

早稲田大学審査学位論文(博士)の要旨  
3661-2

2003 早大学位記 報告番号  
3661 甲乙 1836

博士学位論文概要書：  
ニューラルネットワークの破壊実験によっ  
て表現された神経心理学的モデルの考察  
浅川 伸一

W  
学位論文  
3661  
2

博士学位論文概要書：  
ニューラルネットワークの破壊実験によっ  
て表現された神経心理学的モデルの考察

浅川 伸一

# 1 序論

## 1.1 神経心理学における道具としてのニューラルネットワーク

本論文は、神経心理学の理論的背景を提供する道具としてニューラルネットワークを用いることを提案するものである。ニューラルネットワークを用いた神経心理学的研究とは、神経心理学に対してニューラルネットワークを応用することにより、計算論的観点から神経心理学的症状の理解を目指すものである。ニューラルネットワークを用いることによって神経心理学的症状のいくつかはまったく別の視点で捉え直すことができることを主張する。

神経心理学の分野で提案されているモデルはすべてニューラルネットワークモデルとして書き下すことができ、コンピュータ上に実現できる。いったんコンピュータ上に実現されたモデルを系統的に破壊することで神経心理学的症状の理解を試みる研究手法がニューラルネットワークを用いた神経心理学の標準的な手法である。

## 1.2 ニューラルネットワークの特徴

ニューラルネットワークにおける特徴を挙げるとすれば、分散表現と統計的構造の漸進的学習、および相互作用の3点に要約できる (Farah, 1994; Plaut, 2001)。

**分散表現:** ニューラルネットワークにおいては、知識はそれぞれのユニット集団の活性化パターンとして表現される。例えば、ある単語の意味は別の単語の意味とは異なる活性化パターンとして表現されており、類似した概念は互いに類似した活性化パターンとして表現される。

**統計的構造の漸進的学習:** 長期的な知識はユニット間の関係、すなわちユニット間の結合強度としてネットワーク内に埋め込まれている。ユニット間の結合強度は学習によって徐々に変化する。すなわち、学習により外界から提供される情報 (単語間の類似度や相互関係など) の統計的性質が徐々に獲得される。

**相互作用:** ユニットの密接に連結されており、相互に影響しあう。すなわちユニット間の結合強度に応じて、互いに活性化パターンを強め合ったり、弱め合ったり、振動したりというような複雑な相互作用をする。

ニューラルネットワークを破壊することで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。人間の認知機構を理解する上でも、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られる。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。

### 1.3 新しい脳観

伝統的な神経心理学では、二重乖離の原則に基づいて特定の認知機能の機能局在が論じられてきた。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。

### 1.4 神経心理学のツールとしてのニューラルネットワークがみたすべき要件

ニューラルネットワークは非常に記述能力が高いためにニューラルネットワークモデルを作れば大抵の心理現象はシミュレートできてしまう。この数学的な拘束によって、外界の刺激入力に対して応答するような、ほとんどの心理現象はニューラルネットワークで説明できることを意味する。ほとんどの心理現象が説明可能であることは、単にシミュレーションを行なって動作するだけのモデルは必ず作ることができることを意味する。従ってそのようなモデルは生理学的、神経心理学的な妥当性を持たず、意味が無い。モデルが神経心理学に貢献するためには、以下のような要件を満たしている必要がある。

1. 無矛盾性. 生理学や機能的脳画像研究から得られたデータと矛盾してはならない。
2. 予測可能性. モデルは当該の症状を説明するだけでなく、症状が進行したときの予測や、新たな症状のおきる可能性を指摘できなければならない。
3. 説明可能性. ニューラルネットワークモデルによって複数の心理指標が同時に説明されなければならない。

### 1.5 本論文の構成

本論文の第一部では伝統的な神経心理学の原則をニューラルネットワークから見た場合の批判を取り上げる。二重乖離の原則に基づく局在性仮説は神経心理学の原則であり、基本的なモデル構築のための推論方法である。この

