令和７年度

「専修学校による地域産業中核的人材養成事業」

動画教材（抜粋）

1

IT講座
高校生のための人工知能入門

～〈G検定〉へ向けて～

高校生のための人工知能入門　～〈G検定〉へ向けて～

[1]はじめに

はじめに(1)　〈G検定〉とは？

■ 〈G検定〉とは？

一般社団法人日本ディープラーニング協会(JDLA)が実施する,

AI(人工知能)・ディープラーニングの活用リテラシー習得のための検定試験

- 「G」は、「ジェネラリスト(generalist)」の「G」を意味する。
- 「リテラシー」とは, 元々は「読み書きの能力」のことを言うが, 広義には, 「特定の分野に関する基礎知識やそれをある程度活用できる能力」のことを指す。
- 〈G検定〉を取得することで, AIをビジネスに活用するために必要なリテラシーを身に付けていることを証明できる。

※ 同じJDLAが実施する〈E資格〉(エンジニア資格)は, ディープラーニングの理論を理解し、適切な手法を選択して実装する能力や知識を有しているかを認定する技術者向けの資格。

はじめに(2) 〈G検定〉は,「Di-Lite(ディーライト)」の推奨資格

■〈G検定〉とは？

- 〈G検定〉は, 経済産業省の「デジタルリテラシー協議会」が定めるデジタルリテラシー範囲「Di-Lite(ディーライト)」の推奨資格です。
- 「Di-Lite」とは,「デジタルを使う人材」であるために, すべてのビジネスパーソンが、共通して身につけるべきデジタルリテラシー範囲のことです。

現在,「Di-Lite」が定義する3領域と, その学習すべき範囲としてシラバスが推奨される3つの試験

- ITソフトウェア領域 ―― 〈ITパスポート試験〉
- 数理・データサイエンス領域 ―― 〈データサイエンティスト検定〉
- **AI・ディープラーニング領域 ―― 〈G検定〉**

これらの領域範囲は, 毎年協議され, アップデートされる。

1
高校生のための人工知能入門 ～〈G検定〉へ向けて～
[1]はじめに

はじめに(3) 〈G検定〉って難しいの？

■〈G検定〉は, 誰でも受けることができるが, 高校生には合格は難しい

〈G検定〉の受験対象は, AI・ディープラーニングに関わるすべての人。

高校生には超難関。目下のところ, 合格者は年間数人。

この教材の目的は, 〈G検定〉に合格することではなく, 〈G検定〉の内容の初歩の部分を利用して, AI・ディープラーニングの世界に入門すること。

高校生が学校で学習する範囲の数学や情報の知識(と, 少し背伸びした知識)でイメージが掴める部分もかなりあり, 充分に興味が持続できるはず。

はじめに(4) 〈G検定〉って，難しい数学が必要なんじゃないの？

■〈G検定〉の問題は選択式で，複雑な数式は出てこない

- あくまで，基礎知識やリテラシーの有無を問う試験なので，高度な数式の読解や立式を求められたり，開発技術者に求められるようなコンピュータ技術の細部の知識は問われたりはしない。
- また，問題は四択のため，あきらかに誤っている記述を除外しやすい。
- 社会人の受験者は，IT技術者以外の専門職，営業職，事務職，企画職などが4割ほどを占める。AIがさまざまな分野に入り込んできた現代では，エンジニア以外の職種にも，基本的なAI知識が求められるため，文科系のビジネスマンも数多く受験している。どのような職種であろうとも，右記の3つが把握できる基礎知識があれば，AIが急速に浸透してきている社会では，大きな武器となる。

AIでなにができて、なにができないのか

どこにAIを活用すればよいか

AIを活用するためにはなにが必要か

1
高校生のための人工知能入門 ～〈G検定〉へ向けて～
[1]はじめに

はじめに(5) 〈G検定〉に触れてみよう

〈G検定〉の内容の初歩を足がかりに，

AIの世界に飛び込んでみよう。

2

IT講座
高校生のための人工知能入門

～〈G検定〉へ向けて～

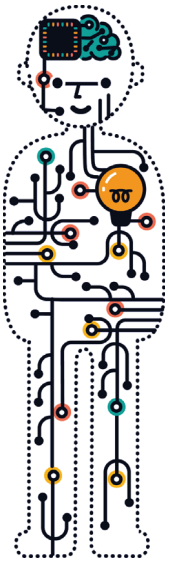
高校生のための人工知能入門 ～〈G検定〉へ向けて～

[2]人工知能(AI)とは

人工知能(AI)とは(1) 人工知能の定義

■ 人工知能とはなにか

- 「人工知能」という言葉には，専門家のあいだで共有されている明確な定義はない。
- 人工知能研究者のジョン・マッカーシーが，1956年にアメリカで開催された「ダートマス会議」で，初めて「人工知能(Artificial Intelligence)」という言葉を用いた。
- 一般的には，推論，認識，判断など，人間と同じ知的な処理能力を持つ情報処理システムを「人工知能」と呼ぶ。だが，「人間と同じ」「知的な」といった言葉に明確な定義はない。

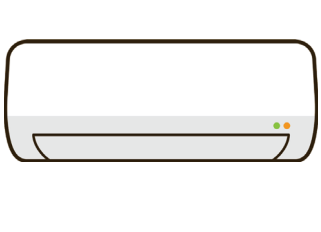


人工知能(AI)とは(2) 人工知能のおおまかな分類(1)

■ 人工知能のレベル別4分類

レベル1 シンプルな制御

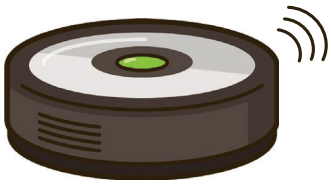
エアコンの温度調整, 洗濯機の水量調整など, すべてのふるまいがあらかじめ決められている制御。制御工学, システム工学などで培われてきている。





レベル2 古典的な人工知能

掃除ロボットや診断プログラムなど, 探索・推論・知識データを利用することで, 状況に応じてきわめて複雑なふるまいをするもの。



2
高校生のための人工知能入門 ～(G検定)へ向けて～
[2]人工知能(AI)とは

人工知能(AI)とは(3) 人工知能のおおまかな分類(2)

■ 人工知能のレベル別4分類

レベル3 機械学習を取り入れた人工知能

検索エンジンや交通渋滞予測など, 非常に多くのサンプルデータの分析を通じて, 入力と出力を結ぶ汎用的なパターンや法則を見い出し学習させる人工知能。そのようにして学習させたあと, 新たに入力された情報に対する判断を行う。



レベル4 ディープラーニング(深層学習)を取り入れた人工知能

「特徴量」と呼ばれる変数を自動的に学習するもの。画像認識, 音声認識, 自動翻訳などの分野で活用が進んでいる。

「特徴量」とは

AIが学習する際に使用する対象データの特徴を定量的な数値として表したもの。たとえば, 犬と猫の写真を見分ける学習をAIにさせるとき, 「どこを見れば効果的に犬と猫を見分ける手がかりとなるか」を数値で表現したものが, 特徴量である。

【レベル3】の機械学習では, 特徴量は人間が定義して数値化するが, 【レベル4】のディープラーニングでは, 大量の犬と猫の写真と, 「これは犬(猫)である」という正解を与えれば, AIは犬と猫を見分けるための特徴量を自動的に見い出し, 学習する。この「特徴量を自動的に見い出し学習する」点こそが, ディープラーニングの画期的なところ。

人工知能(AI)とは(4) 人工知能のおおまかな分類(3)

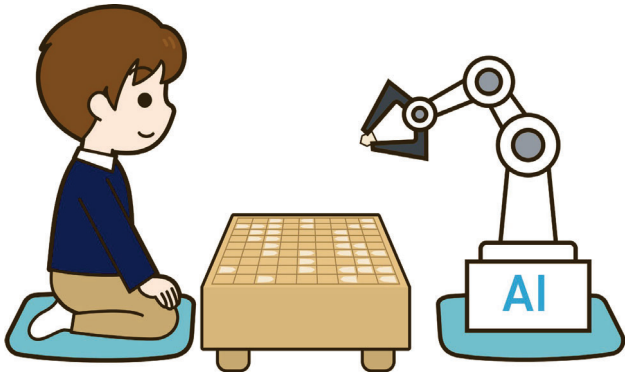


2
高校生のための人工知能入門 ～(G検定)へ向けて～
[2]人工知能(AI)とは

人工知能(AI)とは(5) AI効果

■ AI効果

人工知能でなにか新しいことが実現され, その原理がわかってしまうと,「それは単純な自動化であって知能とは関係ない」と結論づける人間の心理的な効果のこと。



3

IT講座
高校生のための人工知能入門

～〈G検定へ向けて〉～

高校生のための人工知能入門 ～〈G検定〉へ向けて～

[3]人工知能研究の歴史

人工知能研究の歴史(1) ダートマス会議

■ ダートマス会議(1956年)

