

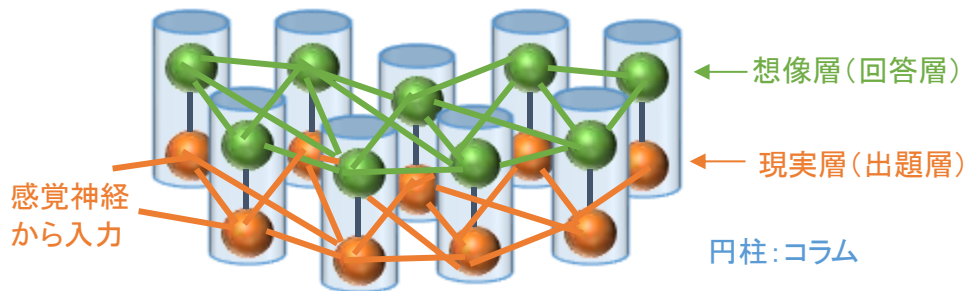
# 二重意識ニューラルネットワーク

© RUSAGI 汎用人工知能研究所 <https://rusagi.com/>

## 目次

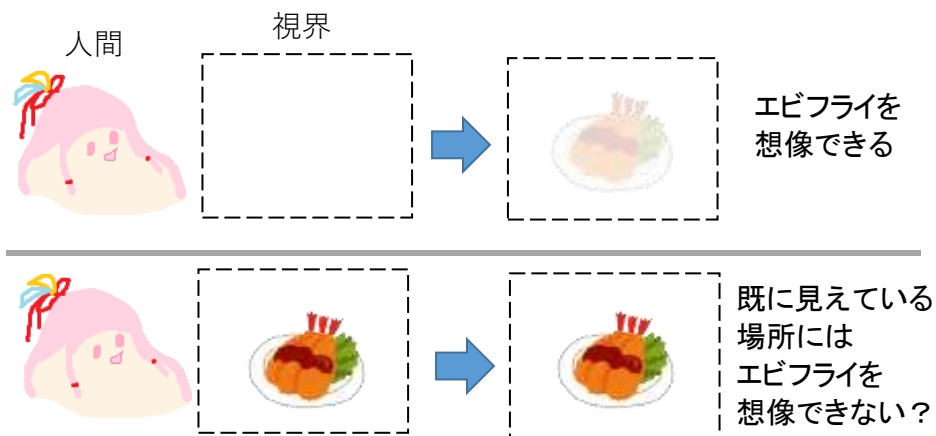
1. 二重意識ニューラルネットワークの構造.....	2
2. 現実と想像、意識の二重性 .....	2
3. 意識と選択 .....	3
4. 二重意識ニューラルネットワークの目的と接続.....	4
5. 学習時の処理イメージ .....	5
6. 意識の集中（注意） .....	7
7. 他の研究との関係.....	7
8. あとがき .....	8

## 1. 二重意識ニューラルネットワークの構造



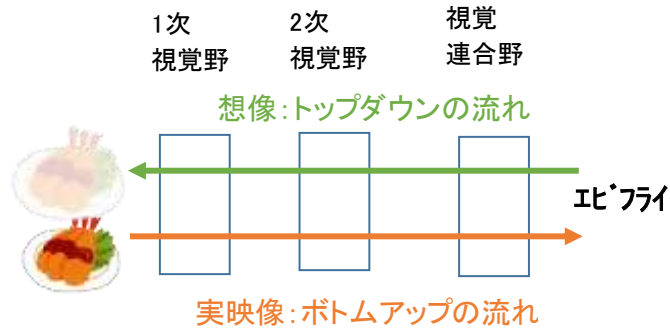
二重意識ニューラルネットワークでは、2つのネットワークが同数のノードを持ち、双方のノードが1対1で繋がっているとす。大脳に対応させると、コラムの上半分と下半分で別々のネットワークがあり、コラム内部で繋がっている。片方は、視覚等から入ってきた現実の情報が流れ、もう片方は、想像している情報が流れる。これを現実層と想像層と呼ぶ。出題する層と、回答する層という役割を持つ。

