神経心理学ツールとしてのニューラルネットワーク

浅川伸一 <asakawa@twcu.ac.jp>

以前にも説明したとおり、ニューラルネットワークを破壊することで言語を含むさまざまな認知機能の障害をコンピュータ上に再現できる。人間の認知機能とニューラルネットワークプログラムとを同一視し、かつ、ニューラルネットワークを部分的に破壊することと脳損傷を同一視することとによって、近年の認知神経心理学は大きく変化してきた。

ここではまとめとして、以前に言及した話題と重複する内容であるが、伝統的な考え方であるボックス–アンド–アロー・モデルに対する批判を紹介し、次に意味記憶の構造とその障害におけるニューラルネットワークモデルを紹介する。

1 ニューラルネットワークによる脳損傷の解釈について

ニューラルネットワークを用いた脳損傷のシミュレーション研究では特定の認知機能を遂行するためのニューラルネットワークモデルをコンピュータプログラムとして実現し、構築されたニューラルネットワークの一部を破壊することによって対応する部位が損傷を受けたときに生ずる症状をプログラムの出力として表現することをめざしている。こうした研究はいわば人工脳損傷とも言える研究である。人間の言語活動を理解する上でも、あるいは実際の脳損傷患者の症状を理解するためにも、コンピュータを用いた人工脳損傷研究は重要だと考える。倫理上の制約から実際の人間の脳を破壊して実験を行なうことは不可能だからである。ニューラルネットワークによる人工脳損傷研究は、実際の脳損傷患者を扱う神経心理学に対して強力な道具を提供していると言えるだろう。

ニューラルネットワークの破壊実検、切除実験は重要な情報を与えてくれる、その解釈には注意を要する。ラジオから抵抗を取り外したらピーッといったからといって、その抵抗がピーの抑制中枢であるとはいえない。ほとんどの人間は噂話が好きであるが、グループは、ある人がそのグループを離れると、その人の噂話をする。そのときその人をそのグループの噂話抑制中枢と

するようなものである。

ニューラルネットワークは相互作用をする複雑なシステムとしての行動と、そのシステムが損傷したときの効果との複雑な関係を推論する手段を提供している。そのような推論が明示的で機能論的な検証、すなわちシミュレーションによる検証が可能になったことが重要である。特定の認知機能の検査結果とその認知システムの障害の部位を特定することの関連は、かつて信じられていた程明らかな関係にはないのである。

2　ボックス–アンド–アロー・モデル批判

神経心理学の分野ではボックス–アンド–アロー・モデルを元にした考察と症例検討が行なわれることが多い。課題遂行に必要な下位機能をそれぞれのボックスで表現し、ボックス間を流れる情報を矢印で表現する。脳内の各領野で行なわれている処理がボックスに対応し、領野間の線維連絡が矢印に相当する。従って神経心理学症状は、ボックスで表現されている機能が損なわれる (損傷) か、または、ボックス間を結ぶ矢印が途切れる (離断) ことによって症状を記述しようとするのがボックス–アンド–アロー・モデルの神経心理学的解釈である。ここには、異なる症状が異なる脳内機構を反映しているという暗黙の了解が存在している。このようなボックス-アンド-アロー・モデルの考え方は一定の成果を挙げてきたし、現在も有効な手法である。

典型的なボックス–アンド–アロー・モデルであるウェルニケ–リヒトハイムの図式を紹介する (図 1)。図中の A は聴覚的イメージの座、M は発話に用いられる運動表象の座。B は言語の概念が保持されている部位、a は聴覚情報の入力系、m は運動機能の言語中枢である。このモデルによれば、言語理解は $a \rightarrow A \rightarrow B$、発話は $B \rightarrow M \rightarrow m$、復唱は $a \rightarrow A \rightarrow M \rightarrow m$ となる。M が損傷されると理解はできるが自発的な発話と復唱ができなくなり (ブローカ失語)、A が損傷を受けると言語理解と復唱の障害が生じる (ウェルニケ失語)。実際この図式は患者の症状を良く説明し、モデルから予測される症状を呈する患者の発見もあって有効性が確認されている (岩田 [2])。$A \rightarrow M$ の結合が断たれると復唱のみができなくなり、言語理解も自発発話も保たれる伝導失語が (図中の 3)、$A \rightarrow B$ 間の結合が切れると復唱はできるが言語理解が不能な超皮質性感覚失語が (図中の 6)、また、$B \rightarrow M$ の結線が切断されると復唱はできるが自発発話が侵される超皮質性運動失語が生じることになる。さらに、$a \rightarrow A$ 間の障害では、聴覚的言語理解が可能だが書字理解が不可能な純粋語聾が、$m \rightarrow M$ では書字障害を伴わず発語に障害が見られる純粋語唖が生じる。

