

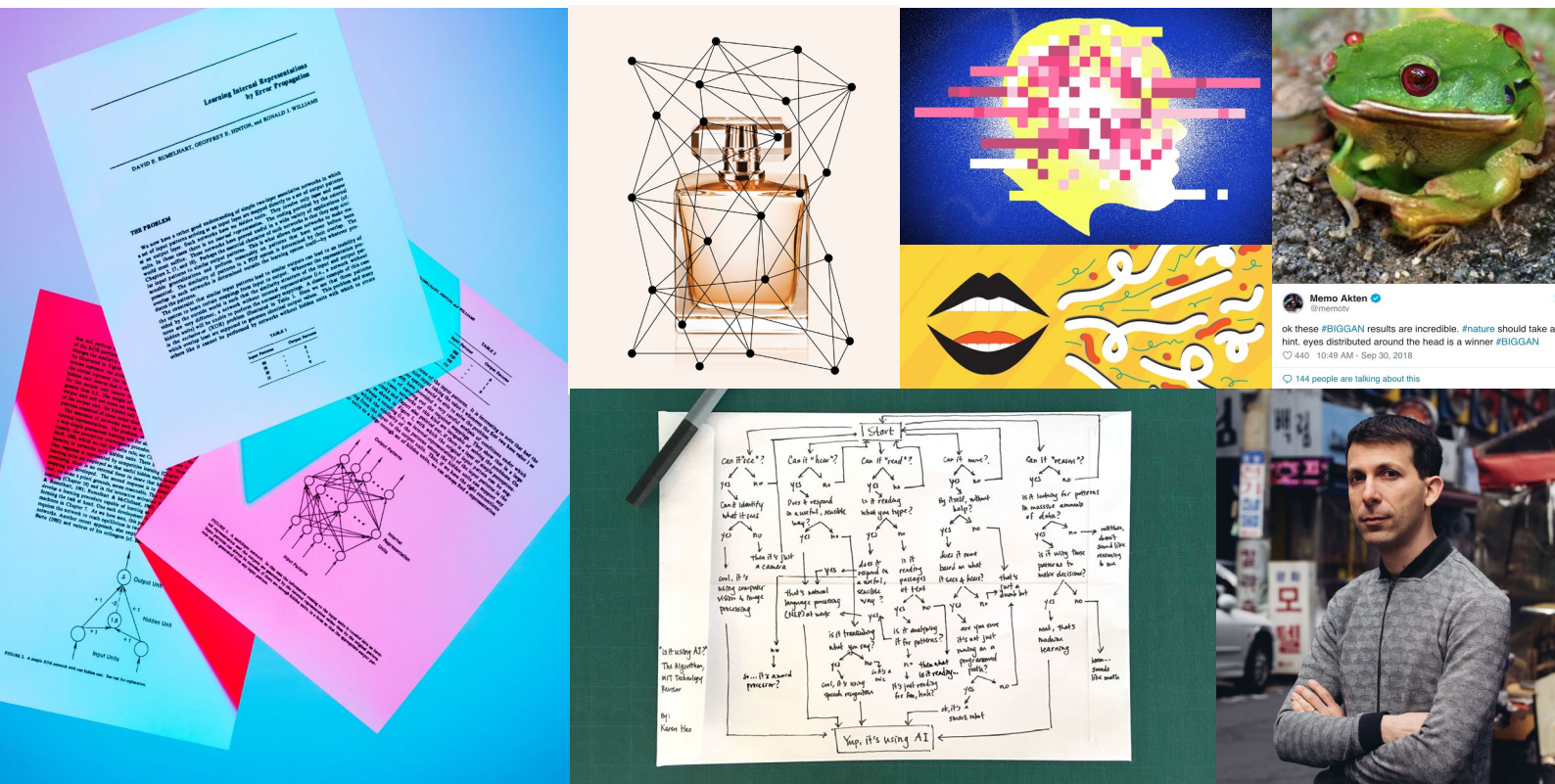
MIT Technology Review

Published by KADOKAWA / ASCII



What is AI?

人工知能への「期待」と「誤解」



CONTENTS

- 001 人工知能バブル
「三度目の冬」はやってくるのか
- 020 「人工知能ってなに？」
あなたはこの質問に答えられますか
- 023 深層学習の終わりの始まり
arXiv 投稿論文に見る AI 研究のトレンド
- 031 いまさら聞けない「機械学習」とは何か？
フロー図を描いてみた
- 036 深層学習万能論に異論、
進化的アルゴリズムがゲームで勝つ
- 041 AI に創造性を吹き込む「GAN」とは
どんなテクノロジーか？
- 044 「人間の知識にとらわれない人工知能を」
アルファ・ゼロ開発者に聞く
- 048 深一般企業の「AI 導入」で
ビジネスが変わらない理由
- 056 機械学習で「真の知性」を持つ機械は作れない、
元ワトソン開発者