のちに人工知能・情報理論の研究で重要な役割を果たす研究者らが，知的に行動したり思考したりするコンピュータ・プログラムの実現可能性について議論した。

代表的なビッグネームの参加者

- マーヴィン・ミンスキー
- ジョン・マッカーシー
- アレン・ニューウェル
- ハーバート・サイモン
- クロード・シャノン

■ ロジック・セオリスト

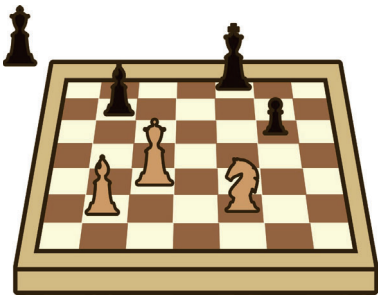
アレン・ニューウェルとハーバート・サイモンがダートマス会議でデモンストレーションした，世界初と言われる人工知能プログラム。コンピュータを用いて数学の定理を自動的に証明することが実現可能であることを示した。

人工知能研究の歴史(2) 人工知能研究のブームと冬の時代(1)

■ 第1次AIブーム(探索・推論の時代:1950年代後半～1960年代)

- 推論と探索に関する研究が進展(対象:パズル, 迷路, チェスなど)
- 明確に定義された特定の簡単な問題(トイ・プロブレム=「おもちゃの問題」)は解けても, 現実の複雑な問題は解けないことがあきらかになったため, ブームは急速に冷め, 冬の時代へと突入。





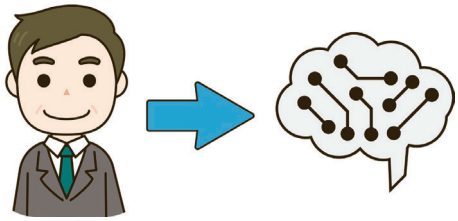
3
高校生のための人工知能入門 ～(G検定)へ向けて～
[3]人工知能研究の歴史

人工知能研究の歴史(3) 人工知能研究のブームと冬の時代(2)

■ 第2次AIブーム(知識の時代:1980年代)

エキスパートシステム

特定分野の専門家(エキスパート)からヒアリングした知識を大量にデータベースに蓄積し, それを基に推論したコンピュータが専門家に代わって回答するシステム。



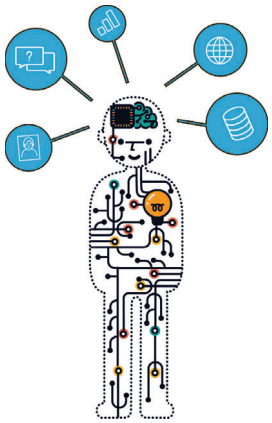
第五世代コンピュータプロジェクト(1982年～1992年)

当時の通商産業省(現・経済産業省)が進めた国家プロジェクト。人工知能コンピュータの開発をめざしたが, 大きな成果は得られなかった。

知識を蓄積し, 管理することの難しさがあきらかになってくるにつれブームは冷め, 1990年代半ばから, AIは再び冬の時代に。

人工知能研究の歴史(4) 人工知能研究のブームと冬の時代(3)

■ 第3次AIブーム(機械学習・特徴表現学習の時代:2010年～)



- ビッグデータと呼ばれる大量のデータを用いることで, 人工知能がみずから知識を獲得する機械学習が実用化される。
  ※人工知能の4分類【レベル3】相当
- 知識を定義する要素(特徴量)を人工知能がみずから自動的に学習するディープラーニング(深層学習)が登場。
  ※人工知能の4分類【レベル4】相当

▼ ディープラーニングの発達

◎ 2012年　画像認識競技で圧勝。

◎ 2015年　ディープラーニングを用いた人工知能 AlphaGo が, 人間の碁のチャンピオンに勝利。

◎ 2010年代中ごろから,「生成AI」(創造的な画像・音楽・文章などを生み出す)が注目を浴びる。

自然な文章を生成できる大規模言語モデル(Large Language Model: LLM)を応用したサービスが次々と出現。中でも, 米国の非営利研究機関「OpenAI」が公開したサービス「ChatGPT」は, 一般の人々に生成AIを認識させた点で, 一部の専門家は, ChatGPTの登場を第4次AIブームのはじまりと位置づけている。

4

IT講座
高校生のための人工知能入門

～〈G検定〉へ向けて～

高校生のための人工知能入門 ～〈G検定〉へ向けて～

[4]人工知能分野の問題(1～3)

人工知能分野の問題(1) トイ・プロブレム(おもちゃの問題)

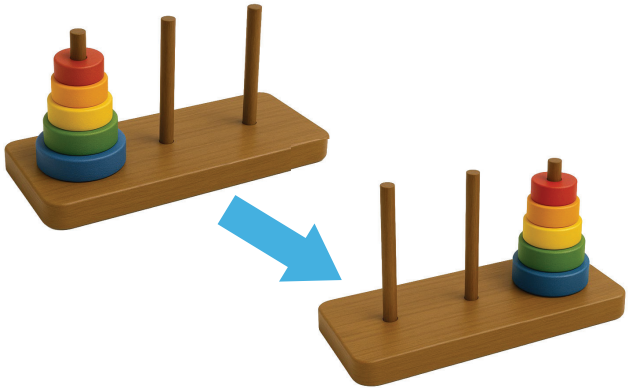
■ トイ・プロブレム toy problem

現実世界の複雑な問題を，本質を損なわない程度に簡略化したもの

AIの黎明期には，オセロ，チェス，三目並べ，ハノイの塔，数学の定理の証明などの問題をコピュータに解かせることが多かった。これらは「おもちゃの問題」，トイ・ブロブレムと呼ばれた。

例 「ハノイの塔」

三本の杭のうち，いちばん左の杭に刺さった大きさの異なる n 個の円盤をすべて，円盤をひとつずつ他の杭に移す動作を繰り返して，最終的にいちばん右の杭に移す。動かす過程で，円盤は，それより小さな円盤の上に乗せてはいけない。



人工知能分野の問題(2) フレーム問題(1)

■ フレーム問題

ジョン・マッカーシーとパトリック・ヘイズが1969年に論文の中で指摘した，人工知能における重要な問題。現在でも未解決という見解が一般的。

ロボット(あるいは，その頭脳である人工知能)が，「いま，しようとしていることに関係のあることがらだけを選び出すことが非常に難しい」という問題。

逆に考えると，「無限の可能性の中から，人間は，いましようとしていることに関係のあることがらだけを，どうやって選び出して思考し，判断しているのか？」という問いでもある。

4
高校生のための人工知能入門 ～(G検定)へ向けて～
[4]人工知能分野の問題（1～3）

人工知能分野の問題(2) フレーム問題(2)

