

# 人工知能の原理と歴史

---

久木田水生

### ウェブで読めるもの

- 久木田水生「人工知能、ロボット、知性」(『社会と倫理』, 第28号、2013年11月、pp. 51-65)
- 久木田水生「人工知能の人工生命への接近」(『PROSPECTUS』, 10号、2007年12月、45-57)
- 久木田水生「帰納論理プログラミング」(『哲学論叢』, 第33号、2006年、103-113)

### ウェブで読めないもの

- 久木田水生「パースと人工知能」(『大航海』, 第6号、2006年、131-139, 新書館)
- 久木田水生、神崎宣次、佐々木拓、『ロボットからの倫理学入門』, 第1章「機械の中の道徳」(名古屋大学出版会、2017年)

# そもそも人工知能とは？

- ジョン・マッカーシーによれば「人間がそれをやれば「知能を持っている」と思われるような振る舞いをする人工的なシステム」。
- 初期の人工知能は特に論理的な推論や言語を使うことに焦点を当てていた。



ジョン・マッカーシー

「人工知能」という言葉を作った。

By "null0" (<https://www.flickr.com/photos/null0/272015955/>)

[CC BY-SA 2.0 (<https://creativecommons.org/licenses/by-sa/2.0/>)]

via Wikimedia Commons

## この定義の困難

- 何が「知的」と見なされるかがかなり曖昧。
- 人間がやっても大して知的と思われなくても、人工知能の歴史において非常に大きな達成と見なされることがある。
- 例：猫の写っている画像とそうでない画像を見分ける。ハンカチをたたむ。

## 普及すると人工知能と呼ばれなくなる？

- 最初は「人工知能」と呼ばれていたものが、普及すると人工知能とは呼ばれなくなる、ということがしばしば指摘される。
- 例えば論理学の定理証明や路線検索なども最初は人工知能と呼ばれていた。
- 機械に実現されてしまうと、「そんなものは本当の知能じゃないのだ」と言いたくなるのかもしれない。

Jerry Kaplan, *Artificial Intelligence: What Everyone Needs to Know*, Oxford University Press, 2016.

- 「知能」という言葉を使うのは無用な混乱のもと。
- 「機械が人間より優れた存在になる」とかいう危惧も、この名前によって掻き立てられているところがある。
- 人工知能の発展とは単なる「**絶え間ない自動化の進展**」である。(p. 17)

- 19 世紀末-1930 年代：記号論理学、計算理論の誕生
- 20 世紀半ば：コンピューターの誕生
- 1956 年：「人工知能」に関するダートマス会議
- 1950 年代：第一次人工知能ブーム。「論理的 AI」の時代。
- 1980 年代：第二次人工知能ブーム。「知識」、「身体化」、「コネクショニズム」などパラダイムの多様化。日本にもブームが波及。
- 2010 年代：第三次人工知能ブーム。「ビッグデータ」と「機械学習」の時代。インターネット上のビジネスに応用され大きな利益を生む。画像・映像処理、言語処理、ロボットなどと組み合わせられることで応用の可能性が広がる。それと同時に社会への影響が懸念されるようになる。

- 厳密に正しい推論についての研究。
- 推論とはいくつかの前提（根拠）から結論を導き出す思考のこと。
- 正しい推論：前提がすべて正しい時には結論も正しい。
- 例：三段論法

前提 1 すべての人間は死ぬ。

前提 2 ソクラテスは人間である。

---

結論 ソクラテスは死ぬ。

前提 1 すべての日本人は生魚を食べる。

前提 2 日本人のノーベル平和賞受賞者がいる。

---

結論 生魚を食べるノーベル平和賞受賞者がいる。

- 三段論法は古代ギリシャのアリストテレスによって研究され、中世まで論理学の中心的なテーマだった。
- より一般的な推論の原理が探求されるのは19世紀になってから。
- 広義の「論理学」は、より幅広い推論や科学的探究の方法一般を指していた。

- 論理的な推論を使えば、正しい前提からは必ず正しい結論が得られる。
- つまり確実に正しい前提だけから論理的推論を繰り返して得られた結論は正しいことが保証される。これが**証明する**ということ。
- 確実に正しいことが分かっていることだけから論理的推論だけを使えば、確実に正しい知識、すなわち真理が得られる。
- これが古代ギリシャにユークリッド幾何学において確立された数学の手法。
- 近代になりデカルトが哲学の方法論として重要視した。
- このように知識獲得の方法として論理的推論を重視する立場を**合理主義**という。

