

THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN

Introduction

最終的に、知覚認識、一般化、想起、思考に関する高等生物の能力を理解するためには、まず 3 つの基本的な質問に対する答えが必要です。

1. 物理的な世界に関する情報は、生物系によってどのように感知または検出されますか？
2. 情報はどのような形式で保存または記憶されますか？
3. ストレージまたはメモリに含まれる情報は、認識と動作にどのように影響しますか？

1 番目の質問は、神経生理学の分野であり、かなりの理解が達成された唯一の質問です。この記事は主に 2 番目と 3 番目の質問に関係しますが、これらは依然として膨大な憶測にさらされており、神経生理学によって現在提供されているいくつかの関連する事実はまだ許容可能な理論に統合されていません。

2 番目の質問に関しては、2 つの代替の立場が維持されています。1 つ目は、感覚情報の保存は、コード化された表現または画像の形式であり、感覚刺激と保存されたパターンの間の 1 対 1 のマッピングのようなものであることを示唆しています。この仮説によれば、神経系のコードまたは「配線図」を理解していれば、原則として、彼らが持っている「記憶の痕跡」から元の感覚パターンを再構築することにより、生物が覚えていることを正確に発見できるはずです。写真のネガを現像したり、デジタルコンピューターの「メモリ」の電荷パターンを変換したりする可能性があります。この仮説は、そのシンプルさとわかりやすさの点で魅力的であり、コード化された表現メモリの概念を中心に理論的脳モデルの大きなファミリーが発達しました（2、3、9、14）。イギリスの経験主義の伝統に由来する代替的なアプローチは、刺激のイメージが実際には全く記録されることはないかもしれない、と中枢神経系は単に保持が活動の中心地間の新しい接続、または経路の形を取る複雑なスイッチングネットワークとして機能することを推測に賭けています。この立場の最近の開発の多く（ヘブの「細胞集合」、およびハルの「皮質の予測目標反応」など）では、刺激に関連する「反応」は、CNS(中枢神経：Central Nervous System)自体に完全に含まれている場合があります。この場合の反応は、アクションというよりも「アイデア」を表しています。このアプローチの重要な特徴は、後の再構築を可能にするコードに従って、刺激のメモリへの単純なマッピングが決していないことです。保持される情報は、何らかの形で特定の応答の設定として保存する必要があります。つまり、情報は位置特異的な

表現ではなく接続または関連付けに含まれています。（このプレゼンテーションの残りの部分では、応答という用語は、外部から検出可能な筋肉活動を伴う場合と伴わない場合がある、生物の識別可能な状態を意味すると理解されるべきです。中枢神経系における細胞のいくつかの核の活性化、例えば、この定義に従って、応答を構成できます。）

情報保持の方法についてのこれら二つの立場に対応して、第三の問題である記憶された情報が現在の活動にどのような影響を及ぼすかについては、二つの仮説が存在する。「コード化されたメモリ理論」は、刺激の認識には、現在の刺激が以前に見たことがあるかどうかを判断し、生物の適切な反応を決定するために、記憶の内容と入ってきた感覚パターンとの照合や系統的な比較が必要であると結論せざるを得ないとしている。一方、経験主義の伝統の理論家は、本質的に 3 番目の質問への回答と 2 番目の質問への回答を組み合わせています。保存された情報は、新しい接続、または神経系の伝達経路の形を取る（または新しい接続と機能的に同等の条件の作成）、新しい刺激は、それらの認識や識別のための別のプロセスを必要とせずに、自動的に適切な応答を活性化し、作成されたこれらの新しい経路を利用することになります。

ここで提示される理論は、これらの質問に関して経験主義者、または「コネクショニスト」の立場をとります。この理論は、パーセプトロンと呼ばれる架空の神経系または機械のために開発されました。パーセプトロンは、特定の生物を保持する特別な、しばしば未知の条件に深く深く入り込むことなく、一般にインテリジェントシステムの基本的な特性の一部を示すように設計されています。パーセプトロンと生物学的システムの間のアナロジーは読者にすぐに明らかになるはずです。