## 2. 現実と想像、意識の二重性



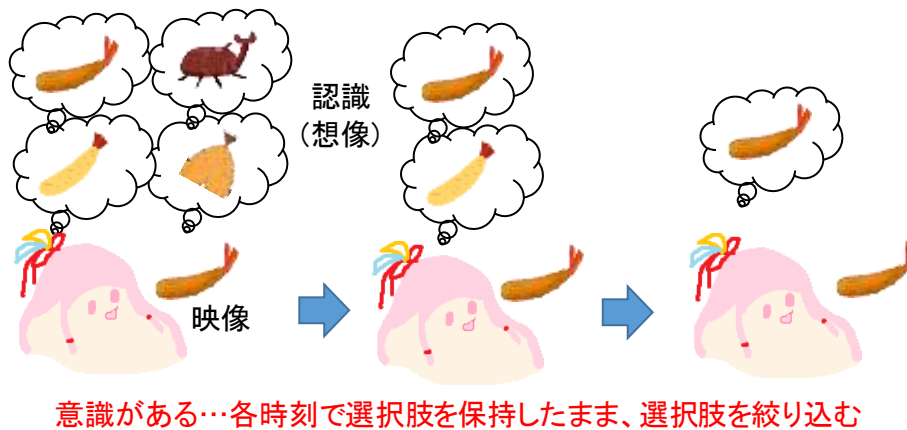
**想像できないのではなく、認識した = 既にその場所に想像している**

人は、何もない場所に、エビフライを想像できる。しかし、エビフライが見えているとき、全く同じ場所に、全く同じエビフライを想像することはできない。これは、想像できないのではなく、エビフライを見た時点で、既にその場所にエビフライを想像しているとも解釈できる。正確には、見えた瞬間はそれが何だか認識できておらず、認識したのと同時にエビフライを想像したと考えられる。



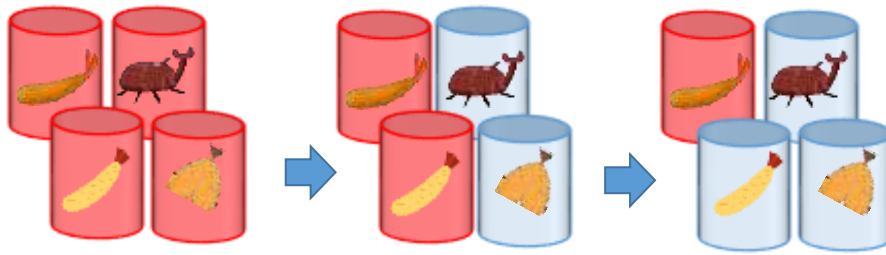
このときに脳内では情報が2方向に流れる。目からの映像の各ピクセルの色から、線の傾きを認識し、さらに図形を認識し、エビフライを認識するというのがボトムアップの流れである。一方、想像するために、図形→傾き→画素というトップダウンの流れが存在する。認識することとは、想像という形で、脳内で外界のモデルを造ることと解釈できる。外界からの情報と、想像の情報が一致すると、それは認識できたといえる。脳の状態は、現実と想像の差が小さくなる方向に遷移し、差が無くなると平衡状態になると考えられる。

### 3. 意識と選択



「選択肢を絞る込むこと」という知能の定義から、エビフライを見たとき、脳の状態は、エビフライを想像している状態と、カブトムシ等を想像している状態等の選択肢があるが、時間とともに徐々に前者の状態を選択するといえる。意識が有るということは、各時刻で選択肢を保持したまま、選択肢を絞り込むことを意味する。

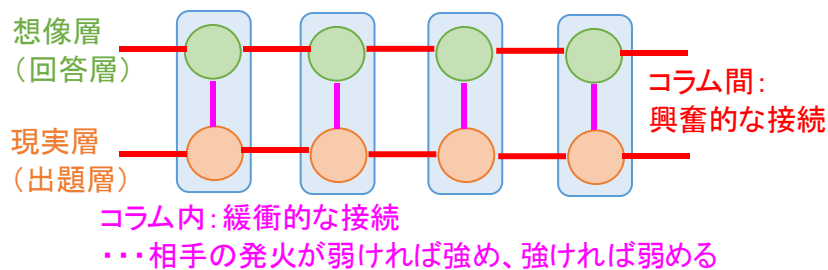
次の時刻の各コラムの発火頻度 =  $f$  (ある時刻の各コラムの発火頻度)  
**脳は「次の時刻の各コラムの発火頻度を選択」**  
 ……各時刻で値を1つに絞り込み……選択肢は保持されない(意識なし)



