れる場合ではばかりではないことも確かである。これには二通りの解釈が可能であ

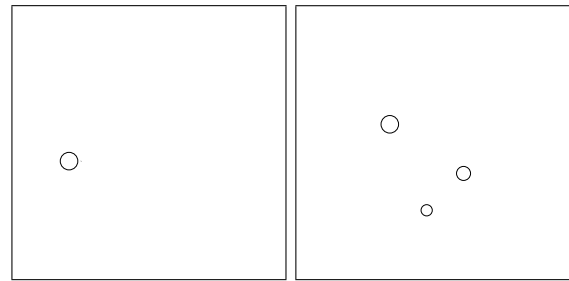


図 2: 2 つの SOM を繋ぐ結合の Hinton グラフ

る。一つは、学習パラメータ β, 忘却パラメータ γ の決定問題である可能性である。もう一つの可能性は SOM によって意味と語彙とを徐々に学習していく過程で、両 SOM 間の対応問題を同時に学習していくという本モデルが持つ機構そのものに起因する限界である。

上述のようなパラメータの恣意性、もしくは限界が考えられるものの、カテゴリ 特異性を説明する試みとしてモデルが実装できた意義は大きいと考える。とりわけ、感覚/機能、領域固有性、そして類似性仮説という 3 つの仮説はニューラルネットワークモデルとして実装した場合、一つのモデルを別の側面から論じているものであると言う可能性を示すものであることは、今まで論じられてこなかった。今後この観点を考慮して検討を進めて行く必要があるだろう。

引用文献

- Hinton, G. E. & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, 98(1), 74-95.
- Humphreys, G. W. & Forde, E. M. (2001). Hierarchies, similarity, and interactivity in object recognition: "category-specific" neuropsychological deficits. *Behavioral and brain sciences*, 24, 453-509.
- Kohonen, T. (1997). *Self-Organizing Maps Second Edition*. Springer.
- Santos, L. R. & Caramazza, A. (2002). The domain-specific hypothesis: a developmental and comparative perspective on category-specific deficits. In M. E. Forde & G. W. Humphreys (Eds.), *Category Specificity in Brain and Mind* (pp. 1-23). Psychology Press.
- Warrington, E. K. & Shallice, T. (1984). Category specific semantic impairment. *Brain*, 107, 829-854.