2015年、グーグルの子会社が開発した「アルファ碁 (AlphaGo)」が人間のプロ棋士を破ったニュースは、多くの人々に衝撃を与えた。それ以降、人工知能 (AI) に対する期待は日々高まり、あらゆる産業界で AI の活用が喧伝されている。だが、過剰な期待は AI に対する誤解を生み、将来に暗い影を落とす可能性がある。あらためて現在、AI 分野で主流となっているテクノロジーの基礎知識を振り返るとともに、MIT テクノロジーレビューに掲載された今後注目の動向をまとめた。

Learning Internal Representations by Error Propagation

DAVID E. RUMELHART, GEOFFREY E. HINTON, and RONALD J. WILLIAMS

THE PROBLEM

We now have a rather good understanding of simple two-layer associative networks in which a set of input patterns arriving at an input layer are mapped directly to a set of output patterns at an output layer. Such networks have no *hidden* units. They involve only *input* and *output* units. In these cases there is no *internal representation*. The coding provided by the external world must suffice. These networks have proved useful in a wide variety of applications (cf. Chapters 2, 17, and 18). Perhaps the essential character of such networks is that they map similar input patterns to similar output patterns. This is what allows these networks to make reasonable generalizations and perform reasonably on patterns that have never before been presented. The similarity of patterns in a PDP system is determined by their overlap. The overlap in such networks is determined outside the learning system itself—by whatever produces the patterns.

The constraint that similar input patterns lead to similar outputs can lead to an inability of the system to learn certain mappings from input to output. Whenever the representation provided by the outside world is such that the similarity structure of the input and output terms are very different, a network without internal mappings (i.e., a network without hidden units) will be unable to perform the necessary mappings. A classic example of this case is the *exclusive-or* (XOR) problem illustrated in Table 1. Here we see that those patterns which overlap least are supposed to generate identical output values. This problem and many others like it cannot be performed by networks without hidden units with which to create

TABLE 1

Input Patterns	Output Patterns
00	0
01	1
10	1
11	0

TABLE 2

Input Patterns	Output Patterns
00	0
01	1
10	1
11	0

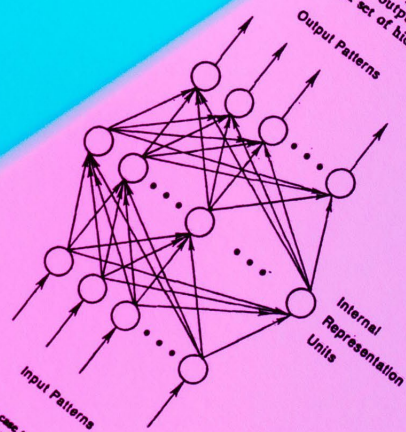


FIGURE 1. A multilayer network. In this case the information coming to the input units is recoded into an internal representation and the outputs are generated by the internal representation rather than by the original patterns. Input patterns can always be recoded, if there are enough hidden units, in a form so that the appropriate output patterns can be generated from any input pattern.

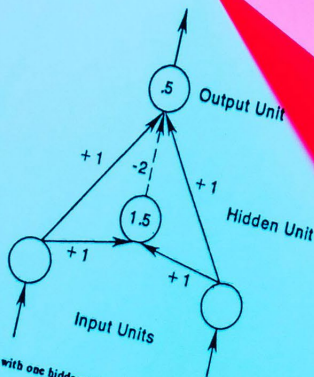


FIGURE 2. A simple XOR network with one hidden unit. See text for explanation.

that will perform a... of the XOR problem... changes the similarity... As illustrated in Figure... arrows represent the strength... the circles represent the strength... hidden unit insures that it will... for the output unit insures that... greater than 0.5. The weight of... output unit will not come on when... of the output unit, the hidden unit... patterns consisted of three rather than... The existence of networks such as th... a very simple guaranteed learning (Chapter... namely, the perceptron convergence procedur... Hoff, 1960, which we call the delta rule; see Ch... learning in networks with hidden units. There... One response is represented by competitive learning (Chapter... learning rules are employed so that useful hidden units develop... promising, there is no external force to insure that hidden... that, on some a priori grounds, seems reasonable. This is... & Rumelhart, 1981; Rumelhart & McClelland, 1982)... develop a learning procedure capable of learning an... forming the task at hand. One such development is... machines in Chapter 7. As we have seen, this procedure... requires the network to reach equilibrium in two... networks. Another recent approach, also employed... Barto (1985) and various of his colleagues (cf. Barto...