ここ数十年の間に、記号論理、デジタルコンピュータ、スイッチング理論の発展は、ニューロンとコンピュータが構成する単純なオンオフユニットとの機能的な類似性を多くの理論家に印象づけ、高度に複雑な論理機能をそのような要素で表現するために必要な分析手法を提供してきました。その結果、刺激のシーケンスに反応して特定のアルゴリズム（「想起」、刺激の比較、変換、および様々な種類の分析を表す）を実行するための論理的な仕掛けに過ぎない脳モデルが大量に生み出されてきた（例：Rashevsky (14)、McCulloch (10)、McCulloch & Pitts (11)、Culbertson (2)、Kleene (8)、および Minsky (13)）。Ashby (1) や von Neumann (17, 18) のような比較的少数の理論家は、多くのランダムな接続を含む不完全なニューラルネットワークが、どのようにして理想化された配線図によって表現されるかもしれない機能を実行できるようにすることができるかという問題に関心を持っていました。残念ながら、記号論理やブール代数の言語は、このような研究にはあまり適していません。総体的な組織しか特徴づけられず、正確な構造が不明なシステムにおける事象の数学的解析に適した言語の必要性から、著者は現在のモデルを記号論理ではなく確率論の観点から定式化することになった。

上述の理論家たちは、知覚や想起などの機能が、どのようにして決定論的な物理システムによって達成されるのかという問題に主に関心を寄せていたが、それは脳によって実際にどのように行われるのかという問題ではなかった。これまでに作られたモデルはすべて、いくつかの重要な点で失敗している（等電位性の欠如、神経経済性の欠如、接続や同期要件の過剰な特異性、細胞の発火に十分な刺激

の非現実的な特異性、神経学的な相関関係が知られていない変数や機能的特徴の仮定、など）生物学的システムに対応している。この一連のアプローチの支持者は、上記の欠点を排除するために、さまざまな種類の物理システムがどのようにして刺激を認識および認識し、または他の脳のような機能を実行することができるかが示されると、既存の原理の改良または変更のみで、より現実的な神経系の働きを理解するとしている。一方、筆者は、これらの欠点は、すでに提案されている原理の単なる改良や改良では、生物学的知性を説明することができないようなものであり、原理の違いが明確に示されているという立場をとっている。ここにまとめられる統計的分離可能性の理論（Cf.15）は、これらの難点のすべてを原理的に解決するものであるように思われる。

これらの理論家-ヘブ(7)、ミルナー(12)、エクレス(4)、ハイエク(6)-は、生物学的神経系と自然環境におけるその活動に、形式的に「類似の機械」ではなく、より直接的に関心を持ってきたが、一般的には定式化の正確さに欠け、その分析の厳密さからは程遠いため、彼らが記述したシステムが現実的な神経系で実際に機能しうるかどうか、必要かつ十分な条件は何かを評価することはしばしば困難である。ここでもネットワーク解析者のブール代数に匹敵する解析言語がないことが大きな障害の一つとなっている。このグループの貢献は、それ自体が完成した理論体系としてではなく、何を探し、何を調査すべきかを示唆するものとして考えられるべきであろう。このような観点から見ると、次の理論の観点から、最も示唆に富む仕事は、Hebb と Hayek の仕事である。

Hebb (7)、Hayek (6)、Uttley (16)、Ashby (1) が特にパーセプトロンの理論の基礎としている立場は、次のような仮定で要約することができる。

1. 学習と認識に関わる神経系の物理的な接続は、ある生物から別の生物まで同一ではありません。誕生時には、最も重要なネットワークの構築は、遺伝的制約の最小数を条件に、大部分がランダムである。
2. 神経活動が一定期間続くと、ある細胞のセットに適用された刺激が他のセットで反応を引き起こす確率は、ニューロン自体の比較的長期的な変化のために変化する可能性があります。
3. 刺激の大量のサンプルへの暴露を通じて、最も「類似」しているもの（ある意味では、特定の物理システムで定義する必要がある）は、同じセットの応答細胞への経路を形成する傾向があります。著しく「異なる」ものは、異なるセットの応答セルへの接続を発達させる傾向があります。
4. ポジティブおよび/またはネガティブな強化（またはこの機能を果たす刺激）の適用は、現在進行中の接続の形成を促進または妨害する可能性があります。
5. そのようなシステムにおける類似性は、神経系のあるレベルで、同じセットの細胞を活性化する同様の刺激の傾向によって表されます。類似性は、特定の形式的または幾何学的なクラスの刺激に必要な属性ではありませんが、特定の環境との相互作用を通じて進化する知覚システムの物理的組織に依存します。システムの構造、および刺激環境の生態学は、知覚世界が分割される「もの」のクラスに影響を及ぼし、主に決定します。