原則および推論方法に対してニューラルネットワークから見れば別の視点が提供できることを示す。

第二部では、応用的なモデルの幾つかを取り上げ実際にどのような研究が可能かを議論する。具体的には、日本語の文字知覚のための 2 段階読字過程モデル、単語の読みにおける労働の分割問題の解としてのエキスパート混合ネットワーク、リハビリテーションへの応用と 3 種類のリハビリテーション技法へのニューラルネットワーク的な意味付けについてである。

本論文の第三部では総まとめとして第一部と第二部の議論を元に脳の構成論的研究の重要性を強調する。

# 第 I 部 神経心理学の基礎概念 及び推論方法への批判

## 2 局在性仮説批判

神経心理学の分野では機能局在性仮説が重要な役割を果たしてきた。機能局在性仮説とは、心の機能が独立した下位モジュールによって構成されているとする考え方である。各モジュールは情報論的にカプセル化されていて、あるモジュールの損傷が別のモジュールの機能に影響を与えることはほとんどなく、あるモジュールの損傷は比較的単純なひとつの認知機能の低下として表出する、というのが局在性仮説である。局在性仮説の背景には、各認知モジュールは専用のハードウェアによって実現されており、局所的な脳損傷によって選択的にその認知モジュールだけが障害を受ける、という暗黙の仮定がある。

もし仮に、我々の認知機構の各モジュールが相互に関連し依存しあい、複雑な相互作用をするのであれば、一つの認知モジュールへの障害は他の認知モジュールにも影響を与える可能性がある。近年の PET や fMRI に代表される機能的脳画像研究によって明らかになったことは、単純な課題を遂行しているときでさえ、多くの部位に活性化が見られることである。すなわち、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、それ程明らかな関係にはないと思われる。

## 3 二重乖離の原則批判

### 3.1 アルツハイマー病の呼称課題における視覚性および語彙性の誤りに関する二重乖離

アルツハイマー病の患者の物体呼称課題における障害には、意味記憶の障害説の他に、視覚性障害、語彙性障害の説明が可能である。Tippett と Farah(1998)によるニューラルネットワークを用いた研究によれば、意味記憶の障害のみによってアルツハイマー病の物体呼称課題における成績低下を説明できる可能性がある。意味的な知識表象と視覚的および語彙的な知識表象は密接に関連しあっているのである。その結果、一つの構成要素、例えば意味記憶に障害があると、その障害の影響は、視覚性の知識にも語彙の知識にも影響を与えられられる。

### 3.2 生物, 非生物概念の二重乖離

認知心理学でしばしば話題になる記憶表象論争に関して, 意味記憶は, 個々の対象についてカテゴリ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, という論争がある. Farah と McClelland(1991) が行なったニューラルネットワークによる研究によれば, モダリティーに依存した意味記憶表象を考えれば, カテゴリに基づく意味記憶表象は説明できる.

彼女らのシミュレーション結果は, モダリティー依存の意味記憶構造を用いれば, カテゴリ依存の障害を説明できることを示している. すなわち生物, 非生物という異なるカテゴリに属する事物は脳内で異なって表象されているのではなく, 意味記憶の構造は入力モダリティーに依存して形成されているという結果が得られた.

### 3.3 機能的失文法と統辞的失文法の二重乖離

二重単純再帰型ニューラルネットワークは言語産出と言語理解が密接に関わっていることを表す, おそらく最も単純なモデルである. 文章産出と文章理解とは密接にからみ合っており, 切り離すことは難しいということである. 学習の成立した二重単純再帰型ニューラルネットワークに対して共有されている文脈層を破壊することによって人工脳損傷をおこさせてネットワークの振舞を観察した.

従来の神経心理学の枠組みでは, 形態的失文法と統辞的失文法とは二重に乖離しているので異なる脳内モジュールが障害を受けたと考えざるを得なかった. ところが, 本節の二重単純再帰型ニューラルネットワークの人工脳損傷のシミュレーションでは同一モデルで二重に乖離した二つの失文法を文脈層ユニットに形成されたと考えられる —文章産出と文章理解のために同時に利用される— 文法知識の障害として説明できる. しかも, この文法知識はア prioriに与えられたものではなくニューラルネットワークの訓練の結果として文法知識が創発されたのである.