...ions of the input patterns. It is interesting to note that in a large... third input taking the value 1 whenever the first two have value 1... case, networks of this kind are incapable of solving the problems. On... of carrying out the required mappings. They show that in a large... (Rumelhart & McClelland, 1982) have provided a very careful analysis of... (Rumelhart & McClelland, 1982) also pointed out, if there is a layer of simple perceptrons... which a recoding (i.e., an internal representation) of the input patterns can be aug... always from the input to the output units. Thus, if we have the right connections... ing from the input to the output units, we can always find a representation... units to a large enough set of hidden units, we can always find a representation

人工知能バブル 「三度目の冬」はやってくるのか

by James Somers

深層学習と呼ばれる人工知能 (AI) テクノロジーがもてはやされ、いずれは人間の仕事を奪うかのように喧伝されている。しかし、現在の AI ブームの根本を支えているのは 30 年前の論文で発表された技術であり、まったく新しいブレークスルーが起こっているわけではない。

2 017 年の秋にオープンするベクター研究所 (Vector Institute) は、ほどなくして世界の中心となる場所だ。カナダ・トロントのダウンタウンにあるピカピカのビルの 7 階の非常に大きな部屋に居を構える同研究所は、人工知能 (AI) の世界的な中心になることを目標としている。

トロントには、現在の AI の興隆の元となった技術「深層学習」の生みの親であるジェフリー・ヒントンが住んでいる。「30 年経って現在を思い返せば、ジェフは単に AI と呼ばれている『深層学習』のアインシュタインだったと呼ばれるでしょう」と語るのは、ベクター研究所の共同設立者の一人に名を連ねるジョーダン・ジェイコブズだ。ヒントンは AI 分野のトップクラスの研究

者で、論文引用件数がトップであるヒントンの論文が引用される数は、2 位から 4 位までの研究者の論文が引用される数を合計したものよりも多い。ヒントンの学生や博士号を取得した研究者たちはアップルやフェイスブック、オープン AI (OpenAI) などの AI 研究所を運営している。ヒントン自身はグーグル・ブレインの AI チームの主席科学者である。実際のところ、翻訳、音声認識、画像認識、ゲームのプレイなど、AI 分野で最近 10 年間に達成された成果のほとんどは、ヒントンの研究とどこかでつながっている。

ヒントンの業績の賜物といえるベクター研究所は、グーグルやウーバー (Uber)、エヌビディア (Nvidia) など、米国およびカナダの企業が AI 技術を商業化するために出資してできた研究セン

Is AI Riding a One-Trick Pony?

ターだ。資金はジェイコブズが思っていたよりもずっと早く集まった。同センターの共同設立者のうち二人がトロント地域の会社を調査したところ、AIの専門家に対する需要はカナダで毎年生まれるAI専門家の数の10倍に達した。ベクター研究所は、深層学習をめぐる現在の世界的な試みである、AI技術を教えて、洗練して、応用して利益を得ようとする動きの中心地と言える。トロントにはデータセンターがいくつも建設され、ビジネスタワーはスタートアップ企業であふれ、学生たちは誰もがAI分野を専攻している。

ベクター研究所のフロアに立ってみると、ガランとして音がよく響くが、もうすぐオフィス機器が運び込まれる予定であり、何かが動き始める現場に立ち会っている印象を受ける。ただし、深層学習において不思議なのは、根幹となる考えの歴史が長いことだ。ヒントンが同僚のデビッド・ラメルハートとロナルド・ウィリアムズと共に書いた大発見の論文が発表されたのは1986年である。この論文はバックプロパゲーション（逆伝搬）と呼ばれる手法について詳しく述べたものだ。

バックプロパゲーションは、プリンストン大学の計量心理学者ジョン・コーエン教授の言葉によれば、「深層学習において文字通りすべての基礎になっているもの」である。

つまるところ、今日のAIとは深層学習のことであり、深層学習とはバックプロパゲーションのことである。これはバックプロパゲーションに30年以上の歴史があることを考えると驚くべきことだ。なぜこのようなことが起こったのか、つまり、ある技術が長い間出番待ちの状態で眠っていて、突然これほど爆発的に必要とされるようになったのかは、知っておく価値がある。なぜなら、バックプロパゲーションの歴史を知れば、AIが現在どんな時期にあるか、特にAIの革命的発展が始まった時点ではなく、おそらくその終わりの時期であることが理解できるようになるだろうからだ。

立証

ヒントンは現在、トロント大学の名誉教授であ

Is AI Riding a One-Trick Pony?