2. The Original of a Perceptron

典型的なフォトパーセプトロン（光パターンを刺激として応答するパーセプトロン）の構成を図1に示す。その構成のルールは以下の通りである。

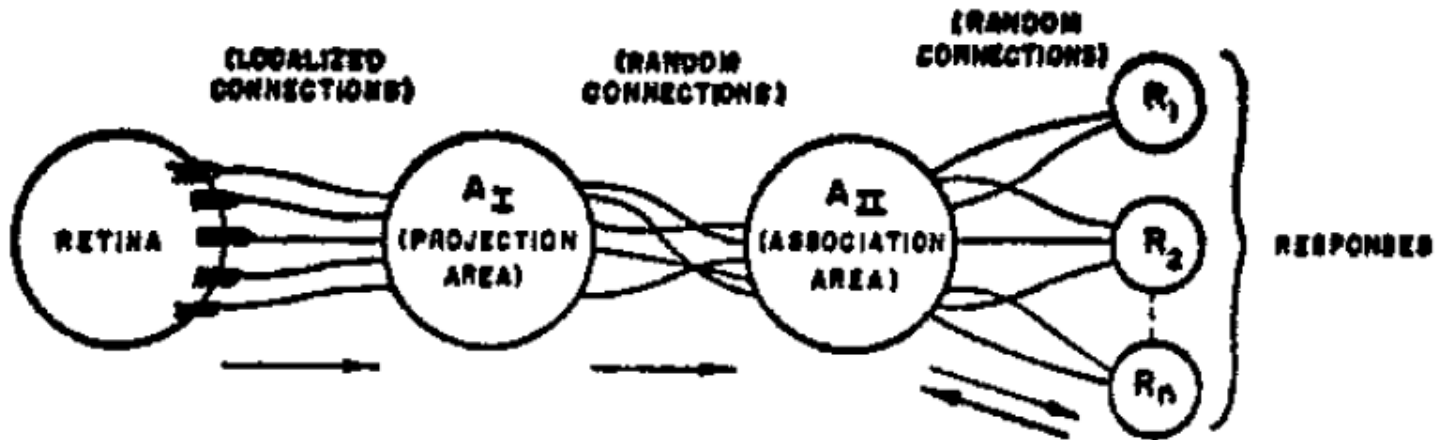


図1. パーセプトロンの構成

1. 刺激は感覚ユニット（S-points）の網膜に衝突するが、これらの刺激は、いくつかのモデルでは All or Nothing ベースで応答すると仮定され、他のモデルでは刺激強度に比例したパルス振幅または周波数で応答すると仮定されている。
2. インパルスは、「投影領域」(A_I)内の一連の連想細胞(A ユニット)に送信されます。この投影領域は、網膜が連想領域 (A_{II}) に直接接続されているいくつかのモデルでは省略されることがあります。特定の A ユニットにインパルスを送信する S-points のセットは、その A ユニットの起点と呼ばれます。これらの起点は、A ユニットへの影響において、興奮性または抑制性のいずれかです。興奮性インパルスと抑制性インパルス強度の代数的合計が A ユニットの閾値(θ)と同じかそれ以上であれば、A ユニットは発火し、再び All or Nothing ベースで(または、ここでは考慮しないいくつかのモデルでは、受信したインパルスの純値に依存する周波数で)発火する。投影領域の A ユニットの起点は、各 A ユニットに対応するいくつかの中心点について、クラスタ化されているか、または焦点化されている傾向があります。問題の A ユニットの中心点からの網膜距離が増加するにつれて、原点の数は指数関数的に減少します。(このような分布は、生理学的証拠によって支持されているようであり、輪郭検出において重要な機能的目的を果たしている)。
3. 投影領域と連想領域 (A_n) の間では、接続はランダムであると仮定される。すなわち、 A_n 集合内の各 A ユニットは、 A_I 集合内の原点からある程度の本数のファイバを受け取るが、これらの原点は投影領域内にランダムに散らばっている。接続分布を除けば、 A_n ユニットは A_I ユニットと同一であり、同様の条件で応答する。

4. 「応答」 R_1, R_2, \dots, R_n はセル（またはセルのセット）であり、A ユニットとほぼ同じ方法で応答します。各応答には、通常、An セット内にランダムに配置された多数の起点があります。

特定の応答にインパルスを送信する A ユニットのセットは、その応答のソースセットと呼ばれます。

（応答の送信元セットは、A システムの送信元セットと同じです。）図 1 の矢印は、ネットワークを介した送信方向を示しています。An まではすべての接続が転送され、フィードバックはありません。An と R ユニットの間の最後の接続セットに到達すると、接続は両方向で確立されます。パーセプトロンのほとんどのモデルで、フィードバック接続を管理するルールは、次のいずれかの選択肢になります。