- 論理的に正しい推論一般を厳密に特徴づけることを目的に、19世紀末から発展し、大きな成果を上げる。
- 推論の構造を明示化するために抽象的な記号によって推論を表現する。
- 限定的な範囲で、少数の原理からすべての論理的推論を導出することに成功する。
- 数学の一部が論理学に還元されることを示す。→形式的数学、数学基礎論。
- それと同時に、数学には記号論理学の手法で厳密に特徴づけられない部分があることも明らかにした。

## 記号論理学における証明の例

公理 1 :  $X \rightarrow (Y \rightarrow X)$

公理 2 :  $(X \rightarrow (Y \rightarrow Z)) \rightarrow ((X \rightarrow Y) \rightarrow (X \rightarrow Z))$

公理 3 :  $(\neg X \rightarrow Y) \rightarrow ((\neg X \rightarrow \neg Y) \rightarrow X)$

推論規則 :  $X$  と  $X \rightarrow Y$  から  $Y$  を推論して良い.

$A \rightarrow A$  の証明

1.  $(A \rightarrow ((A \rightarrow A) \rightarrow A))$  (公理 1)
2.  $(A \rightarrow ((A \rightarrow A) \rightarrow A)) \rightarrow ((A \rightarrow (A \rightarrow A)) \rightarrow (A \rightarrow A))$  (公理 2)
3.  $(A \rightarrow (A \rightarrow A)) \rightarrow (A \rightarrow A)$  (1, 2 と推論規則)
4.  $A \rightarrow (A \rightarrow A)$  (公理 1)
5.  $A \rightarrow A$  (3, 4 と推論規則)

1. 0 は自然数.
2.  $x$  が自然数ならば  $sx$  も自然数.
3. 任意の自然数  $x, y$  について  $(sx = sy \rightarrow x = y)$
4. 任意の自然数  $x$  について  $sx \neq 0$
5. 任意の自然数  $x$  について  $x + 0 = x$
6. 任意の自然数  $x, y$  について  $x + sy = s(x + y)$
7. 0 が性質  $P$  を持ち、任意の自然数  $x$  が  $P$  を持つことから  $sx$  もまた  $P$  を持つことが示されるならば、すべての自然数が  $P$  を持つ.

## $x + y = y + x$ の証明

1.  $0 + 0 = 0$ . (公理 5)
2.  $0 + x = x$  と仮定する.
3.  $0 + sx = s(0 + x) = sx$ . (公理 6 と 2)
4. 任意の自然数  $x$  について  $(0 + x = x)$  とすると  $(0 + sx = sx)$  が言える. (2, 3 より)
5. すべての自然数  $x$  について  $0 + x = x$ . (1, 4 と公理 7 より)
6. すべての自然数  $x$  について  $0 + x = x + 0$ . (5 と公理 5 より)
7.  $sy + 0 = sy = s(y + 0)$ . (公理 5)
8.  $sy + x = s(y + x)$  と仮定する.
9.  $sy + sx = s(sy + x) = ss(y + x) = s(y + sx)$ . (公理 6 と 8)
10. 任意の自然数  $x$  について  $sy + x = s(y + x)$  とすると  $sy + sx = s(y + sx)$  が言える. (8, 9 より)
11. すべての自然数  $x$  について  $sy + x = s(y + x)$ . (7, 10 と公理 7 より)
12.  $y + x = x + y$  と仮定する.
13.  $sy + x = s(y + x) = s(x + y) = x + sy$ . (12 と 11 より)
14. 任意の自然数  $y$  について  $y + x = x + y$  とすると  $sy + x = x + sy$  が言える. (12, 13 より)
15. 任意の自然数  $y$  について  $x + y = y + x$  (6, 14 と公理 7 より)

# Intelligence とは？

- 推論や計算する
- 問題、パズルを解く
- 言葉を使う
- 計画を立てる。
- チェスなどのゲームをする
- 発見をする
- 創造をする
- などなど

- 1956年、ダートマス大学にて開催。
- ジョン・マッカーシー、マービン・ミンスキー、ネイサン・ロチェスター、クロード・シャノンなどが中心となる。
- 初めて「人工知能 (Artificial Intelligence)」という用語が大々的に使われた。