り、グーグルにあるヒントンのオフィスでほとんどの時間を過ごしている。ベクター研究所からヒントンのオフィスまで歩く道すがらは、少なくとも夏の間は、トロントの看板的存在になっている。なぜ英国出身のヒントンがピッツバーグのカーネギー・メロン大学（CMU）勤務の後、1980年代にここに引っ越したかが分かるだろう。

たとえ金融街近くのダウンタウンであっても、一歩外に出るとまるで自然の中に足を踏み入れたような気分になるのだ。たぶん空気に含まれる湿った土の匂いのせいだろう。トロントは森林に覆われた溪谷の上にできた町で、「公園の中の都市」と言われている。トロントの都市化が進むにつれ、自治体は林冠（森林の上層部）を保護するために厳しい規制を課した。飛行機でトロントに近づくと、町の外側がアニメと見まがうほど青々とした木々で覆われているのだ。

**おそらく実際には、革命的発展が始まった時期に
いるのではない。**

トロントはメキシコシティー、ニューヨーク、ロサンゼルスに次ぐ北米第4位の大都市だが、人種的に最も多様である。人口の半分以上がカナダ以外で生まれた人間なのだから。これは町を歩いても感じられる。ハイテク街の人々も、多くの若い白人男性がフード付きスウェットシャツを着ているようなサンフランシスコ的ではなく、もっと国際的だ。医療は無料で公立学校もよく、人々は親切で、政治的には比較的左寄りで安定している。こうしたことが、イラン・コントラ事件に嫌気がさして米国を去ったと言うヒントンなどの人々を引き付けている。昼食時間の少し前にヒントンの会いに行くと、最初にいつもイラン・コントラ事件の話になる。

イラン・コントラ事件は、レーガン政権がイランへの武器売却で得た資金でニカラグアの反共ゲリラ「コントラ」を支援していたことが明るみに出た事件だ。「カーネギー・メロン大学のほとん

どの人々は、米国がニカラグアを侵略するのはまったく当然だと考えていました」とヒントンは言う。「どういうわけか、み

んなニカラグアは米国のものだと思っているようでした」。ヒントンは、最近プロジェクトが大きく進歩したと語る。「非常に優秀な若い技術者がいて、一緒に働いているのです」。彼女の名前はセーラ・セイバー。セイバーはイラン人であり、米国で働くビザを発行してもらえなかったため、ゲーグルのトロント・オフィスが雇ったのだ。

69歳になるヒントンは、ディズニー映画の「BFG: ビッグ・フレンドリー・ジャイアント」のような柔和で痩せた英国風の顔をしている。唇は薄く、耳は大きく、鼻は誇り高い。英国ウィンブルドンの生まれで、子供向け科学本のナレーターのように話す。好奇心を掻き立て、人を引き付け、いろいろなことを熱心に説明してくれる。面白おかしく話すので、ときに少し芸人的でもある。話すときヒントンはずっと立ちっぱなしだ。後で分かったことだが、ヒントンには座っている方が苦痛なのだそう。ヒントンは「2005年6月には座りっぱなしでしたが、あれは間違いでした」と言う。何のことだろうと思っていたら、ヒントンはそれ以降、椎間板に疾患があるのだと説

明してくれた。そのせいで飛行機に乗れない。歯医者では、サーフボードのような奇妙な装置を診療室に持ち込み、そこに寝ないとひびの入った歯根を診てもらえなかったそう。

1980年代、ヒントンは現在と同様に、ニューラル・ネットワークの専門家だった。ニューラル・ネットワークは、人間の脳のニューロン（神経細胞）とシナプス（接合部）で構成するネットワークを非常に簡単に模式化したものである。しかし、当時、ニューラル・ネットワークはAI研究において将来性のない分野だと固く信じられていた。1960年代に開発された最初期のニューラルネットワークであるパーセプトロン（Perceptron）は人間レベルの人工知能への最初のステップとして歓迎されたが、マサチューセッツ工科大学（MIT）のマービン・ミンスキーとセイモア・ペイパートによる1969年の本『パーセプトロン』は、ニューラル・ネットワークが最も基本的な機能しか実行できないことを数学的に証明した。当時のニューラル・ネットワークは入力層と出力層というたった2層のニューロンしかなかったか

らだ。入力ニューロンと出力ニューロンの間に多くの層を持たせれば、ニューラル・ネットワークは理論的には非常に多くの問題を解決できるはずだった。しかし、ネットワークを訓練する方法がわからなかったため、実際にはまったく役に立たなかった。ヒントンなど数人の例外を除いて、パーセプトロン（書籍）のせいで、ほとんどの研究者たちはニューラル・ネットワークを完全に諦めてしまった。