ボックス–アンド–アロー・モデルは患者の症状を良く説明しているため有効であると考えられている。しかし、ボックス–アンド–アロー・モデルに対して以下のような 2 つの本質的な欠点を指摘できる。

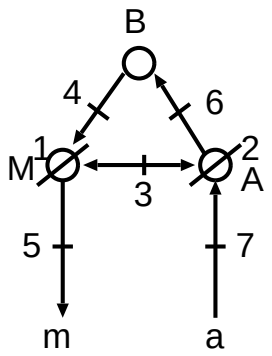

図 1: ウェルニケ–リヒトハイムの図式。

1. モデル内を流れる情報についての言及が曖昧であるため、典型的な反応パターンがどのようにして生じ、反応パターンがなぜ一貫しているのかが説明できない。それぞれの経路における情報処理様式や神経回路の損傷によって生じる誤処理、誤動作のモデルを考える必要がある。

2. 伝統的な神経心理学では、2 重乖離の原理 (principle of double dissociation)、離断の原理 (principle of disconnection) など (山鳥 [1]) によって特定の認知機能の機能局在が論じられる。この立場は特定の認知機能の責任領野が、任意の部位に局在していることを暗黙のうちに仮定している。しかし、ニューラルネットワーク研究の破壊実験では単一部位の損傷によって複数の認知機能障害を同時に説明できる可能性がある。反対に、複数の部位を破壊または離断しなければ高次認知機能障害は発生しないかも知れない。特に高次認知機能においては、シミュレーションによって確認してみるまで分からないという側面が強い。実際の患者に対して破壊実験を行なうことは不可能であるから、コンピュータ上での破壊実験は、病態の真の理解のためには欠かせない研究である。わが国においてはこのようなシミュレーション研究は極めて少ないと言わざるを得ない。

ニューラルネットワークモデルの存在意義はまさにこのような点にある。一旦学習が成立したネットワークを破壊することで症状を模倣させ、モデルの挙動から人間の情報処理モデルを検討することがニューラルネットワーク研究の神経心理学への貢献であると言えよう。

3　意味記憶の構造のシミュレーション

3.1　生物、非生物の二重乖離

意味記憶は、個々の対象についてカテゴリー毎に構成されているのかそれとも、それとも意味記憶はモダリティー毎に構成されているのか、についてファラーとマクレランド [3] が行なった研究を紹介する。彼らのモデルによれば、モダリティーに依存した意味記憶表象を考えれば、カテゴリーに基づく意味記憶表象は説明できることを意味し、認知心理学でしばしば話題になる記憶表象論争に対する一つの解答を与えている。

3.2　神経心理学的症状

実際の脳損傷患者の中には、動物や植物などの生物の知識について障害がある一方で、非生物の知識については健常のまま保たれている患者が存在する [4]。古典的な二重乖離の原則から、生物と非生物の知識の脳内での意味記憶には、生物と非生物とを独立に表象している意味記憶が存在すると仮定される。しかし、ウォリントン (Warrington) とシャリス (Shallice) は、生物の知識と非生物の知識との間で選択的な障害が起こるのは、異なる感覚運動経路からの情報の重みの差異を反映しているためではないか、と述べている。すなわち、生物は主に感覚的な性質によって互いを区別することが多いが、非生物は主に機能によって分類される。ある動物、例えばヒョウは、他の肉食動物と比べて主に視覚的な特徴によって差別化される。これとは対照的に、机の知識については、他の家具との違いを記述するときには主に機能、すなわち何のために使うのか、によって差別化される。それゆえ、障害のある知識と健全に保たれている知識との違いは、生物–非生物の違いなのではなく、対象を記述している特徴が感覚–機能の違いであるかも知れない。

ファラー (Farah) とマクレランド (McClelland)[3] のモデルは上記の感覚–機能仮説が意味記憶障害を説明できることを例示するために作成された。分散表現された 2 種類の意味記憶、すなわち視覚的意味記憶と機能的意味記憶とが損傷を受けたとき、生物–非生物の知識が選択的に障害されることを説明できる。

3.3　モデル

彼女らのモデルを図 2 に示す。 3 つのユニット群、すなわち記憶を表現する意味記憶系と、入出力を表現する 2 つの周辺系、視覚ユニット群と言語ユニット群とがある。言語ユニット群と視覚ユニット群との間を除いて、全てのユニットに群間および群内結合が存在する。各ユニットの活性値は $(-1, +1)$ の間をとる実数値であり、しきい値は無い。各ユニットの活性値は同期的に