### 3.4 まとめ

以上のように, 脳内で各認知モジュールは情報論的にカプセル化されていて各モジュール間は単純な情報が伝達されるだけであるという局在性仮説は, ニューラルネットワークモデルによって全く異なる見方が提供できる. 少なくとも二重乖離という現象を盲目的に当てはめることで, 異なる認知モジュールを想定することはすべきではないだろう.

# 第II部 ニューラルネットワークによる神経心理学的モデルの検討

## 4 2段階読字過程モデル

今日までに、多くの読みのモデルが認知心理学、神経心理学の分野で提出されて来ている。中でも Dejerine が提唱し Geschwind(1965) によって再評価された視覚言語離断仮説は、今日でもなお純粹失読症状を説明するモデルとして妥当であると考えられる。しかし、近年異なる説明モデルも提出されている。たとえば、Farah(1990) は視覚障害が純粹失読を生起させる可能性を指摘している(視覚性障害説)。

本章では紡錘状回後部から角回/縁上回の経路に着目してモデル化を行なった。紡錘状回後部は視覚入力とその特徴に従って大まかに分類する役割を担っていると考えられる。一方、Geschwind(1965) によれば角回は視覚文字-音韻変換の中継路ではなく、異種感覚記憶痕跡間の連合を営む領域である。この説が正しいとするならば、角回で表現されている情報は特定の感覚に依存しない抽象化された文字の表現でなければならない。本章ではこれらの事実と現在までにニューラルネットワーク研究で知られている2つの回路、自己組織化マッピングとパーセプトロンが対応すると考えてモデルを構築した。

### 4.1 モデル

本章で提案するモデルは2段階で構成されている。入力パターンはこの2段階を逐次経過して処理される。

第1段階では、入力された視覚パターンはその形態的特徴に従って高速に、大まかに分類される。大分類を担当する第一段階の回路は、類似した入力刺激を2次元平面上の類似した位置に写像する機構である。入力刺激を2次元平面上に写像することは、大脳皮質が2次的に広がっていて、類似した特徴に反応するニューロンが皮質の近傍に分布するというトポグラフィックマッピングに対応すると考えられる。本章では、第1段階での処理は自己組織化アルゴリズムによって訓練されると仮定した。第1段階で学習すべき文字の候補数を絞り込むことは、第2段階での認識機構において、学習に必要な時間を少なくし、かつ、学習の収束可能性を改善するという計算上の効率性が指摘できる。加えて、2段階の処理を仮定することは、全ての刺激入力を単一の大回路で処理する場合に比べて、ネットワークの経済性、すなわち、素子間の総結合数をより少なく押える効果も存在する。



第2段階での処理の目標は、入力画像の詳細な特徴に従って、視覚入力に対応する入力画像に依存しない文字コードに変換することである。本章では、第2段階の文字同定処理にパーセプトロン型の3層の回路を用いた。

いったん訓練が終了したネットワークに損傷を加えると、失読症、とりわけ形態失読に似た出力を生じることが期待できる。本章では2種類の損傷シミュレーションを行った。

## 4.2 計算機実験

### 4.2.1 離断仮説

第2段階の回路の任意の結合を切断すれば視覚言語離断仮説をシミュレートできる。障害の程度が重ければ重い程、認識率が低下することが示された。切断率とモデルの示す成績との関係は単調減少の傾向にあるが、変化は一定ではない。切断率が0.3付近まで、急激に認識率が低下する。このことは損傷の程度が軽い場合でも認識不能な文字が多いことを示している。もし実際の損傷あるいは病変が本モデルで記述しているとおりならば、以下の2点が演繹できる。

1. 本モデルの第2段階に相当する領野に進行性の病変が存在すれば、患者の読字能力は急激に悪化する。
2. 損傷部位の面積と課題成績とは直線的ではなく、損傷の程度が少なくても急激に成績が低下すると予想される。すなわち、失読症の患者間で成績を比較する場合、小さな損傷領域の差異が大きな成績差となって現われる可能性がある。