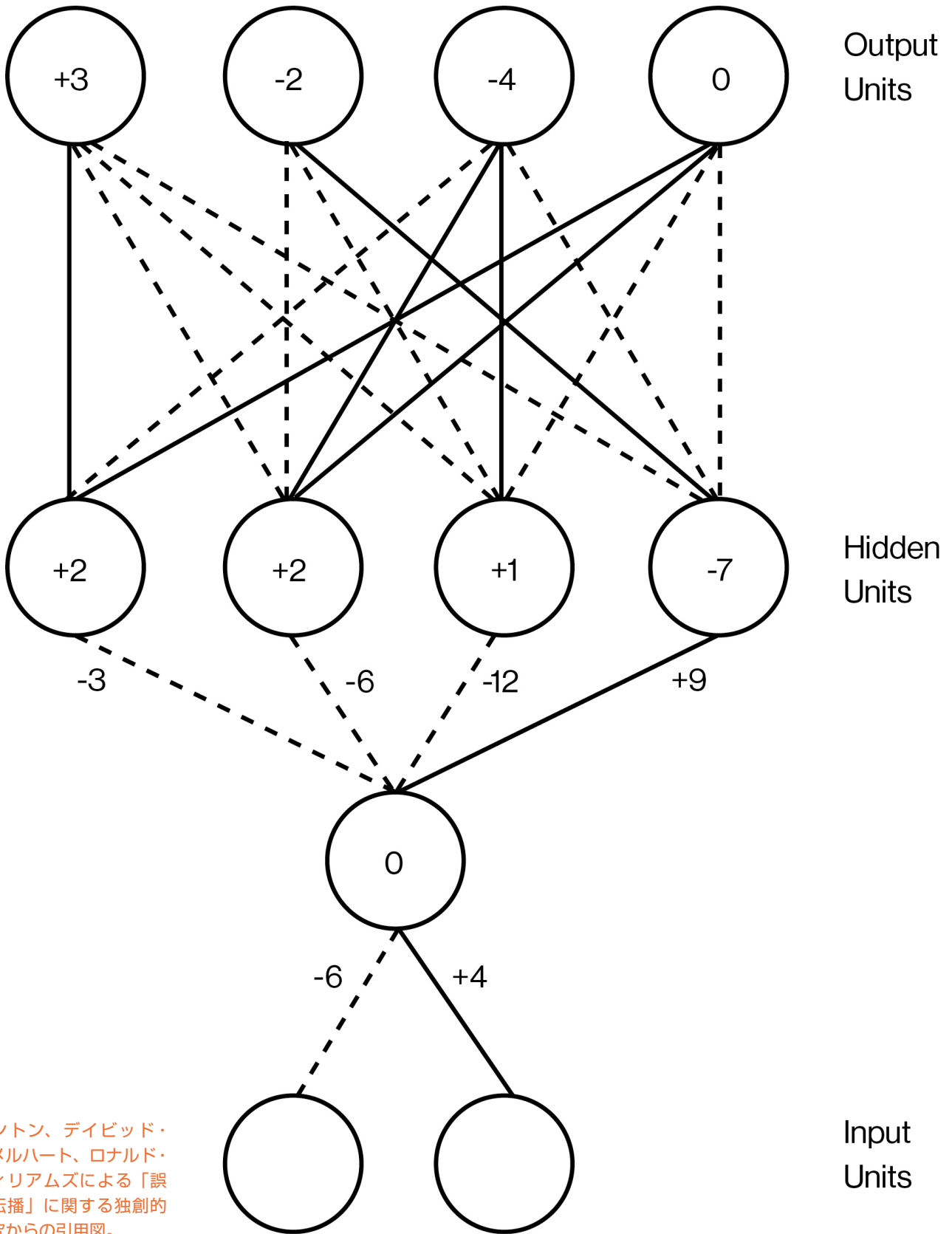
1986年にヒントンがもたらしたブレークスルーは、バックプロパゲーションを使えば3層以上の階層を持つ「深層ニューラル・ネットワーク」を訓練できることを証明したことだ。しかしこの発見の価値が世に広まるためには、コンピュータの処理能力が向上する必要があり、さらに26年が必要だった。ヒントンとトロント大学の学生二人による2012年の論文は、バックプロパゲーションで訓練をした深層ニューラル・ネットワークが、画像認識の最新鋭システムに勝つことを示した。これで「深層学習」の人気が出た。AI業界を知らない人には、AIが一晩で有名

になったように見えるだろうが、ヒントンは長年待ちに待ってようやく報われたのだ。

現実歪曲空間

ニューラル・ネットワークは通常、サンドイッチのように、1つの層が別の層の上に重なった階層状態で描かれる。それぞれの層はわずかな処理能力しかない小さな計算ユニットである人工ニューロンの集まりで構成されている。実際のニューロンと同じように、人工ニューロンは1つ前の層の複数の人工ニューロンから自分自身を励起させる信号を受け取り、自分に接続されている次の層の人工ニューロンに対して相手を励起させる信号を送る。それぞれの人工ニューロンの励起状態はたとえば0.13とか32.39といった数値で表現される。2つの人工ニューロン間の接続部ごとに重要な数字がもう1つある。自分の励起状態をどれだけの強さで相手に伝えるかを決める数値だ。この数値は、脳内のニューロン間のシナプス（結合部）の強さをモデル化したものだ。数値が

Is AI Riding a One-Trick Pony?



ヒントン、デイビッド・ラメルハート、ロナルド・ウィリアムズによる「誤差伝播」に関する独創的研究からの引用図。

Is AI Riding a One-Trick Pony?