「本研究は、学習やその他のあらゆる知能の側面が原理的には、機械がそれをシミュレートできるくらい正確に記述できるという推測に基づいて進められる。どのようにすれば機械が言語を扱えるか、抽象を行い概念を形成できるか、現在は人間にしか扱えないような種類の問題を解けるか、そして自分自身を改良できるか。これらを発見するための試みが行われるだろう。ひと夏の間、慎重に選ばれた科学者のグループがともに取り組めば、これらの問題のいくつかににおいて重要な進展が成し遂げられると私たちは考えている。」

J. McCarthy, M. L. Minsky, N. Rochester, C.E. Shannon, "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence", August 31, 1955.

<http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>

- ダートマス会議で発表された、アレン・ニューウェル、ハーバート・サイモン、J・C・ショーが開発したコンピュータープログラム。
- 人間の論理的推論をシミュレートすることを意図して作られ、アルフレッド・ノース・ホワイトヘッドとバートランド・ラッセルの *Principia Mathematica* (『数学の原理』) の冒頭の 52 の定理のうち 38 を証明することができた。

- アレン・ニューウェルとハーバート・サイモンは「例えばコンピューターのような、物理的記号システムは知的な行動のための必要にして十分な手段を持っている」と主張した。
- A. Newell and H. Simon. “Computer science as empirical inquiry: Symbols and search.” *Communications of the ACM*, 19(3):113–126, 1976.
- この考え方は「物理的記号システム仮説」と呼ばれる。
- 初期の人工知能においてはこのようなパラダイムが支配的であった。
- このようなパラダイムを**記号的 AI**、**論理的 AI** と呼ぶ。

- コンピューターの理論的基礎を築いたアラン・チューリングは、振る舞い上人間と見分けがつかないほど流暢に人間とコミュニケーションができる機械は思考しているとみなすという考えを提唱した。
- Turing, Alan, “Computing Machinery and Intelligence”, *Mind*, LIX (236): 433–460, 1950.



画像はパブリックドメイン。

## 「チューリング・テスト」

- 上記の論文の中でチューリングは、現在「チューリング・テスト」として知られる、機械の思考についてのテストを提案した。
- そこでは被験者が人間またはコンピューターとチャットで会話をする。
- 被験者は相手が人間かコンピューターかを推測する。
- コンピューターが十分な頻度で被験者をだますことができるならば、そのコンピューターは人間と同等の思考を持っていると見なされる。

- これに対して哲学者のジョン・サールは、コンピューターはルールに従って記号を操作しているだけで、その**記号の意味を理解していない**から、本当に思考しているとは言えない、と反論した。
- その際にサールは「中国語の部屋」という有名な思考実験に訴えた。

- 部屋に中国語を知らない人間と、中国語の応答表が置いてある。
- 中国語で質問を書いた紙を部屋に入れると、中の人間は応答表を手掛かりに答えを紙に書いて出す。
- 外から見ている人間には、中の人間が中国語で支障なくやり取りができてるように見える。
- しかし中の人間は中国語を知らないのだから、自分が何を聞かれて何を答えているのかを理解していない。
- したがって中の人間は質問への応答に対応する思考を持っていない。

- スティーヴン・ハーナッドはサールの批判を受けて、「どうすれば人工智能が扱う記号が現実世界の事物と結びつきをもてるか」という問題を**記号接地問題**と名付けて、その解決に取り組んだ。
- Harnad, S. (1990) “The Symbol Grounding Problem.” Physica D 42: 335-346.
- それ以降、記号接地問題は人工智能、ロボット工学において重要なテーマの一つになっている。

- ハーナッド自身の提案は，人工知能に外界を認知し，そこから範疇的表象を形成する能力を持たせることによって，記号接地を実現させるというものである．
- ハーナッド以降，多くの研究者が多種多様なやり方でこの問題に取り組んでいる．

- 表象主義
  - ハイブリッドモデル (Harnad 1990)
  - 機能モデル (Mayo 2003)
  - 意図モデル (Sun 2000)
- 準表象主義
  - 認識論的モデル (Davidsson 1995)
  - 推測ゲーム (Steels and Vogt 1997)
  - 時間遅延と予測意味論に基づくモデル (Rosenstein and Cohen 1998)
- 非表象主義
  - コミュニケーションに基づくモデル (Billard and Dautenhahn 1999)
  - 行動に基づくモデル (Varshavskaya 2002)