■ 哲学者ダニエル・デネットによるフレーム問題の例示

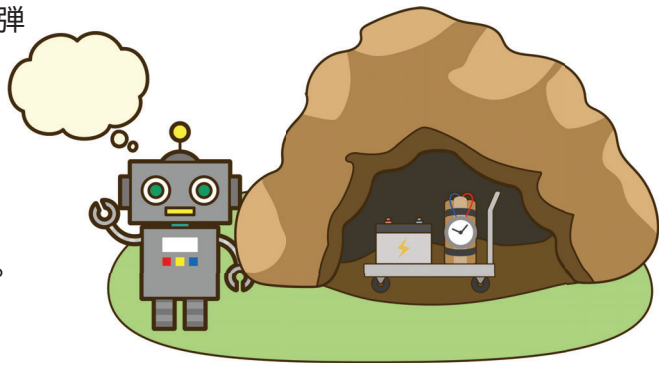
洞窟の中に台車があり，台車の上にバッテリーと時限爆弾がある。

バッテリーを持ち出すようにロボットに命令すると，どうなるか……。

ロボット1号 …… 台車ごとそのまま持ち出してきて，時限爆弾は爆発。

ロボット2号 …… 持ち出す方法の計算中に爆発。

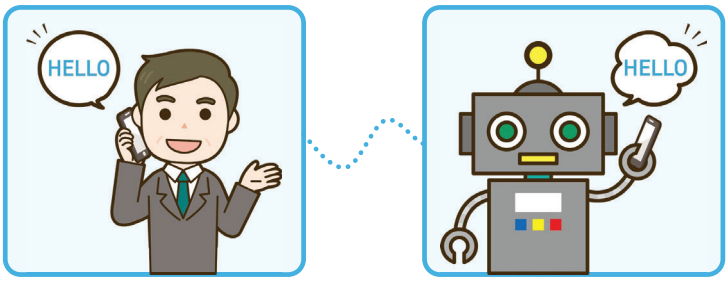
ロボット3号 …… 持ち出す方法の計算方法の計算中に爆発。



人工知能分野の問題(3) チューリングテスト

■ チューリングテスト

- 人工知能ができたかどうかを判定する方法として有名。イギリスの数学者, アランチューリングが提唱。
- 別の場所にいる人間がコンピュータと会話をし, 人間に相手がコンピュータだと見抜けなければ, そのコンピュータには知能があるとするテスト。
- 知能のメカニズムが判明していないのだから, あくまで外部から観察できる行動のみを見るしかないという考えかた。



4

IT講座
高校生のための人工知能入門

～〈G検定〉へ向けて～

高校生のための人工知能入門 ～〈G検定〉へ向けて～

[4]人工知能分野の問題（4～7）

人工知能分野の問題（4） 強いAIと弱いAI

■ 強いAIと弱いAI アメリカの哲学者，ジョン・サールの分類

強いAI　適切にプログラムされたコンピュータは，人間が心を持つのと同じ意味で心を持つとする立場。

弱いAI　コンピュータは人間の心を持つ必要はなく，有用な道具であればよいとする立場。

■ サールの「中国語の部屋」

サール自身は，強いAIを作ることは不可能と考えており，その論拠としての思考実験「中国語の部屋」を提示

［中国語の部屋］

ある密室に英語しかわからない人が閉じ込められている。その部屋の中には，中国語の質問に対する答えを中国語で書くための完璧なマニュアルがあり，そのマニュアルは英語で書かれている。部屋の外にいる中国語がわかる人が，中国語で書いた質問を壁の穴からその部屋に投げ込むと，部屋の中の人は英語のマニュアルどおりに，完璧な中国語の回答を書いて返す。部屋の外から見れば，部屋全体は中国語が理解できるかのようにふるまっている（チューリングテストをクリアするかもしれない）が，部屋の中には中国語を理解できる人はおらず，ただただ機械的な作業を行っているにすぎない。これは知能とは言えない。意味を理解せず，記号操作をしているだけである。

人工知能分野の問題(5) シンボルグラウンディング問題(記号接地問題)

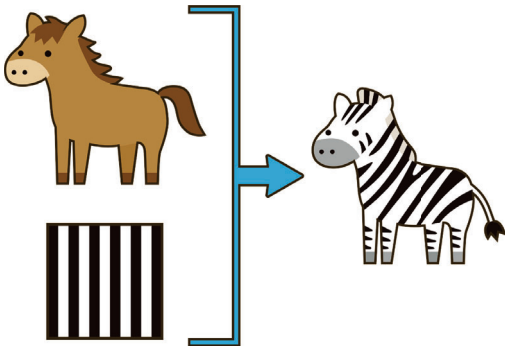
■ シンボルグラウンディング問題(記号接地問題)

認知科学者のスティーブン・ハルナッドが1990年に提唱。

記号(シンボル)で指し示されるものが, いかにして現実世界の実体と結びつくかという問題。

人間であれば,「シマ」という概念を知っており, かつ,「ウマ」という概念を知っておれば, 初めてシマウマを見ても,「これが話に聞く“シマウマ”なるものにちがいない」と推測できる。

ところが, コンピュータは「シマ」や「ウマ」といった記号の“意味”がわかっていないため,「シマウマ」という記号を現実の実体と結びつける(グラウンディングする)ことができない。これが, フレーム問題と並んで人工知能分野の難問とされている, シンボルグラウンディング問題である。

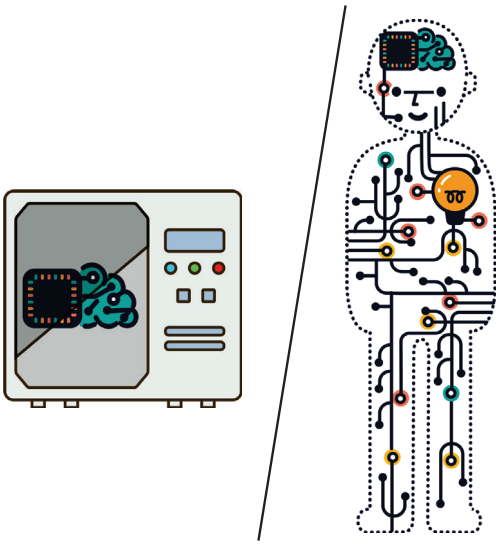


4
高校生のための人工知能入門 ～(G検定)へ向けて～
[4]人工知能分野の問題（4～7）

人工知能分野の問題(6) 身体性

■ 身体性とは

- 知能が成立するためには, 環境と相互作用するための“身体”が必要であるという考えかたがあり, そうした考えかたに立ったアプローチを「身体性」に着目したアプローチといった言いかたをする。
- 身体性は, シンボルグラウンディングにも深いかかわりがある。



人工知能分野の問題(7) シンギュラリティ(技術的特異点)

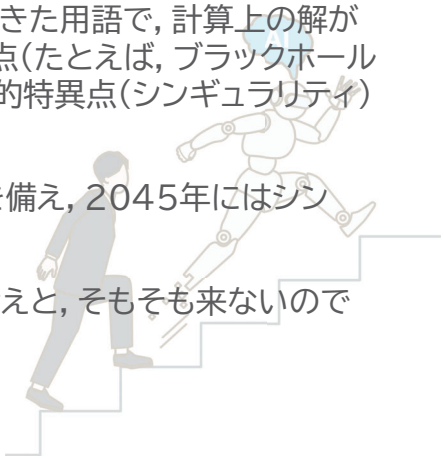
■ シンギュラリティとは

将来, 人間の知能を超えるAIが誕生し, さらにそのAIが自身を超えるAIを作り出し続け, 超知能が生み出されるという概念のこと。

シンギュラリティという言葉自体は, 数学や物理学でふつうに使われてきた用語で, 計算上の解が無限大に発散して, 一般的な法則や概念を当てはめることができない点(たとえば, ブラックホールの中心など)のことを言う。この概念を技術の発達に喩えたのが, 技術的特異点(シンギュラリティ)である。