大きければ接続は強くなり、あるニューロンの励起状態が次のニューロンの励起に与える影響は強くなる。

深層ニューラル・ネットワークが最も成功した応用例の1つに画像認識がある。ケーブルテレビ局 HBO の番組「シリコンバレー」で、研究チームが画像にホットドッグが写っているかどうかを認識できるプログラムを作った時のシーンを思い出してほしい。こうしたプログラムは今でこそ普通にあるが、10年前には構築は不可能だった。このプログラムを動作させるには、最初のステップとして画像を用意しなければならない。話を簡単にするため、縦も横も 100 ピクセルの小さな白黒の画像だとしよう。この画像をニューラル・ネットワークに入力して、入力層の 1 万 (=100 × 100) 個の人工ニューロンそれぞれの励起量が、対応するピクセルの明るさに等しくなるように設定する。これがサンドイッチの最下層になる。人工ニューロンで画像中の各ピクセルの輝度を表現するわけだ。

次に、1 万個の人工ニューロンで構成する入力

層を、その上の大きな人工ニューロン層（たとえば、数千ニューロンとしよう）に接続し、その層をさらに次の層の同じく数千ニューロンに接続する。これを何層かにわたって繰り返す。最後はサンドイッチの最上層、つまり出力層になる。出力層には人工ニューロンは 2 つしかない。1 つは「ホットドッグである」を表し、もう 1 つは「ホットドッグではない」を表す。この場合、ニューラル・ネットワークを訓練するということは、画像にホットドッグが写っていれば出力層の 2 つのニューロンのうち 1 つめのニューロンだけを励起させ、写っていなければ 2 つめのニューロンだけを励起させるようにすることである。ヒントンの偉大な実績となったバックプロパゲーションは、この訓練を実行するための手法なのだ。

バックプロパゲーションは意外なほど単純だが、膨大なデータを対象とする場合にもっともうまく機能する。それが AI においてビッグデータが非常に重要になるゆえんであり、フェイスブックやゲーム会社がビッグデータを求める理由である。ベクター研究所がカナダ最大級の病院が 4 つも立ち並ぶ通

Is AI Riding a One-Trick Pony?

りにオフィスを構え、それらの病院とデータ共有契約を結ぶことを決めたのもそのためだ。

先ほどの例でいうと、データとは、ホットドッグが写っているか、あるいは写っていない数百万枚の画像のことだ。この際に大事なものは、ホットドッグが写っている画像にラベル付けをしておくことだ。最初にニューラル・ネットワークを作るときは、人工ニューロン間の各接続部の重み付けは無作為になるだろう。つまり、自分の励起量をどれだけ強く次のニューロンに伝えるかを示す数値を無作為に設定する。これは、脳のシナプスが調整されていない状態と同じだ。バックプロパゲーションの目的は、ニューラル・ネットワークが正しく機能するように、重み付けの数値を修正することだ。つまり、ホットドッグの画像を最下層の人工ニューロンに入力したとき、出力層にあたる最上層の「ホットドッグ」ニューロンが強く励起されるようにするのだ。

これで訓練の最初のイメージがつかめたと思う。ここに 100×100 ピクセルのピアノの画像があるとす。まず、合計1万個ある各ピクセル

の輝度を、最下層ニューラル・ネットワークの1万個の人工ニューロンのそれぞれに設定する。すると、隣接する層の人工ニューロン間の接続の強さにしたがって、励起状態が順にニューラル・ネットワークの上の層に広がっていき、最後には、画像にホットドッグが写っているかどうかを示す2つのニューロンだけで構成する最上層に到達する。ピアノの画像なので、理想的には「ホットドッグがある」ニューロンの励起量はゼロで、「ない」ニューロンの励起量は大きくなるはずである。しかし、そうはならなかったとしよう。つまり、ニューラル・ネットワークが画像認識を誤ったとする。バックプロパゲーションとは、こうした場合にニューラル・ネットワークの人工ニューロン間のそれぞれの接続の強度を調整することで、訓練用データを入力したときに表れた出力エラーを修正する手順を示したものだ。

バックプロパゲーションの仕組みは次のとおりだ。まず出力層の2つの人工ニューロンの励起量が、どれだけ間違っていたかを計算する。つまりそれぞれの人工ニューロンの実際の励起量の値

が、あるべき数値とどれだけ誤差があったのかを調べる。それが終わったら、出力層の人工ニューロンに接続されている1層下の各ニューロンに対して、出力層の2ニューロンの誤差にどれだけ関与しているかを計算する。この作業をニューラル・ネットワークの最下層、すなわち入力層に至るまで続ける。この時点で、それぞれの接続部が最終的な誤差にどれだけ関与しているかが判明するので、最後に全体的な誤差が最小になるように各接続部の重み付けを調整する。「バックプロパゲーション」と呼ばれるのは、誤差がニューラル・ネットワーク内で出力層から入力層に向けて逆方向に伝搬（プロパゲーション）していくからである。