- コンピューターは数学的な計算、論理的な推論、チェスなどのように、問題と答え（目標）、使える資源、考慮すべき要素が明確に定まっている課題を遂行するのはそれなりにうまくできる。
- また非常に単純化されたシミュレーションにおける課題（トイ・プロブレム）を遂行することもしばしば得意である。
- しかし現実の問題は必ずしもそうではない。

- 現実の問題を解決しようとするとき、私たちは自分の行動がどのような副次的な帰結を持つかを考えなければならない。
- しかし現実世界においてある行為は極めて多くの帰結を持ちうる。
- 教卓を動かす→床が傷つく、教壇から机が落ちる、机が壊れる、黒板との距離が変わる、机の上のチョークが落ちる、パソコンが落ちる、ペンが落ちる、大きな音がする、教室の外の鳩が驚いて飛び立つ、木の枝が揺れる、などなど。
- 論理的には、一つの前提は無限の多くの帰結を持ちうる。
- 人間はそういった帰結のうちのいくつかを無意識のうちに選んで考察している。
- しかしコンピューターにそのような選択をさせることは極めて難しい。

- 記号論理学をベースとした推論に基づいた人工知能では現実の問題に対処することは難しい、ということが認識されるようになる。
- 例えばコンピューターには状況や文脈を理解することや、人間にとって当たり前の常識を前提して思考することが致命的に苦手。
- テリー・ウィノグラードの例：「市評議会はデモ隊にデモの許可を出さなかった。彼らは暴動を【恐れていた／辞さない姿勢だった】からである」
- ここでの「彼ら」の指す対象は【】内の語句に応じて変化する。
- ある程度の常識のある人間ならば「彼ら」が何を指すかはすぐに理解できるが、コンピューターにはこれが極めて難しい。

- コンピューターは、人間の知能をシミュレートして、人間の代わりに課題を遂行するのではなく、あくまでも人間の知能をサポートするものとして使うことが提案される。
- Cf. テリー・ウィノグラード、フェルナンド・フローレス、『コンピューターと認知を理解する——人工知能の限界と新しい設計理念』、平賀譲訳、産業図書、1987。

## ドレイファスによる批判

- 哲学者のヒューバート・ドレイファスは、1964年にランド研究所に提出した報告書「錬金術と人工知能」において、人工知能研究者が実現できないことを吹聴して研究資金を詐取している、と強く非難。
- 人間の知能においては経験によって培われる勘、体に染みついた感覚、暗黙の知識などが重要な役割を果たす。
- こういったものは論理的な規則のようなもので表現することは不可能であり、したがって機械は真の意味で知能を実現することはできない、と主張した。
- ちなみに日本の数理論理学者、前原昭二は1961年に「人工頭脳と錬金術」（『科学基礎論研究』5巻3号、p. 108-11）というタイトルの論文を書き、人工知能の限界についてコメントしている。

- 計算機に自動的に何かを判定させるには、要求を完全に客観化しなければならない。
- しかしそのような客観化は多くの場合、不可能である。
- 「科学上の重要問題の多くが、客観的な問題ではなく、人間感情に密着した主観的問題である、ということは、十分注意するに値する事実」
- また要求を客観化できたとしても、必ずしも計算機が十分に素早くその答えを見つけてくれるわけではない。
- 上のような問題を指摘した上で、前原は「人工頭脳の能力の限界について述べて来たのは、決して、人工頭脳にケチをつけよう、という意志からではない、そうではなくて、適用範囲の限界を明示することこそ、科学者の良心というものであろう、と思ったからである」と述べる。