AI研究者のレイ・カーツワイルは, 2029年にAIが人間と同等の知能を備え, 2045年にはシンギュラリティが来ると考えている。

シンギュラリティについては, それが来れば人類の終焉だと警戒する考えと, そもそも来ないのではないかという懐疑的な考えがある。



5

IT講座
高校生のための人工知能入門

～〈G検定〉へ向けて～

高校生のための人工知能入門 ～〈G検定〉へ向けて～

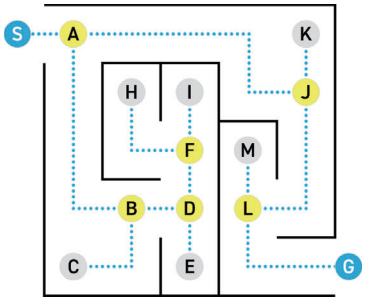
[5]人工知能をめぐる動向（1） 探索・推論

人工知能をめぐる動向（1） 探索・推論（1）

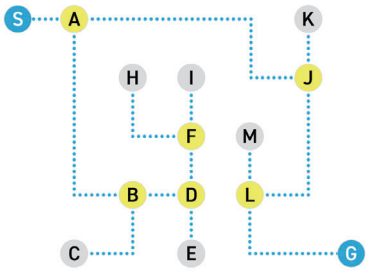
■ 探索

最初に与えられた状態から，目的状態に至るまでの道筋を，場合分けや試行錯誤をしながら探ってゆくこと。

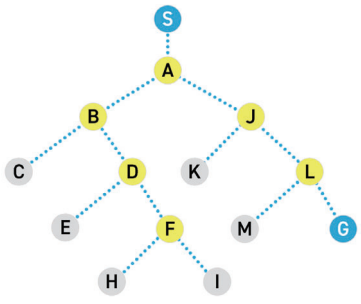
例）コンピュータに迷路の問題を解かせるには，どのように考えさせるか？



① 分岐や行き止まりを認識できるように印をつける。



② 迷路の壁を消し，S（スタート）からG（ゴール）までの分岐を表す。



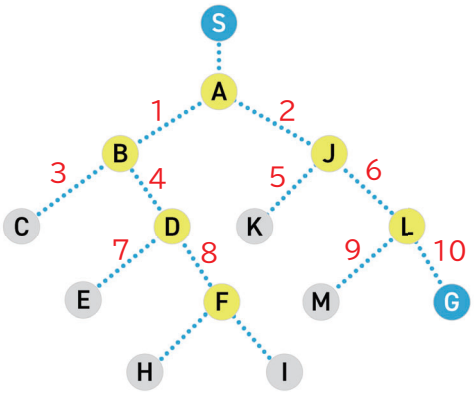
③ 木のような構造（ツリー構造）として捉えることができる。これを「探索木」と呼ぶ。

人工知能をめぐる動向(1) 探索・推論(2)

■ 幅優先探索

幅優先探索

出発点に近いノードから，同じ階層を順にしらみ潰しに当たり，目的のノードを発見できなければ，次の階層に進んで検索してゆく探索方法。



有利な点

最短距離でゴールにたどり着く解を必ず見つけることができる。

不利な点

立ち寄ったノードをすべて憶えておかなくてはならないため，メモリが不足する可能性がある。

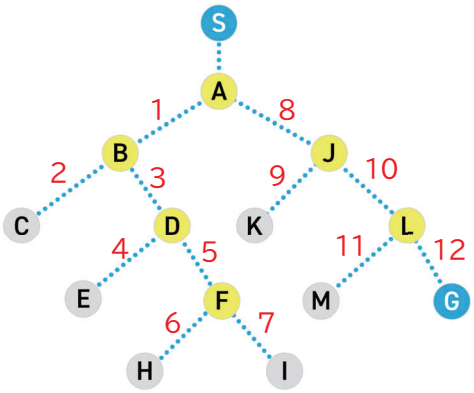
5
高校生のための人工知能入門 ～(G検定)へ向けて～
[5]人工知能をめぐる動向（1）
探索・推論

人工知能をめぐる動向(1) 探索・推論(3)

■ 深さ優先探索

深さ優先探索

あるノードからいったん深さ方向に行けるところまでノードを辿ってゆき，解が見つからなければ引き返して，次の枝に移る探索方法。



有利な点

解が見つからない枝のことは根本から忘れてしまってもよいので，幅優先探索よりもメモリが節約できる。

不利な点

運が悪いと，解が見つかるまでたいへんな時間がかかってしまうおそれがある。

人工知能をめぐる動向(1)　探索・推論(4)

■ ボードゲーム(チェス・将棋・囲碁など)も探索の問題

- ボードゲームをコンピュータで解くには，やはり迷路と同じように，考え得るすべての盤面の状態を探索してゆく問題になる。
- 代表的なボードゲームの組み合わせは，おおよそ右のとおり。

チェス　……　約10の120乗通り
将棋　……　約10の220乗通り
囲碁　……　約10の360乗通り

ちなみに，観測可能な宇宙全体にある水素原子の総数は，約10の80乗個と推定されている。

- 人間の頭脳が行なっている探索や推論では，探索木のすべてをしらみ潰しに探索してゆくのではなく，あらかじめ得られている知識や経験で，「これは解に繋がる可能性が低い」と思われる“枝”の探索を省略していると考えられる。コンピュータによる探索でも，これと同じように，経験的な知識を使って探索木の“枝刈り”をすることで，総当り探索を避けさせる。このような“経験則”のことを，「ヒューリスティックな知識」と呼ぶ。「ヒューリスティック heuristic」とは，「発見的な」という意味。
- ボードゲームでの「定石(じょうせき)」は，いわば，こうしたヒューリスティックなゲーム展開の体系。

5
高校生のための人工知能入門 ～(G検定)へ向けて～
[5]人工知能をめぐる動向（1）
探索・推論

人工知能をめぐる動向(1)　探索・推論(5)

■ ボードゲーム(チェス・将棋・囲碁など)も探索の問題

さまざまな探索方法が研究され，コンピュータはボードゲームで強くなっていったが，基本的には，コンピュータの計算能力を頼りにした“力任せ”(ブルートフォース)での探索の延長線上にあるものだった。

ゆえに，囲碁のように，組み合わせ数が飛び抜けて膨大なゲームでは，コンピュータは，長年，人間の名手に勝つことはできなかった。

2015年，DeepMind社の囲碁プログラム「AlphaGo」が，ヨーロッパチャンピオンの棋士を破り，翌 2016年には，さらに段位の高い韓国人のトップ棋士に大勝した。

AlphaGoが，旧来の計算力頼りの探索ではなく，“ディープラーニング”というまったく異なるアプローチの技術で人間の棋士を破った点が，歴史的な出来事である。

5

IT講座
高校生のための人工知能入門

～〈G検定〉へ向けて～

高校生のための人工知能入門 ～〈G検定〉へ向けて～

[5]人工知能をめぐる動向（2） 知識表現

人工知能をめぐる動向（2） 知識表現（1）

■ エキスパートシステム

ある専門分野の知識を取り込んで，その分野の専門家（エキスパート）のようにふるまうプログラムをエキスパートシステムと呼ぶ。1970年代から開発されはじめ，1980年代の第2次AIブームの牽引役となった。

マイシン（MYCIN）

ある専門分野の知識を取り込んで，その分野の専門家（エキスパート）のようにふるまうプログラムをエキスパートシステムと呼ぶ。1970年代から開発されはじめ，1980年代の第2次AIブームの牽引役となった。

■ 知識ベース

マイシンの知識を蓄える「知識ベース」は，「IF ○○ THEN □□」といった形式で書かれていた。「もし○○という条件が成立すると，そうしたら，その微生物は□□である」という専門家の知識の表現である。このような 知識の表現のしかたは，「IF-THEN ルール」と呼ばれた。

人工知能をめぐる動向(2) 知識表現(2)

■ 知識獲得のボトルネック

エキスパートシステムの開発が進むうち, 人間の専門家から知識を獲得するのは, 非常に困難であるということがわかってきた。

- 専門家の持っている知識は暗黙的(言語化できない／言語化されていない)。
- 専門家が自分の持つ知識を自発的に漏れなく表出するのは困難。
- 専門家からインタビューで知識を聞き出すのも容易ではない。

また, 構築した知識ベースが大きくなってくると, その保守が困難であることもわかってきた。

- 獲得した知識同士が互いに矛盾してくる。
- 知識が一貫性を欠いてくる。

エキスパートシステムの限界があきらかになってきた。

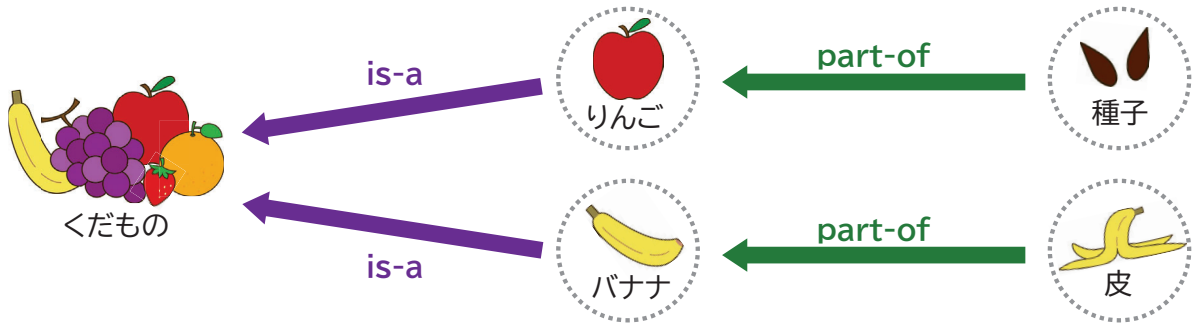
5
高校生のための人工知能入門 ～〈G検定〉へ向けて～
[5]人工知能をめぐる動向（2）
知識表現