驚くべきことに、この作業を数百万から数十億枚の画像に対して実行すると、ニューラル・ネットワークは画像にホットドッグが写っているかどうかをかなり正確に言い当てるようになる。さらに注目すべきなのは、画像認識ニューラル・ネットワークの個々の層が、ある意味で人間の視覚と同じように画像を「見る」ことができ始

めることだ。つまり、最初の層は、境界線があれば励起され、なければ励起されないという方法で物の境界線を検知できるようになるかもしれない。その上の層は境界線の集合、たとえば角を認識できるようになるかもしれない。その上の層は形を見始めるかもしれない。さらにその上の層は「切って開いたパン」と「開けていないパン」のそれぞれに反応するニューロンを持っているかのようになるかもしれない。加えて、ニューラル・ネットワークは、自分自身を組織し直すことができる。すなわち、明示的にプログラミングされていないのに、自らを階層構造に組織化できるのだ。

本当の知能であれば、解決しようとしている問題の条件を多少変えても混乱は起こさない。

これこそがだれもがニューラル・ネットワークに夢中になっている理由だ。ニューラル・ネットワークは単にホットドッグか何かの画像をうまく分類できるだけではない。考えを表現できるようだ。文章を使えばもっとはっきり分かる。何十億、何百億語ものウィキペディアの文章を単純

なニューラルネットワークに入力し、単語ごとに、ある層の各ニューロンの励起に対応した数値から成る長いリストを出力するように訓練したとしよう。これらの数値のそれぞれが複素空間における座標だと考えれば、今していることは本質的に、それぞれの単語に対応する複素空間中の点(いわゆるベクトル)を見つけることだ。このようにニューラル・ネットワークを訓練して、ウィキペディアのページ上にある近くの言葉同士が似た座標を持つようになれば出来上がりだ。すると素晴らしいことが起こっている。似た意味を持つ言葉同士が複素空間内で近くに現れるのだ。つまり、「気が狂った」と「錯乱した」は似た座標位置になり、「3」と「7」も似た座標位置になる。さらに、ベクトル演算をすることで、たとえば「パリ」のベクトルから「フランス」のベクトルを引いたり、「イタリア」のベクトルを足して「ローマ」の近くにしたりできる。その際、ローマとイタリアの関係はパリとフランスの関係と同じだなどとニューラル・ネットワークに指示をする必要はない。

「これは驚くべきことで、衝撃的でもあります」

とヒントンは言う。ニューラル・ネットワークは色々なもの(画像、単語、誰かの話の録音、医学データなど)を取り入れて、それを数学で高次元ベクトル空間と呼ばれるものに変換する。そこでは物と物との距離に、実際の世界の重要な特徴が反映される。ヒントンは、これはまさに人間の脳がしていることそのものだと考えている。「もし、思うとはどういうことか知りたかったら、それを言葉のつながりで表現してみましよう」とヒントンは言う。「『ジョンは“しまった”と思った』と言った場合、『思ったとはどういうことですか。ジョンがそう思ったということはどういう意味ですか』と聞かれても、ジョンの頭の中に引用符の始まりがあり、『しまった』があり、引用符の終わりがあるわけではなく、引用符をすべて取り払った同じ文章があるわけでもありません。ジョンの頭の中には、ニューラル活動の大きなパターンがあるのです」。数学者は、ニューラル活動の大きなパターンを、ベクトル空間内において、各ニューロンの活動が数値に対応し、それぞれの数値が非常に大きなベクトルの座標に対応したものと捉え

ることができる。ヒントンの考えでは、それこそが「思考」なのだ。思考はいわば「ベクトルの舞踏」なのである。

トロント大学の旗艦となる AI 研究所がこの事実になんで名づけられたのは偶然ではない。実はヒントンこそがベクター(ベクトルの英語読み)研究所という名前を思い付いた人間なのだ。

ヒントンが作り上げる、確信と熱意に満ちた現実歪曲空間で話を聞いていると、ベクトルにできないことは何もないと感じてしまう。結局のところ、これまでにヒントンたちが成し遂げた成果を見てみると、自動運転車、がんを検知するコンピューター、話された言葉を聞いて即時翻訳するコンピューターなどである。この魅力的な英国人科学者が、高次元空間における勾配降下法について話す姿をぜひご覧いただきたいものだ。

ただし、「深層学習」システムは時々賢いと思うこともあるが、結局今も脳なしだと感じることもある。テーブルにある山盛りのドーナツの写真を見て自動的に「テーブルの上のドーナツの山」と説明文を付け加えてくれたりすると、おお、世