- (a) 各応答には、独自のソースセット内のセルへの興奮性フィードバック接続があります。
- (b) 各応答は、それ自身のソースセットの相補体への抑制的フィードバック接続を有する（すなわち、それは、それに送信しない任意のアソシエーション細胞での活動を禁止する傾向がある）。

R ユニットはそれぞれのソースセットと同じ皮質領域にある可能性があり、適切なソースセットの R ユニットと A ユニット間の相互励起が高い確率で行われるため、これらのルールの最初のルールは解剖学的にもっともらしいようです。ただし、代替ルール（6）を使用すると、システムをより簡単に分析できるようになるため、ここで評価するシステムのほとんどで仮定されます。

図2 は、統計的分離可能性の理論に簡単に入ることができる単純化されたパーセプトロンの構成を示しています。

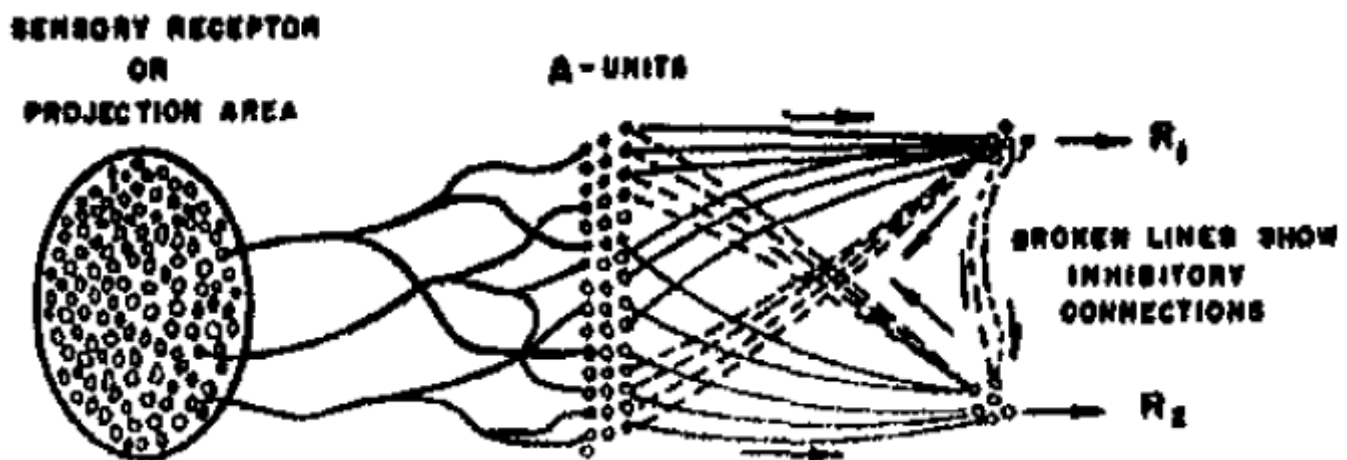


図2.A. 単純なパーセプトロンにおける接続の模式表現。

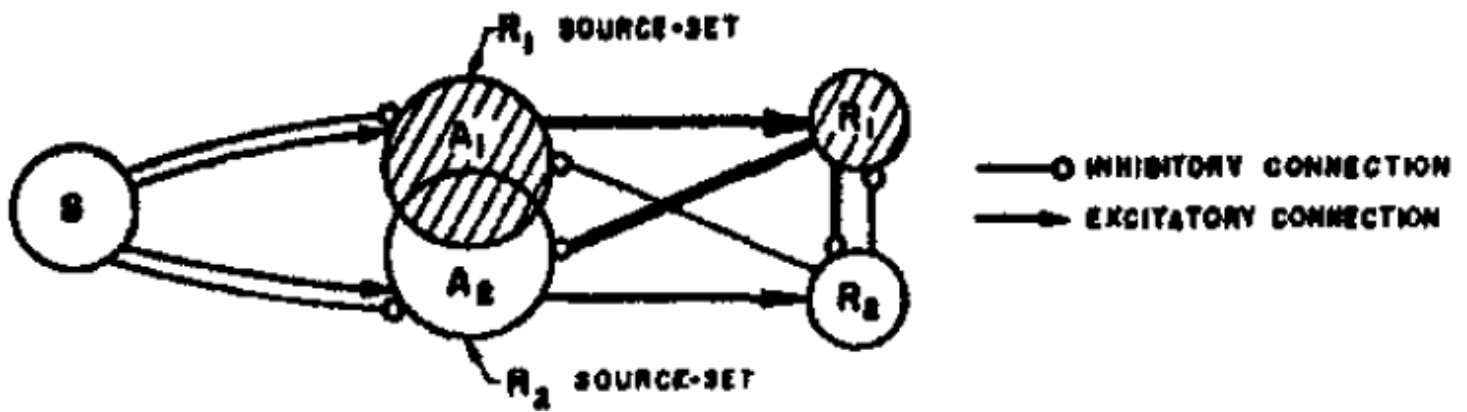


図2.B. 同じパーセプトロンのベン図（網掛けはR1応答のアクティブセットを示す）。

この単純化されたモデルについて理論が展開された後、図1のシステムの利点について議論することができます。図2に示すフィードバック接続は抑制的であり、それが元になる応答のソースセットの補数に行きます。ここに示されているシステムは、3つの段階のみを持ち、第1の連想段階は除去されている。各Aユニットは、網膜内にランダムに配置された起点のセットを有する。このようなシステムは、輪郭や輪郭の類似性ではなく、刺激の一致領域に基づいて類似性の概念を形成する。このようなシステムは、多くの弁別実験では不利であるが、その能力は非常に印象的である。図2のシステムは2つの応答しかないが、その数に制限はないことは明らかである。

このように構成されたシステムにおける応答は、相互に排他的である。R1が発生すると、それはR2を阻害する傾向があり、R2のソースセットも阻害する。同様に、もしR2が発生すれば、それはR1を阻害する傾向がある。1つのソースセット内のすべてのAユニットから受信される総インパルスが、代替(拮抗)応答によって受信されるインパルスよりも強いのか、またはより頻繁である場合、最初の応答は、他のものよりも有利になる傾向があり、発生するものとなる。もしそのようなシステムが学習が可能であるならば、あるクラスの刺激がR2ソースセットよりもR1ソースセットでより強い衝動を呼び起こす傾向がある一方で、別の(異質な)クラスの刺激がR1ソースセットよりもR2ソースセットでより強い衝動を呼び起こす傾向があるような方法で、Aユニットまたはそれらの接続を変更することが可能でなければならない。