人工知能をめぐる動向(2) 知識表現(3)

■ 意味ネットワーク

概念にラベルを付与して, 概念と概念との関係性を記号で表すネットワーク。知識表現によく用いられる。

- is-a 関係 ―「……である」 リンゴはくだものである。
- part-of 関係 ―「……の一部」 皮はバナナの一部である。



人工知能をめぐる動向(2)　知識表現(4)

■ オントロジー

知識･概念を共有･活用しやすいように体系化するための枠組み

※ 元々は哲学用語で,「存在論」という意味だが, 情報科学では上記のような意味で使われる。エキスパートシステムの開発において, 知識ベースの開発･保守に非常に労力がかかったことから, 知識の共有と再利用を研究する学問「知識工学」が期待された。オントロジーは, 知識工学の中心的な研究。

■ オントロジーの種別

ヘビーウェイトオントロジー(重量オントロジー)
- 記述を哲学的に考察し厳密に設計することを重視。
- 労力がかかる。
- 定義と知識の追加を人間が行う。

ライトウェイトオントロジー(軽量オントロジー)
- 効率化重視。
- 関係性の正当性より, 使えることを重視。
- 自動化しやすい。

5

IT講座
高校生のための人工知能入門

～G検定へ向けて～

高校生のための人工知能入門 ～〈G検定〉へ向けて～

[5]人工知能をめぐる動向(3) 機械学習(1～4)

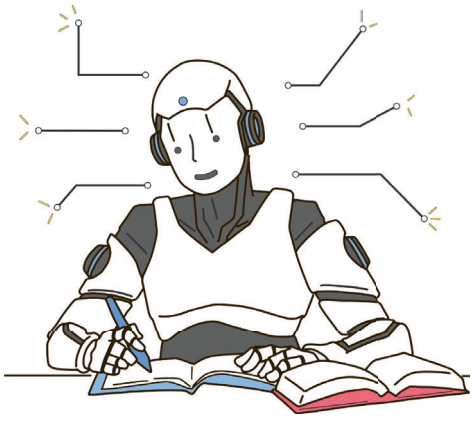
人工知能をめぐる動向(3) 機械学習(1)

■ 機械学習とは

- 人工知能のプログラム自身がデータから学習するしくみ。
- 学習に用いるサンプルデータから，データに潜んでいるパターンを学習してゆく。

よって

サンプルデータが多ければ多いほど，
理屈の上では，
望ましい学習結果が得られる。



人工知能をめぐる動向(3) 機械学習(2)

■ 次元の呪い

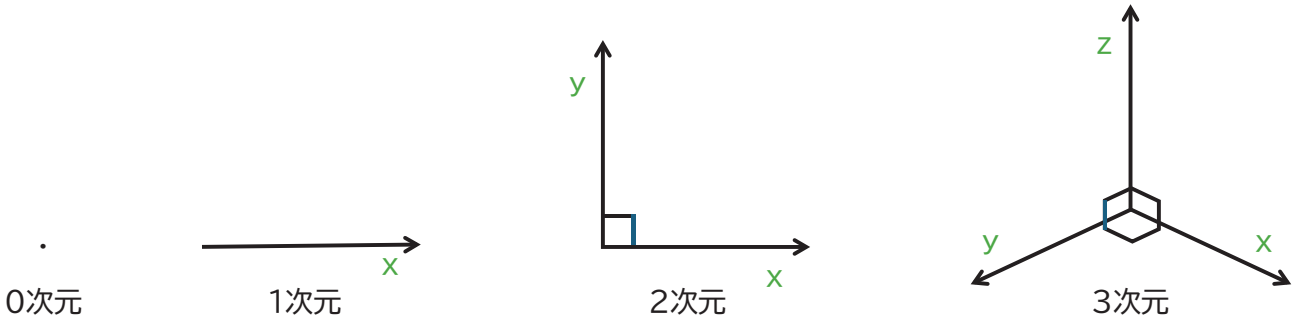
数理的な問題を解こうとするとき, 次元の数が増えると, 関与する要素の数や計算回数などが指数関数的に増大してしまい, 有限な現実的な時間内で解くことが困難になってしまうことを,「次元の呪い」と呼ぶ。

機械学習の場合は, 特徴量の数が多いと, 傾向を学習するために必要なデータの数が爆発的に増えてしまい, 学習に必要なサンプルデータの数が, 現実に調達できるサンプルデータの数を大幅に超えてしまうことが問題になる。

5
高校生のための人工知能入門　～〈G検定〉へ向けて～
[5]人工知能をめぐる動向（3）
機械学習（1～4）

人工知能をめぐる動向(3) 機械学習(3)

■ 次元とは(1)

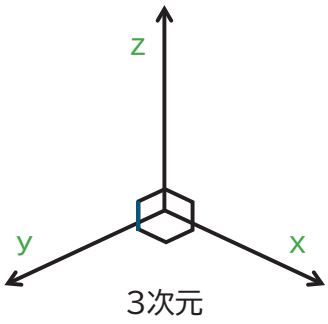


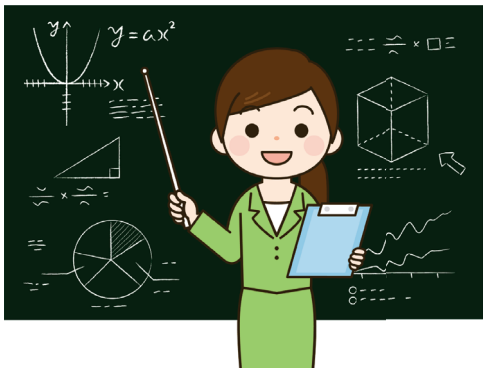
x軸に直交する(垂直に交わる)
y軸をプラス

x-y平面に直交するz軸をプラス

人工知能をめぐる動向(3) 機械学習(4)

■ 次元とは(2)





4次元以上の次元は，“絵”としてはイメージできないが，0次元から3次元まで次元を拡張してきたようにして，抽象的には考えることができる。
まず，x軸・y軸・z軸すべてに直交する“四つめの軸”の方向へ広がる空間を，“絵”としてではなく，抽象的に“概念”として想像してみよう。
さらに……。

5

IT講座
高校生のための人工知能入門

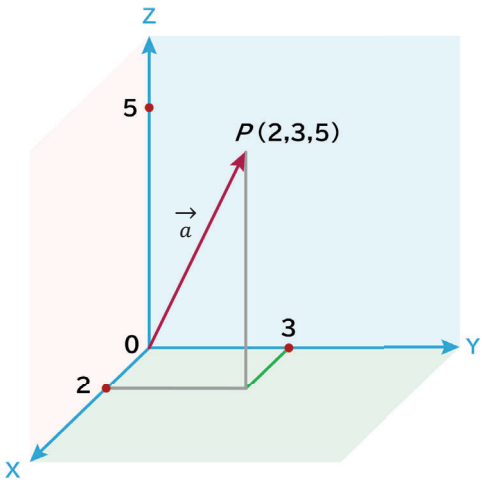
～〈G検定〉へ向けて～

高校生のための人工知能入門 ～〈G検定〉へ向けて～

[5]人工知能をめぐる動向(3) 機械学習(5～6)

人工知能をめぐる動向(3) 機械学習(5)

■ ベクトルと次元



$\vec{a}$ = (2,3,5) …… ①

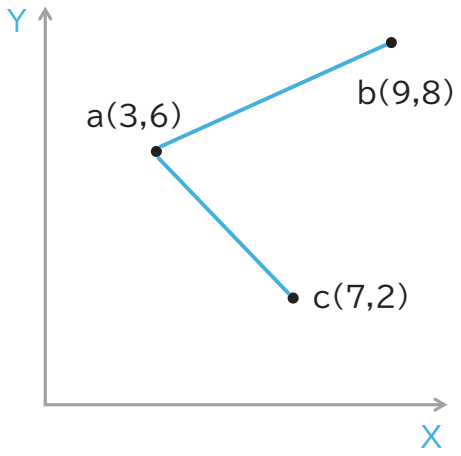
$\vec{b}$ = (2,3,5,7,2,0,15,8,3) …… ②

ベクトルを詳しく学ぶには,「線形代数」という数学を学ばなければならない。

人工知能をめぐる動向(3) 機械学習(6)