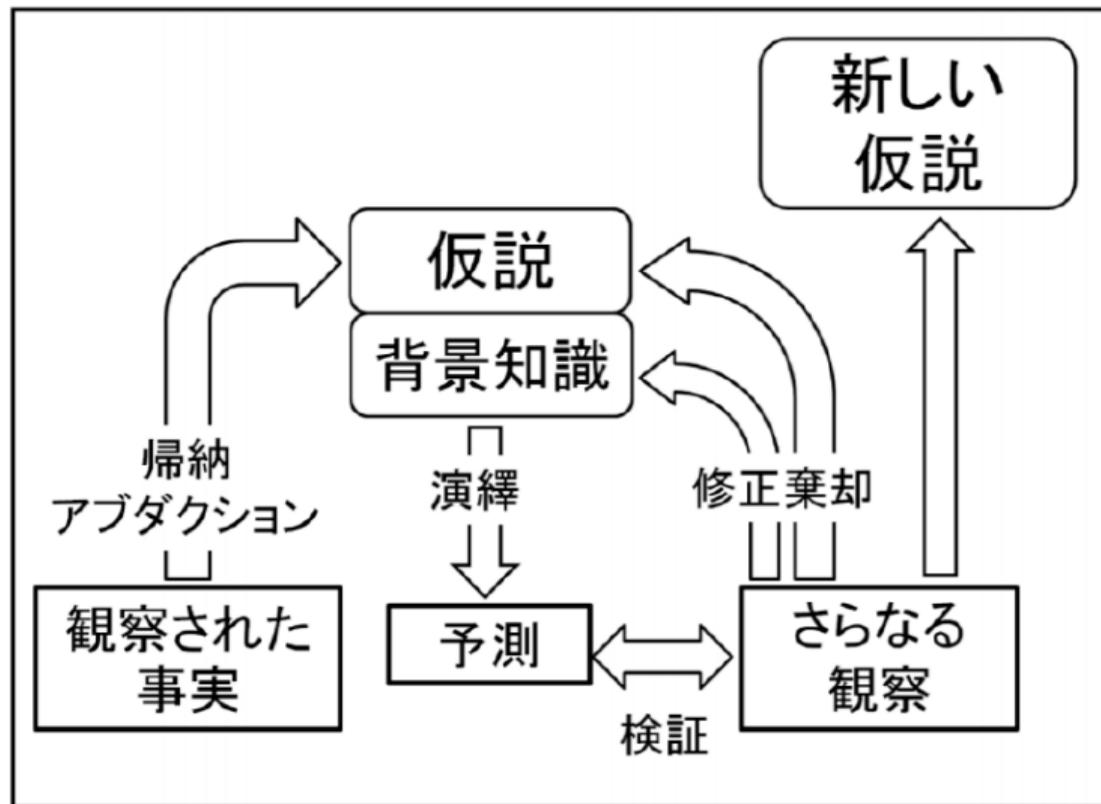
## 論理が得意なことと苦手なこと

- 体系だった仕方で、明確に定義された問題の解を求めるのは得意。
- 例えば数値計算、演繹的推論、プログラムのコンパイル（人工言語から人工言語への翻訳）など。
- 状況や文脈の理解、暗黙の知識が前提とされる思考、正解の規準が曖昧な問題、現実の状況で適時かつ適切に行動することが致命的に苦手。
- 例えば人の顔を見分ける、自然言語を翻訳する、物を掴むなど。
- テリー・ウィノグラードの例：「市評議員会はデモ隊にデモの許可を出さなかった。彼らは暴動を【恐れていた／辞さない姿勢だった】からである」

- 機械学習、帰納的推論、アブダクション
- エキスパート・システム
- 人工生命、遺伝的アルゴリズム（進化的計算）
- ニューラルネットワーク
- 身体化された認知
- AI から IA(Intelligence Augmentation) へ

- 次の三つの命題を考えよう。
  - (1) すべての哺乳類は授乳をする
  - (2) 鯨は哺乳類である。
  - (3) 鯨は授乳をする。
- (1)(2) から (3) を推論するのは**演繹**である。
- (2)(3) から (1) を推論するのは個別の事例からの一般化であり、**帰納**と呼ばれる。
- (1) という背景知識を持っていて (3) が観察されたときに (2) を推論するのは、観察された事実を説明する仮説を立てることであり、**アブダクション**あるいは**仮説形成**と呼ばれる。

# 演繹、帰納、アブダクション



- 1965年、J. A. ロビンソンが、「導出原理」という推論規則を持つ、コンピューター向けの一階述語論理の体系、節形論理 clausal logic が発表される。
- 導出原理（命題論理の場合）： $A \vee C$  と  $B \vee \neg C$  から  $A \vee B$  を導出。
- 節形論理を応用したプログラミング言語、Prolog が人工知能の開発に盛んに使われる。
- Prolog では一階述語論理式によって**知識**を表現する形でプログラムを書くことができる。