### 4.2.2 視覚性障害説

純粹失読の患者には、ある種の視知覚的な障害を持つものがあることを報告されている (Farah, 1990)。この仮説を調べるために雑音の付加によるシミュレーションを試みた。

離断仮説によるシミュレーションと対照的に、視覚性障害説では、雑音の程度に対して頑健な傾向がある。雑音のレベルが高くなれば成績が悪くなるが比較すれば成績低下の割合は少ない。このことから、障害の影響が2つの型の失読症で異なることが予測できる。

## 5 単語の読みにおける労働の分割問題に対する解としてのエキスパート混合ネットワーク

### 5.1 二重経路モデルとトライアングルモデル

二重経路モデルは Coltheart ら (Coltheart, Curtis, Atkins & Haller, 1993; Coltheart & Rastle, 1994) によって開発された記号処理モデルである。印刷文字を音韻へ変換するための明示的な規則に基づく直接経路と、規則にあてはまらない単語を読むための単語ベースのルックアップテーブルをもつ間接経路(語彙経路)とから構成されている。印刷された文字を読むときには、直接文字を音韻へと変換する直接経路と、例外語の語彙テーブルまたは意味を介して発話に至る間接的な経路との2つの経路を仮定するのが二重経路モデルである。

これに対してトライアングルモデル (Seidenberg & McClelland, 1989; Plaut, McClelland, Seidenberg & Patterson, 1996) では、同時的、相互作用的处理が仮定される。トライアングルモデルにおける直接経路では、一貫語と高頻度の例外語が学習される。一方、低頻度の例外語は意味系に依存すると仮定される。すなわちトライアングルモデルにおける直接経路では一貫語と高頻度例外語が学習される。一貫語および高頻度例外語と低頻度例外語との処理の違いには労働の分割と呼ばれる作用が関与する。

### 5.2 論争

二重経路モデルかトライアングルモデルかという論争の論点は、単一経路か二重経路かという議論ではなく、直接経路と間接経路の処理の違いに関してである。

二重経路モデルの特徴は明示的で記号处理的なルールすなわち単語毎のルックアップテーブルを用いることである。すなわち二重経路モデルでは2つの経路間の離散的なスイッチが仮定されている。これに対して、トライアングルモデルではそのようなルールあるいはスイッチの存在を仮定しない。代わりに、トライアングルモデルではニューラルネットワークの特徴である同時的、相互作用的处理が仮定される。2つの経路間の離散的なスイッチを持つ代わりに、ニューラルネットワークアプローチでは同時的、相互作用的处理を含みより軟らかい労働の分割が行なわれる。

### 5.3 労働の分割問題

トライアングルモデルはニューラルネットワークの特徴である相互作用をする。いったん書記素層へ入力された文字は、意味層を介する間接経路と音

韻層に直接出力を送る直接経路との両方の影響を受ける。ある単語がどちらの経路をたどって読まれるかは、単語ごとに、また、個人ごとに異なるとされる。直接経路では規則語と高頻度例外語とが処理され、低頻度例外語は間接経路である意味層のサポートを必要とする。このように単語ごとに2つの経路の影響が異なって表現されることを労働の分割問題という。

ただしトライアングルモデルでは労働の分割問題を完全に実装したわけではない。

#### 5.4 エキスパート混合ネットワークによる労働の分割問題の解

もし、書記素-音韻対応規則を学習し、非一貫語と一貫語を自動的に分類して学習できるアルゴリズムが存在すれば、表層失読と音韻失読という神経心理学的症状を説明するための労働の分割問題を解決できるモデルになるだろう。この語彙の自動分類機構を実現するのがエキスパート混合ネットワーク Mixtures of experts (以下 ME と略記) とよばれるニューラルネットワークモデル (Jacobs, Jordan, Nowlan & Hinton, 1991; Jordan & Jacobs, 1994) である。