■ “似ている”ということを, どうやって判断するか？ (1)

- ユークリッド距離 …… 2点のあいだを直線で結んだときの距離



距離 $(a,b)=\sqrt{(3-9)^2+(6-8)^2}=6.32\cdots$

距離 $(a,c)=\sqrt{(3-7)^2+(6-2)^2}=5.66\cdots$

マップ上での直線距離が近いほうが“似ている”と判断する方法。
3次元以上にも拡張できるが, 高次元になるほど「次元の呪い」が効いてきて, あまり距離に差がつかなくなってくる。

n次元での点xと点yのあいだのユークリッド距離

$x=(x_1, x_2,\cdots,x_n)$, $y=(y_1,y_2,\cdots,y_n)$

ユークリッド距離(x,y) $=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}$

5

IT講座
高校生のための人工知能入門

～〈G検定〉へ向けて～

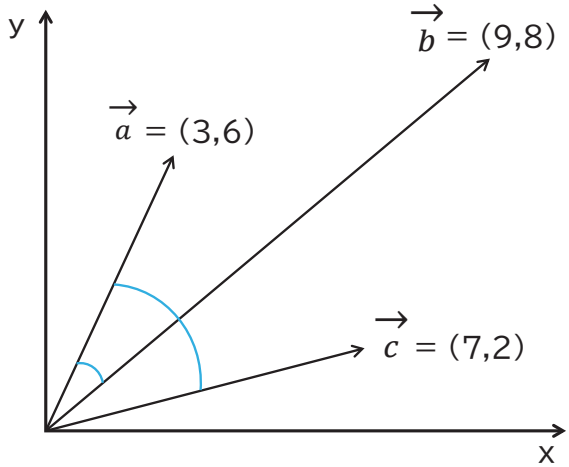
高校生のための人工知能入門 ～〈G検定〉へ向けて～

[5]人工知能をめぐる動向(3) 機械学習(7～11)

人工知能をめぐる動向(3) 機械学習(7)

■ “似ている”ということを, どうやって判断するか？ (2)

コサイン類似度 …… 2つのベクトルがなす“角度”を計算で求める

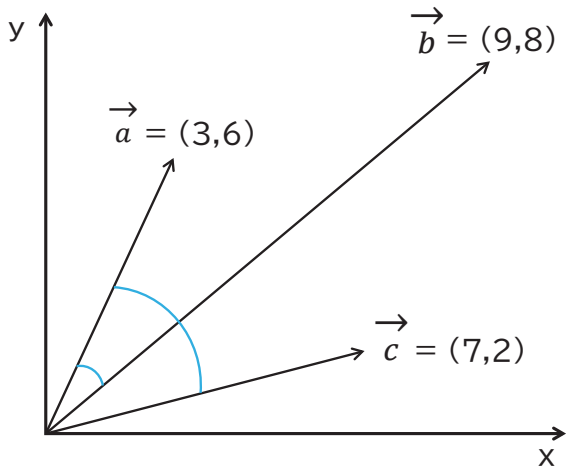


コサイン類似度を求める式

$$\cos_\mathrm{sim}(\mathbf{a},\mathbf{b}) = \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\,\|\mathbf{b}\|} = \cos\theta \quad \cdots\cdots ①$$

人工知能をめぐる動向(3)　機械学習(8)

■ “似ている”ということを，どうやって判断するか？（3）



コサイン類似度

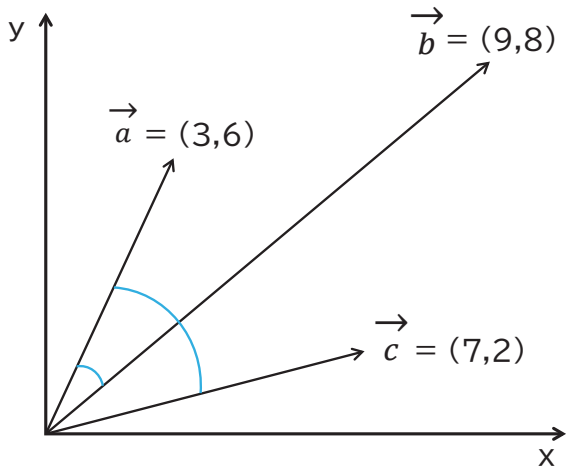
$$\cos_\mathrm{sim}(\mathbf{a},\mathbf{b}) = \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\,\|\mathbf{b}\|} = \cos\theta \qquad \cdots\cdots ①$$

- 太字のアルファベットはベクトルを表す。
- 通常の数式では，二重の縦線 ‖　‖ で数値や変数を囲むと「絶対値」を表すが，ベクトルの場合は，ベクトルの大きさ(「ノルム」と呼ぶ)を表す。
- 通常の数式では，「a･b」は「aとbの積」を表すが，aやbがベクトルの場合は，ふたつのベクトルの「内積」を表す。

※ ふたつのベクトル　$\mathrm{a} = (a_1,a_2,\ldots,a_n)$, $\mathrm{b} = (b_1,b_2,\ldots,b_n)$ があるとき
内積(a・b)は　$\mathrm{a}\cdot\mathrm{b} = a_1b_1 + a_2b_2 + \ldots + a_nb_n$　　…… ②

人工知能をめぐる動向(3)　機械学習(9)

■ “似ている”ということを，どうやって判断するか？（4）



コサイン類似度

$$\cos_\mathrm{sim}(\mathbf{a},\mathbf{b}) = \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\,\|\mathbf{b}\|} = \cos\theta \qquad \cdots\cdots ①$$

- ここではベクトルのノルム自体には意味はないので，実際の計算では，ベクトルのノルムが「1」になるよう「正規化」しておく。

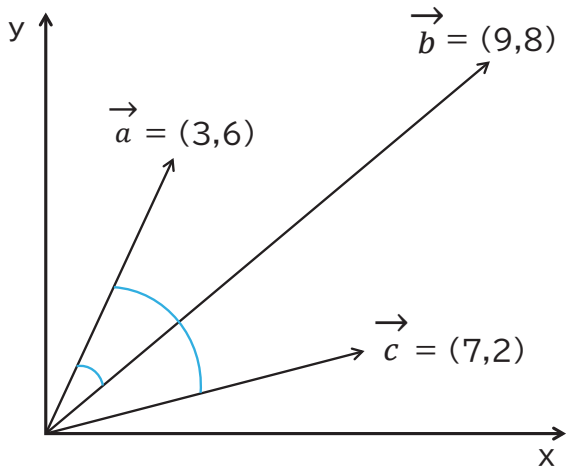
※ ベクトルの正規化の手順
ベクトル a = (3,6) を正規化しよう。

まず，ベクトル a のノルムを求める。
ノルムは，$\sqrt{3^2+6^2} = \sqrt{45} = 3\sqrt{5}$

次に，ベクトルの各成分をノルムで割る。
$\left(\frac{3}{3\sqrt{5}},\frac{6}{3\sqrt{5}}\right) = \left(\frac{1}{\sqrt{5}},\frac{2}{\sqrt{5}}\right)$　　…… ②

人工知能をめぐる動向(3) 機械学習(10)

■ “似ている”ということを，どうやって判断するか？（5）



コサイン類似度

$$\cos_\mathrm{sim}(\mathbf{a},\mathbf{b}) = \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\,\|\mathbf{b}\|} = \cos\theta \quad \cdots\cdots ①$$

正規化されたベクトル a = $\left(\frac{1}{\sqrt{5}},\frac{2}{\sqrt{5}}\right)$

正規化されたベクトル b = $\left(\frac{9}{\sqrt{145}},\frac{8}{\sqrt{145}}\right)$

正規化されたベクトル c = $\left(\frac{7}{\sqrt{53}},\frac{2}{\sqrt{53}}\right)$ ⎫ …… ②

ベクトルを正規化したら，式①の分数部分の分母の $\|a\|\ \|b\|$ は「1」となって消せることになり，内積の値そのものが $\cos\theta$ になる。内積を計算すれば，ふたつのベクトルのなす角度がわかる。