**脳は「次の時刻でどのコラムが発火するべきか選択」**  
 ……各時刻で、選択肢は保持される (意識あり)

このとき脳内で何が起こったか考える。大脳のコラム単位で考えると、ある時刻に、各コラムは発火頻度という値を持っている。脳のネットワークは、ある時刻の各コラムの発火頻度配列を入力として、次の時刻の発火頻度配列を出力する関数と解釈できる。しかし、脳が発火頻度を選択していると考え、各時刻では発火頻度が幾つであるべきかという選択肢を1つに絞り込んでいることになる。しかし、意識が存在するのなら、その定義から、選択肢は維持されなければならない。そこで、脳は「どのコラムの発火頻度が高くあるべきか」選択していると考え。たくさんのコラムが発火している状態が、まだ選択肢を絞り込んでいない状態で、徐々に発火するコラムが減る形で選択肢が絞り込まれる。興奮と抑制は非対称的である。

#### 4. 二重意識ニューラルネットワークの目的と接続



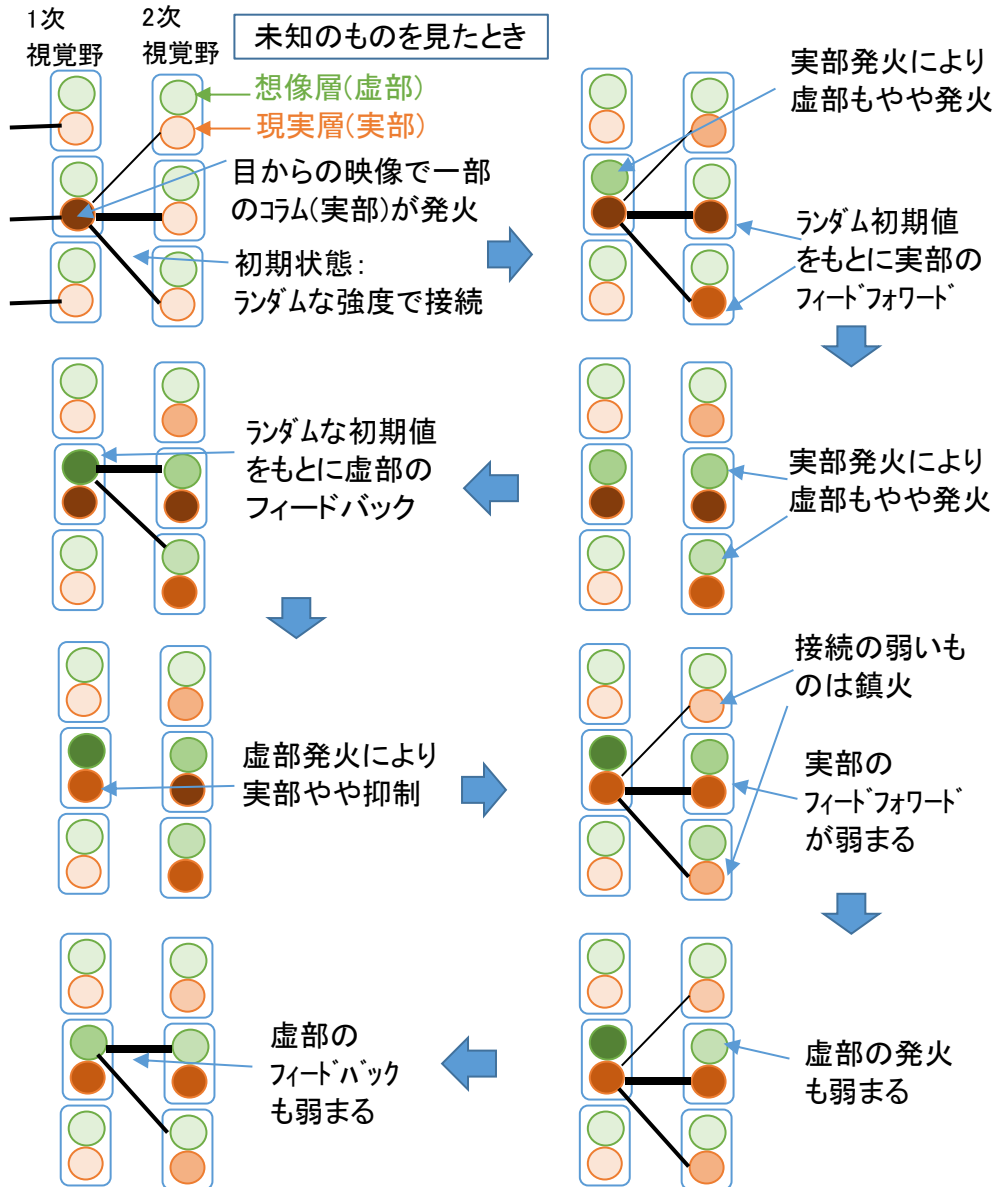
**目的: 2層の発火の同期化**

「発火すべきコラム」を選択するという形で意識を持たせるとして、その「発火すべき」の指針が必要となる。二重意識ニューラルネットワークでは、感覚神経から入力を受ける現実層と、想像層の対応コラム発火頻度の同期化を目的とする。ただし、全コラムの発火頻度が0または1の状態、同期が取れても意味がないため、発火頻度が中間的な状態で同期化するのを目的とする。