ME とは入力データ空間をいくつかの小領域に分割し、その分割された各領域に対してひとつのニューラルネットワークを割り当てることによって、複雑な問題の解を求めるため手法である。複雑な問題を分割して小領域に区切ることによって、1つの大きなニューラルネットワークを学習させるよりも効率の良い学習をさせようというモデルである。

ME はエキスパートネットワークとゲーティングネットワークから成り立っている教師あり学習の一手法である。ゲーティングネットワークは問題空間を分割するために用いられ、各エキスパートネットワークは分割された領域内での局所的な解を出力する。

トライアングルモデルは直接経路と意味層を介した間接経路という2つのエキスパートネットワークを持つ ME とみなすことができる。ME において出力が混合されることと労働の分割問題は直接対応すると考えられる。すなわち、トライアングルモデルにおける労働の分割問題は ME におけるゲーティングネットワークによる領域の分割そのものである。

#### 5.5 数値実験

全てのモデルはある程度人間のデータを再現できているとみなすことができた。したがって労働の分割問題の解としてはどれを選んでも大差が無いように見うけられた。したがってモデルの優劣をきめるのは、課題成績だけではなくその背後にあるモデルの持つ本質を議論しなければならないだろう。ある程度万能なアルゴリズムを用いれば任意の精度で望む関数可以实现できる

ことはノーフリーランチ定理が主張しているところでもある (Duda, Hart & Stork, 2001). この意味において, ME は労働の分割問題を自然に取り入れた実装になっており, 同じ成績を示すのであれば最も自然なモデルだと言うことができる.

## 5.6 考察

意味層を介した間接経路のネットワークでは, 例外語に特化した局在化した領域分割がなされていると考えれば, 労働の分割問題が説明できる. 分散パラメータ  $\sigma^2 \rightarrow 0$  とすれば領域はディラックのデルタ関数に等しくなり, この各々の低頻度例外語ごとに限局された極限では, その単語のみに応答するルックアップテーブルと同一視できる. すなわち ME モデルの観点からトライアングルモデルを解釈し直すとすれば, なぜ低頻度例外語は意味経路を通るのかを見通しよく説明することができる. ME により, 二重経路モデルにおけるルックアップテーブルも, トライアングルモデルにおける労働の分割問題も統一的に記述できる. この意味において, 二重経路モデルとトライアングルモデルの間に本質的な違いはない. 両者の間に存在する違いとは, 分散パラメータ  $\sigma^2$  による入力空間の領域分割の大きさという量的な問題に帰結され, 本質的には二重経路モデルとトライアングルモデルは同じ ME という, より一般的な範疇のニューラルネットワークモデルとして同一視できるのである.

# 6 大脳基底核と皮質領野の相互作用を仮定したモデルによるリハビリテーションのシミュレーション

## 6.1 はじめに

本章では失語症の回復過程に関するニューラルネットワークモデルを提案し, 言語治療の技法に理論的な根拠を与えることを目的とした.

失語症の言語機能改善のアプローチには次の 3 つの方法がある. 1. 再活性化: 通常の情報処理過程の強化を図るもの, 2. 再編成: 活性化できる情報が限られている場合, 異なる情報処理過程を強化して同じ言語出力に至るようにするもの, 3. 再学習: 活用できる情報が失われてしまった場合, 新たな情報の獲得を行なうことで機能の改善を図るもの. 本章では, これら 3 つのリハビリテーション技法を用いたとき, 失語症患者の内部でどのような変化が起こっているのかをニューラルネットワークを用いて表現することを目指した.