- 導出原理を逆転させて帰納的推論、アブダクシヨンの推論を機械によって実行させる試み → **帰納論理プログラミング、仮説形成的論理プログラミング**
- **機械学習**と呼ばれるテクニックの先駆け。

- 医者、化学者など、個別的な領域における**知識**や**ヒューリスティック**を機械に与えて、専門家の判断をシミュレートするシステム。
- ヒューリスティックというのは、何かの問題に取り組むときの指針、定石のようなもの。
- エキスパートは、経験に基づいた「直観」を働かせて状況に応じた判断をしている。
- 常にうまくいくわけではないが、許容できる程度にうまくいく。
- この直観を、ヒューリスティックとして明示化し、機械に実装することが試みられた。

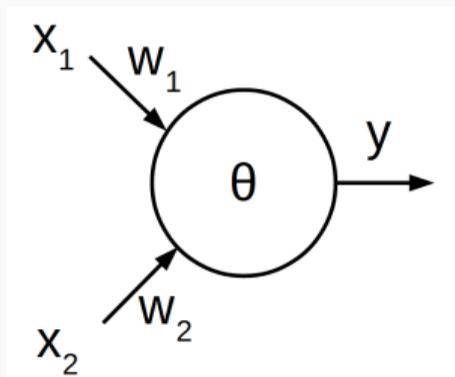
- 生物の**進化のメカニズム**を様々な問題解決に応用。
  - 1. ランダムにいくつかの解を生成。
  - 2. 解を評価。
  - 3. 評価の高い解を選択して、それらから次世代の解を生成。
  - 4. 2-3 を繰り返す。
- これで複雑な問題にうまく解が得られることがある。

- **生物の脳の構造を模したネットワーク**で問題の解を与える。
- 下図のようなユニット（パーセプトロン）を組み合わせて複雑な計算が可能になる。
- パーセプトロンは入力（下図の  $x_1, x_2$ ）のそれぞれに異なる重み（下図では  $w_1, w_2$ ）をかけて、その和、 $w_1x_1 + w_2x_2$  を出力（下図では  $y$ ）に変換する。
- $y$  は典型的にはある閾値  $\theta$  を用いて

$$y = \begin{cases} 1 & (w_1x_1 + w_2x_2 \geq \theta) \\ 0 & (w_1x_1 + w_2x_2 < \theta) \end{cases}$$

のように定義される。

## パーセプトロン



- 例えば入力が0か1のどちらかだとして、 $w_1 = w_2 = 1/2$ 、 $y$ は

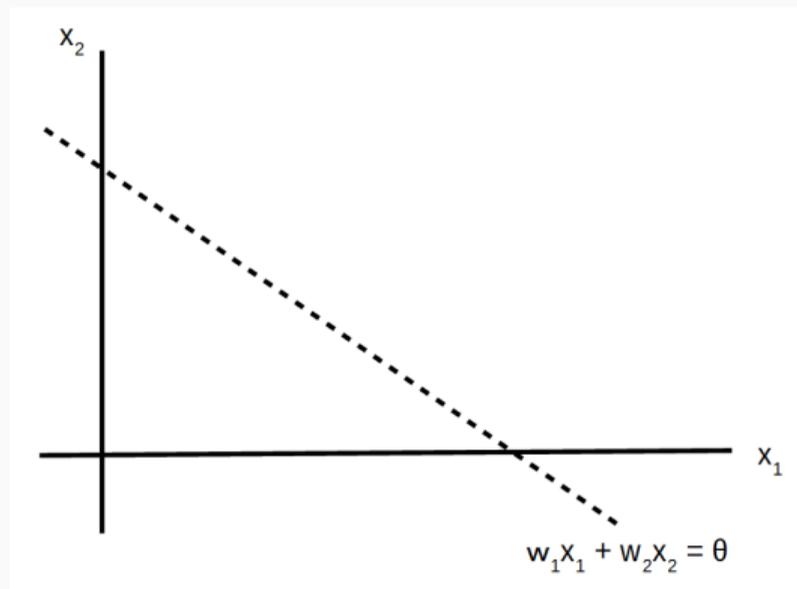
$$y = \begin{cases} 1 & (w_1x_1 + w_2x_2 \geq 1) \\ 0 & (w_1x_1 + w_2x_2 < 1) \end{cases}$$