2つのネットワークそれぞれ、コラム間は興奮型の接続をする。それは連鎖的に情報を伝えるためである。コラム内では、二つのノードが互いに緩衝的に働く。すなわち、相手のノードの発火頻度が低いときは高め、高いときは低める。コラム間の接続の強さは、可塑的であり、それぞれ接続強度の値を持つ。全てのコラム間で接続している必要はなく、周辺のコラムにラ

ランダムな接続強度で接続した状態からスタートする。

## 5. 学習時の処理イメージ



**出題(サーチ): 上層の実部の広範囲のコラムへ発火を撒き散らす**

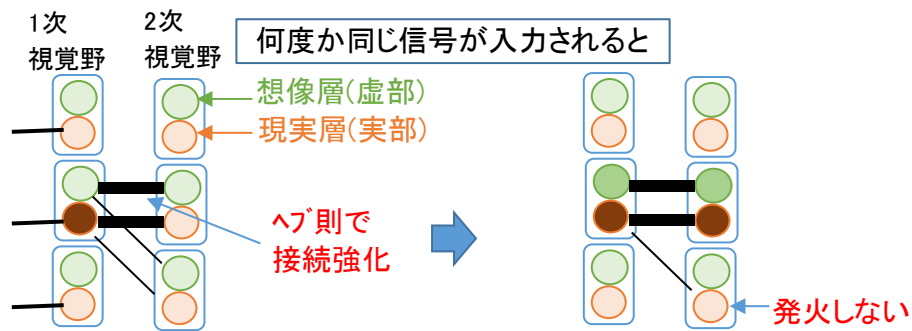
**回答: 上層の虚部からのフィードバックで抑制を受ける**

**抑制されるとフィードフォワードが弱まり、サーチが終わる。**

**フィードバックが強すぎるとフィードフォワードも弱められるため、一定値に落ち着く**

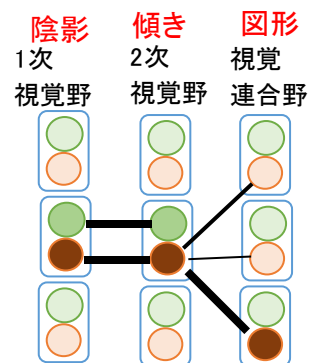
未知なものを見て、1次視覚野があるパターンで発火したとき、2次視覚野へ発火がフィードフォワードする。また、現実部(実部)が発火すると、コラム内の緩衝的接続で、想像部(虚部)もやや発火する。コラム間の接続は初期状態ではランダムだが、なかにはフィードバックするように接続しているものもある。虚部でフィードバックを受けると、緩衝的接続により、実部の発火はやや抑制される。入力映像は、実部へ出題され、上層の広範囲に発火を撒き散らしてサーチを行い、フィードバックでの抑制として回答されると、サーチは終わる。フィード

バックが強すぎればフィードフォワードが弱くなるため、ある一定のところで落ち着く。



何度か同じ信号が入力されると、ヘブ則で接続が強化される。  
 →フィードフォワードは、いきなり最大出力ではなく、徐々に立ち上がるが、  
 立ち上がりきる前にフィードバック抑制されるため、  
 上層の広範囲に発火を撒き散らさなくなる