脳における言語の処理には 3 つの機構が存在する. 単語が表す概念を蓄える概念系, 単語・文章生成系, および, この 2 つを繋ぐ媒介系の 3 種類である. しかし, この 3 種類の系に加えて, 他の感覚入力からの経路を仮定しな

ければ、再編成や再学習をモデル化することはできないと考える。

他の感覚入力から言語産出に至る経路の候補としては大脳基底核が挙げられる。概念系、単語・文章生成系、媒介系の皮質上の領野は、運動野や運動前野と直接に結合し、また、左大脳基底核や左視床前部の核を介して間接的にも結合している。この二重の経路の存在は、失語症患者のリハビリテーションにおいて、皮質と皮質下の両経路を通じて治療が行なわれることに対応すると思われる。すなわち新たに単語を学習するときには、その意味や他の感覚モダリティとの相互結合を介して新たな回路が形成されると考えられる。

ニューラルネットワークの観点から上記の3つのリハビリテーションを考えれば、1. 再活性化の場合は、概念系から単語・文章生成系への経路の再訓練とみなすことができる。一方、2. の再編成は、音韻手がかりや絵画・文字マッチングによるリハビリテーションと対応させて考えることができ、大脳基底核経由で学習が進むと考えられる。3. の再学習はリハビリテーションによって大脳基底核内部での自己組織化による応答の変化として考えることができよう。

## 6.2 リハビリテーション

モデルに学習を行なわせ、全ての刺激に対して正しく応答するようになるまで学習を繰り返した。この学習終了時のネットワークの結合強度を健常者のモデルと見なすことにした。そしてニューラルネットワークの一部のを破壊することによって失語症をモデル上で表現した。

ニューラルネットワーク上での脳損傷のシミュレーションを実施し、破壊後残されたユニットのみを用いてニューラルネットワークを再訓練した。すなわち残されたユニットのみを用いて最初の訓練時と同じ学習方式で学習を行なわせた。

前述のとおり失語症の治療には3つの方法がある。本章で提案したモデルに対応させて考えれば、1. 再活性化の場合は、概念系から単語・文章生成系への経路の再学習とみなすことができる。一方、2. の再編成は、音韻手がかりや絵画・文字マッチングによるリハビリテーションと対応させて考えることができ、大脳基底核経由で学習が進むと考えられる。3. の再学習は dead unit を利用した再学習で、リハビリテーションによって大脳基底核内部での自己組織化による応答の変化として考えることができよう。

## 6.3 考察

本モデルを用いれば、リハビリテーション訓練中の失語症患者の内部でどのような変化が起こっているのかを説明でき、リハビリテーション技法である、再活性化、再編成、および再学習において、媒介系、および大脳基底核

のユニットの再構成，再学習を視覚化して表現することが可能であることを示した。本モデルによって言語治療の技法に理論的な根拠を与えることができるのではないかと考えている。

## 第III部 総合考察

### 7 ニューラルネットワークを神経心理学へ適用する意味について

#### 7.1 ニューラルネットワークモデルと症例研究

コンピュータ上では実際の患者には要求できないような極端な要求や課題を課すことができ、その結果得られるデータも現象の理解に役立つ。こうして得られた知識や予測をもう一度症例の検討に戻って考えてみる、という研究手法がこの分野の研究を一步前に進める力になりうるだろう。モデルだけが一人歩きした測定無き理論 (theory without measurement) も現象 (症例) の記述のみに終始した理論無き測定 (measurement without theory) も、どちらも研究としては不完全である。このうちの理論的考察を行う際に、推論の道具としてニューラルネットワークを用いることが本論文で主張した神経心理学に対するニューラルネットワークの適用である。ニューラルネットワークモデルによる脳の構成論的研究と、実際の症例研究とは車の両輪をなしていなければならないと考える。

#### 7.2 脳の構成論的研究の意義

本論文では、人間の認知機能とニューラルネットワークモデル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によって、神経心理学に対して新しい視点を提供できることを示してきた。その適用範囲は、神経心理学の伝統的な推論方法である二重乖離の原則に基づく局在性仮説への批判から、具体的なリハビリテーション技法に対する理論的意味付けまで、多くの概念を説明することが可能である。

第一部で示してきたように、ニューラルネットワークモデルによる推論を行なうことで従来の神経心理学的症例分類学をみなおす必要があるのではないかと思える。ニューラルネットワークモデルを用いた脳の構成論的研究の意義はまさにこの点にあると思われる。