によって定義される関数とすると、これは論理結合子 AND として機能する。

- すなわち入力がどちらも1ならば1を出力し、どちらか一方でも0ならば0を出力する。
- 同様にして論理結合子 OR (二つの入力のどちらかが1ならば1を出力する) や NAND (二つの入力のどちらかが0ならば1を出力する) として機能するパーセプトロンを作ることできる。

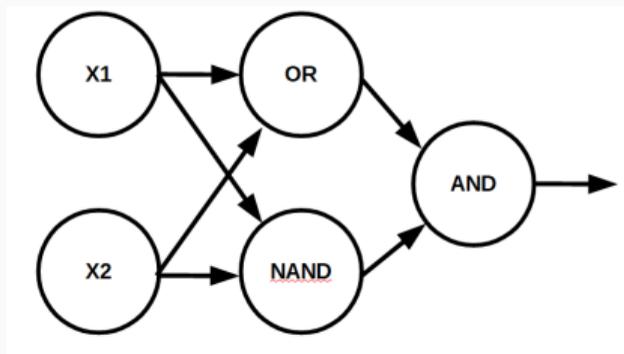
# ニューラル・ネットワーク

- 入力が  $n$  個のパーセプトロンは、 $n$  次元空間を  $w_1x_1 + \dots + w_nx_n = \theta$  という境界で分割するものである。
- 下図は二次元空間の分割の例。



## ニューラル・ネットワーク

- 一個のパーセプトロンでは XOR（どちらか一方だけが真のときにのみ真）を表現することはできない。
- しかし複数のパーセプトロンを組み合わせることで XOR が表現可能になる。



XOR を表すネットワーク

- 多数のパーセプトロンのを組み合わせた**ニューラル・ネットワーク**による複雑な関数の表現。
- 出力と正解の誤差から重みを調整する方法（誤差逆伝播法）や、ネットワークの構成に関する様々なアイデアが試みられる。
- 画像や音声のパターン認識に応用される。

- ロドニー・ブルックスなどによって提唱されたアイデア。
- 頭の中での表象（記号）を使った計算では現実世界において適応的な行動を取ることはいできない。
- 知能は脳の中で閉じているものではなく**現実世界との身体を介したインタラクション**によって実現される。
- 身体は脳という一つの中枢でコントロールされているのではなく、階層化された構造の中で分散的に制御されている。（**包摂アーキテクチャ**）

- 物理的記号システム：計算や論理的思考は得意
- ニューラル・ネットワーク：曖昧なパターン認識は得意
- 包摂アーキテクチャ：物理的世界における行動は得意

これらのいずれも人間の知能の一部をシミュレートしている。また特定のタスクをある程度こなすことはできるが、人間のような汎用性は持たない。

人工知能の発展は、人間の心の仕組みについての哲学的理論に大きな影響を与えてきた。

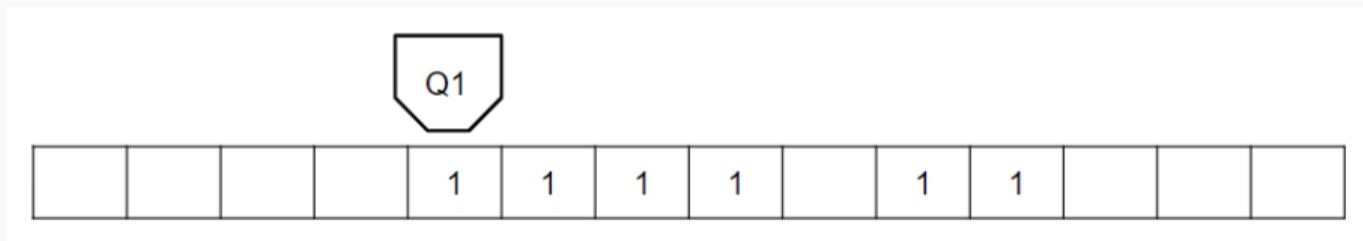
- 機能主義
- 心の計算理論
- 思考の言語仮説
- コネクショニズム
- 反表象主義

心的状態は、入力に対する出力の関数として捉えることが適切であるとする考え。こう考えると人工知能などもセンサーとアクチュエーターなどを備えることで人間と同等の心的状態を持つことができると考えられる。