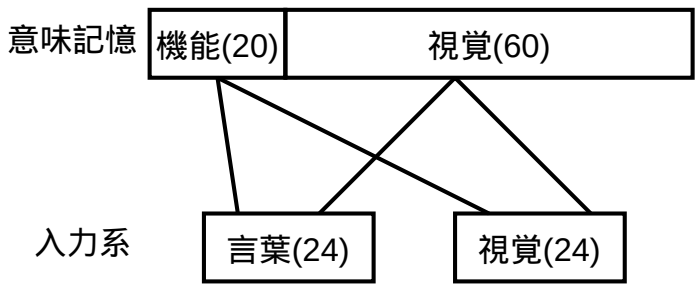

図 2: ファラーとマクレランド (1991) の意味記憶モデルの概念図。カッコ内の数字は数値実験で用いられたユニット数を表す。意味記憶内で機能的記憶と視覚的記憶のユニット数が異なるのは、彼らの論文中の実験 1 (心理実験) の結果を反映している。

更新された。すなわち、それぞれのユニットの活性値を計算するタイミングは全ユニットについて同時である。各周期ごとに、そのユニットに対して出力を送っている全てのユニットの活性値と結合係数からそのユニットの活性値が計算される。

それぞれ 10 個づつ、生物と非生物を表す刺激が乱数を用いて作成され、モデルに提示された。生物と非生物とを表す項目は、視覚情報と名前情報との比率が変えられた。生物項目では平均して 16.1 の視覚意味記憶ユニット、2.1 の機能意味記憶ユニット。非生物では 9.4 の視覚意味記憶ユニット、6.7 の機能意味記憶ユニットを用いて表現された。

誤差修正学習手続きによってネットワークが訓練された。絵画パターンが提示されたときに対応する意味記憶パターンと言語パターンが産出されるように、また、言語パターンが提示されたときには対応する意味記憶パターンと視覚パターンが産出されるよに訓練された。

各訓練試行では、生物、もしくは非生物に対応する絵の名前が名前ユニット群あるいは絵画入力ユニット群に対して提示され、そして、ネットワークは解が安定するまで 10 サイクルの活性値の更新が行なわれた。ユニット間の結合係数は学習についてはデルタルールが用いられた。

3.4 破壊実験

彼女らは、実験 3 でモダリティに依存した意味記憶障害、すなわち、耳で聞いたときには理解できるが、目で見たときには特定のカテゴリーについての知識に障害を生じる患者のシミュレーションを行ない。実験 4 では、雑音を加える方法によるシミュレーションを行なっている。

結果から、モダリティ依存の意味記憶構造だけを用いて、カテゴリー依存

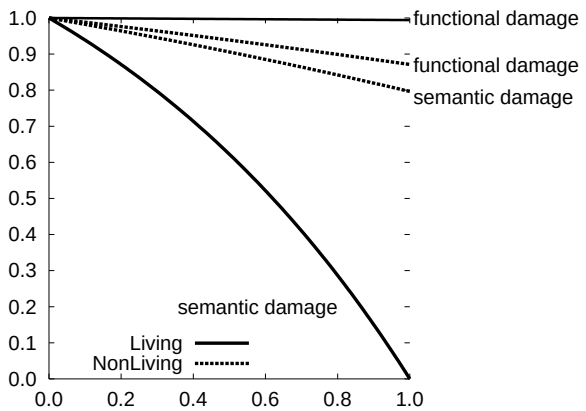

図 3: 生物–非生物別の意味記憶内の損傷の程度と課題成績との関係 (ファラーとマクレランド (1991) の表 3 と図 2 より改変)。彼らは各条件で 0, 0.2, 0.4, 0.6, 0.8, 0.99 の 6 点について各 5 回づつしか数値実験を行なっていないため実際の曲線は滑らかにならない。そこで、指数関数に回帰させてプロットしなおした。指数回帰を用いれば唯一のパラメータを変化させることで 4 つの条件に対応する曲線を描くことができる。

の障害を説明できることが確認された。

参考文献

[1] 山鳥重. 神経心理学入門. 医学書院, 1985.

[2] 岩田誠. 脳とことば —言語の神経機構. 共立出版, 1996.

[3] M. J. Farah and J. L. McClelland. A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, 120(4):339–357, 1991.

[4] E. K. Warrington and T. Shallice. Category specific semantic impairments. *Brain*, 107:829–854, 1984.