各Aユニットによって送達されるインパルスは、振幅、周波数、遅延、または送信完了の確率のいずれかである値Vによって特徴付けられると仮定する。Aユニットが高い値を持つ場合、その出力インパルスは、低い値を持つAユニットからのインパルスよりも、より効果的、より強力、またはそれらのエンドバルブに到達する可能性が高いと考えられます。A-ユニットの値は、おそらく細胞と細胞膜の代謝状態に応じて、かなり安定した特性であると考えられていますが、それは絶対的に一定ではありません。一般的には、活動期間中は細胞の値が増加する傾向があるが、活動がないと値は(いくつかのモデルでは)減衰することがあると仮定されている。最も興味深いモデルは、細胞が代謝物質を求めて競争すると仮定されたモデルであり、活動的な細胞は、活動的でない細胞を犠牲にして得をする。このようなシステムでは、活動がない場合、すべての細胞は比較的一定の状態に留まる傾向

があり、（活動に関係なく）システム全体で見た場合、システムの正味の値は常に一定のままである。値のダイナミクスが異なる 3 種類のシステムを定量的に調査しました。それらの主な論理的特徴を表 1 に示します。アルファシステムでは、アクティブセルはインパルスごとに単純に値を増加させ、この増加を無期限に保持します。ベータシステムでは、各ソースセットには一定のゲイン率を与えられており、そのインクリメントはソースセットのセルの活性度に比例して各セルに配分される。ガンマ・システムでは、アクティブなセルは、そのソース・セットの非アクティブなセルを犠牲にして値を増加させるので、ソース・セットの合計値は常に一定である。

	α - System(Gain System)	β - System(Feed System)	γ - System
補強あたりのソースセットの総バリューゲイン	N_{ar}	K	0
1 単位時間の間アクティブな A ユニットの ΔV	+1	K / N_{ar}	+1
ドミナントセット外の不活性 A ユニットののための ΔV	0	K / N_{Ar}	0
ドミナントセットの不活性 A ユニットののための ΔV	0	0	$\frac{-N_{ar}}{N_{Ar}-N_{ar}}$
A システムの平均値	増援数に応じて増加		