心の機構を計算と類比的に捉える考え方。心はハードウェアとしての脳（あるいは身体）が行っているある種の計算と見なされる。ここでは計算の単位（コンピューターの中での0と1の信号に対応するもの）と、それを操作する規則（コンピューターではプログラムに対応するもの）、そしてそれを実現する仕組み（コンピューターでは電子回路に対応するもの）を人間の脳（あるいは身体）が備えていることを明らかにすることが課題になる。この立場はチューリング・マシンのような計算モデルに大きな影響を受けて発展してきた。

## チューリング・マシン

左右に無限に伸びたテープがあり、そのテープはマス目で区切られている。それぞれのマス目は、1が書き込まれているか、空白になっている。そのテープの上に機械が置かれており、それが置かれたマス目上の記号を読み取る。読み取った記号と内部状態に従ってテープ状の位置を変えたり、それが置かれたマス目の記号を書き換えたりする。内部状態と振る舞いはプログラムによって与えられる。



チューリング・マシンは、動作を始める時のテープの状態を入力、停止したときのテープの状態を出力とする計算機であると考えることができる。

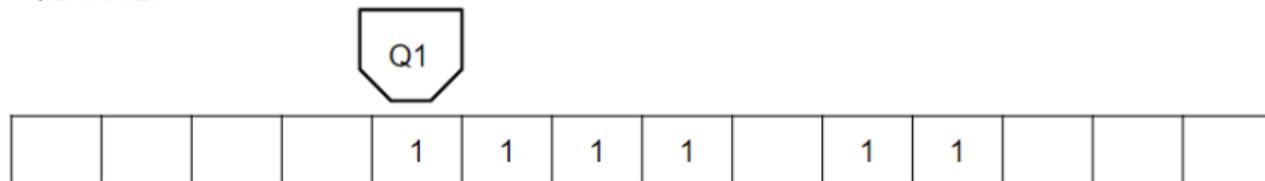
例えばチューリング・マシンに次のプログラムを与えるとする。

- Q1：1 を読み取ったら右に進み Q1 を実行する。空白だったら 1 を書き込み右に進み Q2 を実行する。
- Q2：1 を読み取ったら右に進み Q2 を実行する。空白だったら左に進み Q3 を実行する。
- Q3：1 を読み取ったら 1 を消し左に進み Q4 を実行する。
- Q4：1 を読み取ったら左に進み Q4 を実行する。空白だったら右に進み停止する。

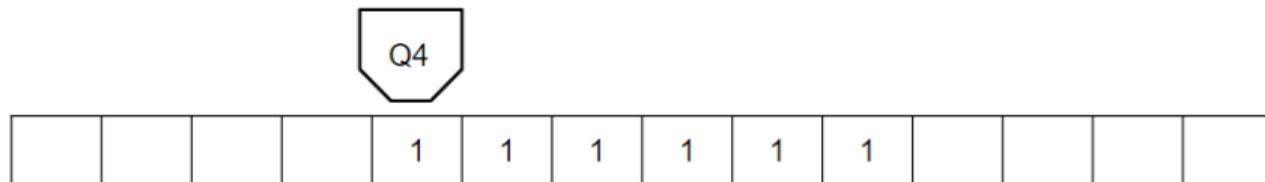
# チューリング・マシン

下図の開始状態で上のプログラムを受け取ったチューリング・マシンは、下図の終了状態で停止する。これは足し算を実行するプログラムと見なせる。

開始状態



終了状態



チューリング・マシンは単純なものだが、しかし適切なプログラムを与えることで、様々な計算をさせることができる。

さらにはチューリング・マシンのプログラムをコード化したものを入力として受け取って

実際、現在のコンピュータで計算できることはすべてチューリング・マシンで計算できる。

思考は頭の中にある、一種の言語によって行われるとする仮説。その精神言語（Mentalese）は、通常の言語と同じような構造（単語、文、文法規則、意味形成規則）を持っているとされる。思考の言語仮説は通常、心の計算理論のバリエーションとして提唱される。

ニューラル・ネットワークについての研究を通じて知能・心理・認知・行動を研究しようというアプローチ。

極端な論者は、人間の心理や行動について語る従来の記述（「素朴心理学」と呼ばれる）はすべて脳についての記述に置き換わるだろうと主張する。このような立場を消去主義という。