- $\cos\theta$ が 1 に近いほど，ふたつのベクトルの方向は似ている
- $\cos\theta$ が 0 のとき，ふたつのベクトルは直交している
- $\cos\theta$ が -1 に近いほど，ふたつのベクトルの方向は逆に近づいている

5
高校生のための人工知能入門　～〈G検定〉へ向けて～
[5]人工知能をめぐる動向（3）
機械学習(7～11)

人工知能をめぐる動向(3) 機械学習(11)

■ “似ている”ということを，どうやって判断するか？（6）

ユークリッド距離とコサイン類似度の比較

	ユークリッド距離	コサイン類似度
なにを測っているか	点と点との物理的な近さ(遠さ)	ベクトルの方向の近さ
機械学習での使い分け	「量のちがい」も重要なとき	「傾向・意味の方向性」が重要なとき
高次元での性質	高次元ではデータに差がつきにくい(「次元の呪い」)	高次元でも差が出やすい

機械学習では，実際の物理的な「量」よりも，意味や傾向を重視し，高次元での処理を行うことが多いため，コサイン類似度が用いられることが多い。

5

IT講座
高校生のための人工知能入門

～〈G検定〉へ向けて～

高校生のための人工知能入門 ～〈G検定〉へ向けて～

[5]人工知能をめぐる動向(4) 深層学習(1～4)

5

人工知能をめぐる動向(4) 深層学習(1)

■ 深層学習(ディープラーニング)への歩み

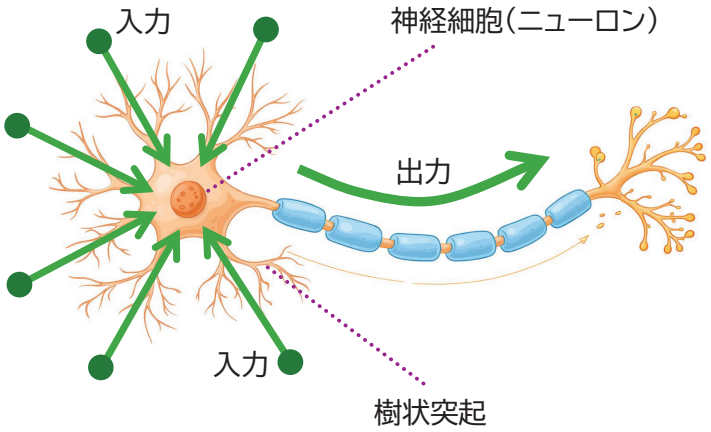
▼ 「ニューラルネットワーク」(神経回路網)の出現

生物の神経回路の働きを真似することで，学習を実現しようとするしくみ。機械学習のひとつと分類される。

- 1943年　神経生理学者のウォーレン・マカロックと数学者のウォルター・ピッツが，生物の神経細胞(ニューロン)を単純化したモデル(形式ニューロンモデル)を発表。ふるまいは固定されており，学習能力はなかった。
- 1958年　心理学者・計算機科学者のフランク・ローゼンブラットが，学習可能なニューロンモデル「パーセプトロン」を提案。

人工知能をめぐる動向(4) 深層学習(2)

■ 神経細胞の刺激の伝達



- 神経細胞に，ほかの神経細胞からの電気信号が，複数のルートで伝達される。
- 入力信号の強さが総和として充分で，ある閾値を超えると，その神経細胞から他の神経細胞へと信号が出力される。これを神経細胞の「発火」と呼ぶ。
- 非常に多くの神経細胞が絡み合い，電気信号のやりとりを続けていると，信号の通りやすい経路とそれほどでもない経路ができてくる。

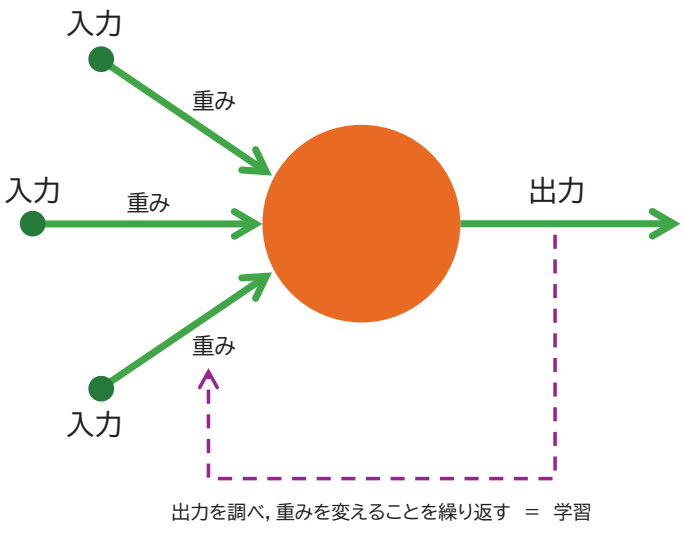
このように信号の通り道に癖がついてくることが

“学習”である。

5
高校生のための人工知能入門　～〈G検定〉へ向けて～
[5]人工知能をめぐる動向（4）
深層学習（1～4）

人工知能をめぐる動向(4) 深層学習(3)

■ パーセプトロン(単純パーセプトロン)



左図のように，入力×重みの総和が，ある閾値を超えれば「発火」し，出力が起動されるようにプログラムしたしくみを「パーセプトロン」と呼ぶ。

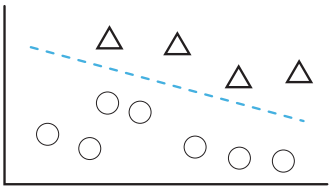
神経細胞において，信号の通りやすさに差がついてくる(学習する)ことを，入力に“重み”を与えて，“重み”を変化させることで表現する。

このような，入力層と出力層のみから成るパーセプトロンを「単純パーセプトロン」と呼ぶ。

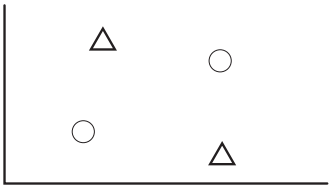
- 単純パーセプトロンは，視覚と脳の機能をモデル化したもので，単純なパターン認識ができる。たとえば，入力を画像のピクセル(画素)にしてやると，「特定の図形を含むかどうか」といった識別問題を解く学習ができた。

人工知能をめぐる動向(4)　深層学習(4)

■ 単純パーセプトロンの限界



直線で分離できる分類問題



直線で分離できない分類問題
※単純パーセプトロンでは解けない

- 直線でふたつに分けられる※分類(線形分離)問題しか解けない。
  1969年, 計算機科学者のマービン・ミンスキーとシーモア・パパートが, この単純パーセプトロンの限界を厳密に数学的に証明。
  ※　3次元では,「3次元空間を切る“平面”でふたつに分けられる」と, 1次元上げて解釈する。同様に, n次元の高次元空間でも, この限界は成り立つ。
- 第1次AIブームでは, パーセプトロンに大きな期待が寄せられたが, ミンスキーらの単純パーセプトロンの限界の発見によって, 急速にブームの熱が冷めた。

では, 入力層・出力層だけではない, 多層のパーセプトロンを作ったら・・・

- 多層パーセプトロンのアイディアは当時にもあったが, どう学習させたらよいのかという理論が未熟なうえに, コンピュータの性能がまだまだ低すぎた。
- 1986年, 多層パーセプトロンとそれに学習をさせる方法(誤差逆伝播法)をデビッド・ラメルハートらが提案。ミンスキーらが示した単純パーセプトロンの限界を超えられることが判明。ニューラルネットワークの研究に火が点き, 第2次ニューラルネットワークブームが到来。

5

IT講座
高校生のための人工知能入門

～〈G検定〉へ向けて～

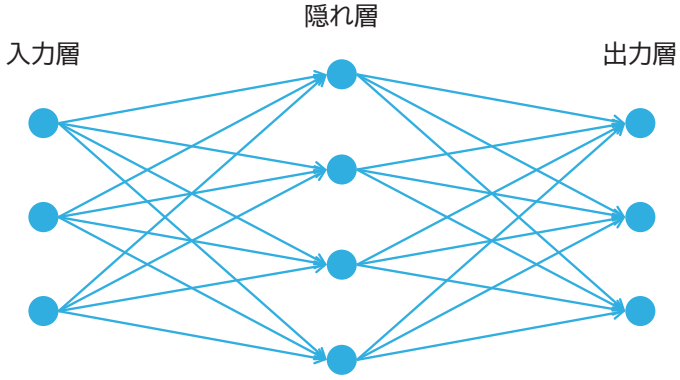
高校生のための人工知能入門 ～〈G検定〉へ向けて～

[5]人工知能をめぐる動向(4) 深層学習(5～9)