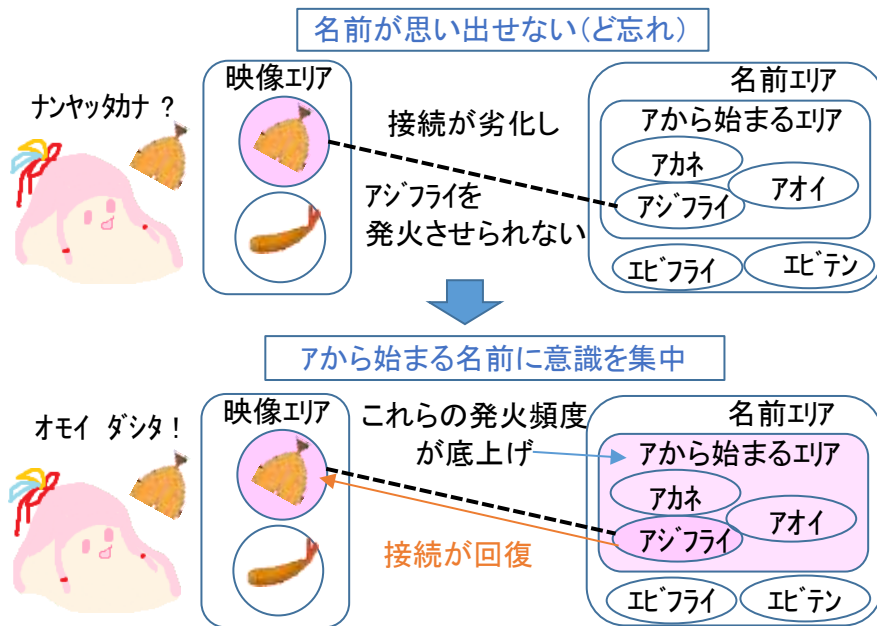
何度も同じ信号が入力されると、ヘブ則で接続が強化される。初めは特定の接続が強化されていない未学習のコラムのうち、初期値でたまたま接続が強かったコラムが選ばれるが、学習後は、そのコラムが選択的に発火する。また、サーチのフィードフォワードは、いきなり最大出力ではなく、徐々に立ち上がるが、強い接続があれば、立ち上がりきる前にフィードバック抑制されるため、広範囲に発火を撒き散らすことはなくなる。



どの層のコラムも同じように働くため、  
 ある次元の出題が、1つ上の次元で回答されると、  
 次は、そのさらに上の次元で回答(説明)しようとする

どの層のコラムも同じように働くため、ある次元の出題が、1つ上の次元で回答されると、次は、そのさらに上の次元で回答(説明)しようとする。つまり、あることを認識すると、次はさらに高次の難しいことを認識しようとする。

## 6. 意識の集中（注意）



**意識を集中：そのエリアの発火頻度が底上げされる(活性化)**  
**活性化されたエリア間では、フィードフォワード・バックループが形成され易い**

ある物の名前を学習しても、ど忘れで名前が思い出せないことがある。何らかの原因で接続が劣化し、その映像のコラムが発火しても、その名前のコラムが発火しない状態にと考えられる。そのとき、例えば「アから始まる名前」に意識を集中（注意）することで、連想しやすくなるということが起こる。このとき、「アから始まる名前」のコラムの発火頻度が底上げされて活性化していると考えられる。その中に思い出したかったものがあれば、発火し、接続が回復するだろう。意識を集中（注意）することは、そのエリアの発火頻度を底上げすることであり、活性化されたエリア間では、フィードフォワード・フィードバックのループが形成され易くなると考えられる。

## 7. 他の研究との関係

本研究は、脳の解剖結果も参考にしている。コラム間のフィードフォワード接続とフィードバック接続は非対称的であるため、それぞれの方向のネットワークを別々に考えることもできる。また、興奮型の接続と比べて、抑制型の接続は、コラム内であまり規則がない接続をしているため、緩衝的な作用をすると推測する。

二つのネットワークを戦わせるというのは、敵対的生成ネットワーク **Generative adversarial networks(GAN)**と似ている。GANはディープラーニングの最終層の出力結果をもとに戦うが、脳は、横方向に層が接続しているため最終層というものはない。映像に限らず、高次の概念でも戦う形にしようとするれば、各層で2つのネットワークが戦う形になり、本研究と近い形となる。

本研究は、予測符号化(predictive coding)とも似ている。それは、実情報と予測の誤差を上

層へ伝えるというもので、本質的にやろうとしていることは大差ない。ただ、差異を伝えるのではなく、2層に分けてしまった方が、より単純であり、心理的な処理に近いと感じる。また、2層構造なら、適当に繋がっている接続がヘブ則で強化されるだけでもそれなりに機能し、ロバスト性が優れる。より単純な仕組みの方が、本当の仕組みに近いと予想する。

## 8. あとがき

脳の処理は、解剖学的に詳しく分かっているわけではなく、不明な部分は予想しなければならない。しかし、知能や意識がどういったものか分かっていたら、それに合うように予想を立てることができる。本研究は、そのようなモデルの一つである。本研究ではプログラムとして実行できるまで詳細を詰めてはいない。それは、詳細を詰めようとすると数十パターンくらいアルゴリズムが考えられてしまうからである。人間レベルの知能さえ達成できれば良いのなら、人海戦術で、そういったアルゴリズム数十個を実験的に検証してしまうのが近道だと思われる。