ニューラルネットワークモデルの記述力の高さによって、5章で見てきたように、論争の一部に結着をつけるような、より高い視点から現象を観察できるようになることもニューラルネットワークモデルを用いることの大きな利点である。エキスパート混合ネットワークを導入することで、失読症のモデルである二重経路モデルとトライアングルモデルにおける直接経路と間接経路の処理の違いに関する論争は、より一般的な読みのエキスパート混合ネッ

トワークの一形態にすぎないということが導き出せる。二重経路モデルにおけるルックアップテーブルの存在も、トライアングルモデルにおける労働の分割問題も、エキスパート混合ネットワークによる領域分割の仕方として記述可能である。すなわちエキスパート混合ネットワークにより、この論争に結着をつけることができると考えられる。

### 7.3 人工脳損傷の意義

ニューラルネットワークを用いた人工脳損傷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。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であるから、人工脳損傷研究は、神経心理学に対して強力な道具を提供していると言えよう。例えば4章に示した2段階読字過程モデルでは視覚性障害仮説と離断仮説によって生じる障害の程度が量的に異なることを予測している。また、例えば6章に示したように障害の程度を変化させて、どのリハビリテーション技法が効率良く治療を促進するかを予測することができる。こうした考え方は人工脳損傷とリハビリテーションによる回復との関係を論じることができるようになったのもニューラルネットワークによるシミュレーションの大きな効用であろう。

### 7.4 モデルに満たすべき要件

モデルが、説明可能性や妥当性を有し説得力を持った神経心理学のツールとなるためには、モデルの計算論的意味付けが可能でなければならない。個々のモデルにおいては、1.4節で述べた要件が満たされている必要がある。そうでなければ、モデルが神経心理学に貢献することはできないだろう。

## 8 背景となる数学的な基礎の整備の必要性

ニューラルネットワークを用いた神経心理学モデルの検討には計算論的研究が不可欠である。5章で示したように個々の症例を説明できるだけでなく、複数のモデルをその特殊な場合として含むような一般的なモデルの構築を目指す努力をすべきである。そのためには単にシミュレーションだけではなく計算論のレベルの考察が重要になってくるだろう。すなわち第二部でも示したようなモデルの数学的な性質を検討する努力がなされるべきである。



## 引用文献

- Coltheart, M., Curtis, B., Atkins, P. & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589–608.
- Coltheart, M. & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1197–1211.
- Duda, R. O., Hart, P. E. & Stork, D. G. (2001). *Pattern Classification, 2/Edition*. John Wiley and sons. パターン識別, 尾上守夫監訳, 2001, 新技術コミュニケーションズ.
- Farah, M. J. (1990). *Visual Agnosia*. MIT press, Cambridge, MA.
- Farah, M. J. (1994). Neuropsychological inference with an interactive brain: A critique of the locality assumption. *Behavioral and Brain Sciences*, *17*, 43–104.
- Farah, M. J. & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, *120*(4), 339–357.
- Geschwind, N. (1965). Disconnexion syndromes in animals and man. *Brain*, *88*, 17–294. (河内十郎訳、「高次脳機能の基礎」、(1984)、新曜社).
- Jacobs, R. A., Jordan, M. I., Nowlan, S. J. & Hinton, G. E. (1991). Adaptive mixtures of local experts. *Neural Computation*, *3*, 79–87.
- Jordan, M. I. & Jacobs, R. A. (1994). Hierarchical mixtures of experts and the em algorithm. *Neural Computation*, *6*, 181–214.
- Plaut, D. C. (2001). A connectionist approach to word reading and acquired dyslexia: Extension to sequential processing. In M. H. Christiansen & N. Charter (Eds.), *Connectionist Psycholinguistics* chapter 8, (pp. 244–278). Westport, CT: Ablex Publishing.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S. & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.

- Seidenberg, M. S. & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523–568.
- Tippett, L. J. & Farah, M. J. (1998). Parallel distributed processing models in alzheimer's disease. In R. W. Parks, D. S. Levine & D. L. Long (Eds.), *Fundamentals of Neural Network Modeling: Neuropsychology and Cognitive Neuroscience* chapter 17. MIT press.