人工知能をめぐる動向(4) 深層学習(5)

■ 多層パーセプトロン

入力層と出力層だけから成る単純パーセプトロンに,「隠れ層」(「中間層」とも呼ぶ)を追加したモデル。単純パーセプトロンの限界であった線形分離不可能な問題にも対応可能。



「隠れ層を増やしていけば, もっと
もっと複雑な処理ができるのでは？」

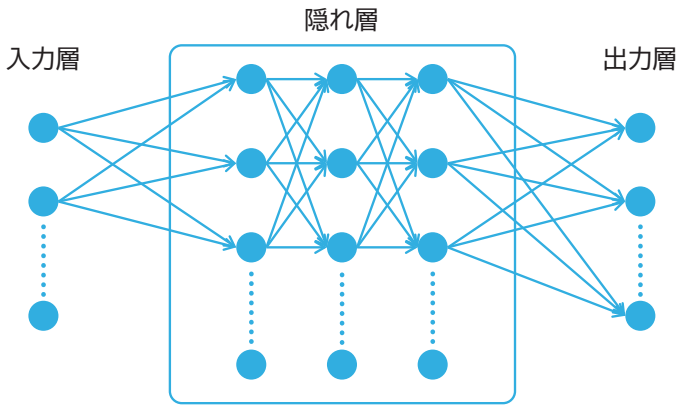
ディープニューラルネットワーク[DNN]
(ディープラーニング)へ

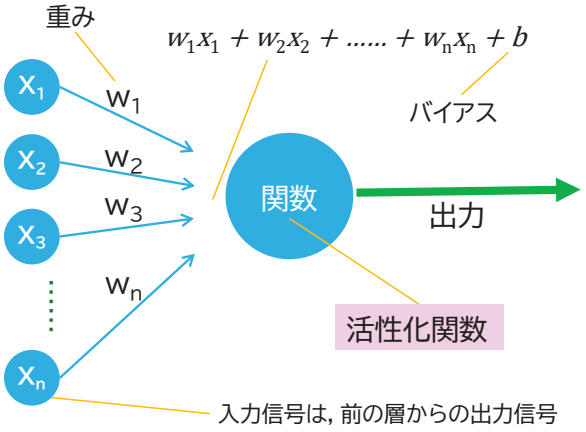
人工知能をめぐる動向(4) 深層学習(6)

■ ディープニューラルネットワークの学習に関わる要素

ニューラルネットワークの“学習”とは，ニューロン間の結合の強さを表すパラメータである「重み」と「バイアス」を最適化すること。

▼隠れ層のニューロンひとつが受け取る入力信号は下記のように処理される。





5
高校生のための人工知能入門 ～〈G検定〉へ向けて～
[5]人工知能をめぐる動向（4）
深層学習(5～9)

人工知能をめぐる動向(4) 深層学習(7)

■ 実装のためには，下記の課題を解決しなければならない

- 活性化関数として，どのような関数を用いるか
- 「重み」や「バイアス」を最適化するには，どのような考えかたやアルゴリズムが必要か

これらを深く理解するには，すべての高校生が詳しく習うわけではない領域の数学(微分，偏微分など)が必要だが，この教材ではそこまで踏み込まないので，下記のような概念的なアウトラインを把握しておこう。

■ ニューラルネットワークの学習のおおまかな手順

① パラメータ(「重み」と「バイアス」)を初期化。
② 学習データを入力層に入力し，入力層→隠れ層→出力層へと伝播させる。
③ 出力層から予測結果を出力させ，出力と“正解”とのあいだの誤差を表す「損失関数」というものを算出する。
④ 誤差が小さくなるようにする「勾配降下法」というアルゴリズムを用いてパラメータを更新する。 ※局所最適解，大域最適解に注意
⑤ 最適解に至るまで，②から④を繰り返す。

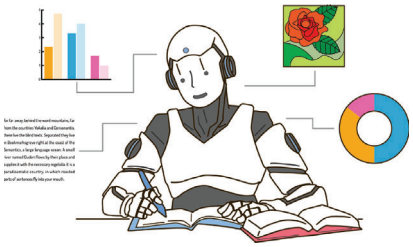


人工知能をめぐる動向(4) 深層学習(8)

ここまでは，ディープラーニングのほんの初歩の概要にすぎない。ここから先，下記のようなディープラーニングのさまざまな要素技術や応用について学ぶのが，〈G検定〉の中核部分となる。

■ ディープラーニングの要素技術

- 隠れ層には，特化した処理を行うさまざまな役割を持たせた層が開発されている。
- 入力層から出力層に至る流れは，必ずしも一方通行ではなく，層をスキップしたり，隠れ層の一部をループさせて，出力を入力にフィードバックするような構造のものもある。



■ ディープラーニングの応用

- 画像認識
- 音声処理
- 自然言語処理
- データ生成(画像・音声・文章)
  いわゆる「生成AI」と「ディープラーニング」が混同されがちだが，これらは同じものを指しているのではない。
  生成AIは，ディープラーニングを応用した“用途”のひとつであり，「生成AIには，ディープラーニングの技術が用いられていることが多い」という関係にある。

5
高校生のための人工知能入門 ～〈G検定〉へ向けて～
[5]人工知能をめぐる動向（4）
深層学習(5～9)

人工知能をめぐる動向(4) 深層学習(9)

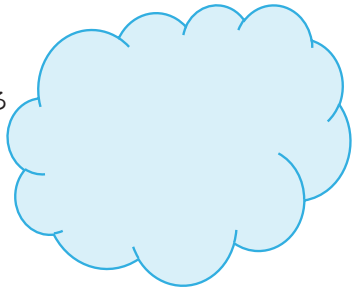
■ ディープラーニングを用いた「生成AI」

複雑で高次元なデータを、ニューラルネットワークが扱いやすい形に要約した低次元ベクトル表現が存在する空間を、「潜在空間」と呼ぶ。この潜在空間を得ることで，実在しないデータを生成することができる。この技術は，画像だけでなく，音声や文章にも適用できる。



実在するデータ

ニューラルネットワークに学習させる



潜在空間

※ 元のデータが持つ本質的な特徴や構造が，ベクトルの形で抽象的に表現されている。

ベクトル表現を解釈し，元のデータの本質的な特徴を持つ画像を生成する





生成されたデータ

6

IT講座
高校生のための人工知能入門

～〈G検定〉へ向けて～

高校生のための人工知能入門 ～〈G検定〉へ向けて～

[6]おわりに

おわりに(1)

■ すべてのAIがディープラーニングになったわけではない



レベル 1 から レベル 4 までのAIは，いまも併存しており，その技術は適所で使われている。

単純な場合分けですむような仕事に，膨大な計算資源を要する手段を用いると，エネルギーの浪費に繋がったり，かえって非効率になってしまうこともある。

「鶏を割(さ)くに焉(いずく)んぞ牛刀(ぎゅうとう)を用いん」

(小さな物事を処理するのに、大げさな手段を用いる必要はない)

おわりに(2)

■ “賢さ”は“正しさ”を担保しない

- ハルシネーション(hallucination＝幻覚)
  生成AIが, もっともらしい偽の情報を生成してしまう現象
- AIは, “正しさ”を求めているわけではなく, “もっともらしさ”(つまり, 確率的にいかにもありそうな状態)を優先して文章を繋いでいる。その結果, たまたま“正しく”なってしまうことも多いだけ。
- 人間と同じくらい賢いAI　≒　人間と同じくらい信用できないAI
- 判断に至る過程や, 判断の根拠が説明できる「説明可能AI(Explainable AI)」も研究されている。
- ハルシネーションを極力起こさせないように生成AIを利用するには, プロンプト(生成AIへの指示)に目的や条件などをしっかりと網羅し, AIに不必要な推測をさせないことが肝要。

おわりに(3)

■ 人工知能や〈G検定〉に興味が湧きましたか？