界が分かっている、と思う。しかし、同じプログラムが女の子が歯を磨いている写真を見て「男の子が野球のバットを持っています」と出力するのを見ると、深層学習システムの理解が本当はどれほど薄っぺらなものが分かる。いや、もし少しでも理解しているとしての話だが。

ニューラル・ネットワークは思考のない、ただのもやもやした(ファジーな)パターンの認識装置に過ぎない。ファジー・パターン認識機としては非常に役に立つので、誰もがこぞってあらゆるソフトウェアに搭載しようとするが、ニューラル・ネットワークはよく言っても、簡単に騙せる、制限付きの知性でしかない。画像認識の深層ニューラル・ネットワークは、ピクセルを1つ変えたり、人間には知覚できない画像ノイズを加えるだけで立ち往生してしまう。実際に、深層学習を適用する新しい方法を考えつくたびにほとんど毎回、深層学習の限界を新たに発見するのである。自動運転車は未知の状況ではナビゲーションに失敗することがある。コンピューターは、世界がどう動くか理解するために常識が必要になる構文を解析し

Is AI Riding a One-Trick Pony?

ジェフリー・ヒント



ようとする困ってしまうのだ。

深層学習はある意味で人間の脳の働きを真似ているが、浅い方法で真似ているに過ぎない。深層学習による知性が時に非常に浅く見えるのは、おそらくそのせいだ。実際のところ、バックプロパゲーションは脳を深く研究して人間の思考自体を復元しようとして発見したものではない。昔ながらの古典的条件付けの実験をして、動物がどうやって試行錯誤しながら学んでいくかをモデル化して生まれたものだ。深層学習の発展の過程で生まれた大きな飛躍には、ほとんどの場合、神経科学についての新しい知見がほとんど含まれていない。飛躍は技術の改良によるものであり、長年にわたる数学と工学の研究によって達成されたものだ。私たちは知性について、ほとんど何も知らないのだ。

トロント大学でヒントンと同じ学部のデビッド・ドゥベノー助教授は、深層学習は物理学以前の段階の工学と同じであると言う。「誰かが論文を書いて『私は橋を作った。橋は壊れなかった』と言います。別の人論文を書いて『私は橋を作っ

たが壊れてしまった。しかし柱を追加したら壊れなかった』と言います。これで柱が話題になります。誰かがアーチを話題に持ち出し、『アーチは素晴らしい』と言います」。物理学を理解していれば「どうすれば橋が壊れず、どうすれば壊れるか分かります」と言うだろう。AIのおかげでこの段階まで実際の理解が進んだのはごく最近のことだという。

ヒントン自身、次のように言っている。「会議のほとんどは小さな変更ばかり時間を使っています。一生懸命考えて『現在の我々の活動に本当に欠けているのは何でしょうか。どこが難しいのでしょうか。そこに焦点を当てようではありませんか』と発言するのは正反対のことをしているのです」。

誇大広告と共に次々に現れる大きな進歩ばかりを目にしていたら、外部から正しく評価するのは難しいかもしれない。しかし AI における最近の一連の成果は、科学的というよりは工学的であり、ティンカリング（いじくりまわすこと）的さえある。どうすれば深層学習システムを改善できる

かは分かり始めてきたものの、深層学習システムがどう動いているのか、果たして人間の心に比肩し得るほど強力な何らかのシステムになるのか、依然としてほとんど何も分かっていない。

バックプロパゲーションから得られるものをほとんど得尽くしたかどうかは問う価値がある。もし得尽くしたというのなら、それはAIの進歩が頭打ちになることを意味しているかもしれない。

忍耐

次に来る大ブーム、つまり今よりずっと自由度の高い知能を持ったコンピュータを作る基礎が何かを知りたいのなら、1980年代にバックプロパゲーションと出会っていなかったら見つけたであろう「何か」に似た研究を調べてみるのがよいだろう。頭のよい人たちは、まだうまくいっていない思いつきに対してもコツコツと努力を続けるものだから。

数カ月前、MITに本部を置く複数機関の共同研究組織「精神・脳・コンピューター研究センター

(Center for Minds, Brains, and Machines)」で、MITの大学院生であるエヤル・デクターが認知科学の論文を発表した。発表が始まる直前、彼の妻エイミー、犬のルビー、娘のスザンナが近くを動き回りながら、発表がうまくいくように言葉をかけていた。スクリーンに映し出されたのはルビーの写真で、隣りに写っているのは赤ん坊のときのスザンナだ。エヤルがスザンナに、自分が写真のどこにいるか指差すように言うと、スザンナは喜んで伸縮式ポインターの先で写真を叩いた。スザンナは母親の後ろについて部屋の出口に向かって、おもちゃのベビーカーを押しながら歩いていき、肩越しに「グッドラック、パパ」と叫んだ。最後の一言は「じゃあね」だった。彼女は2歳だ。

「うまくいかないのは一時的な問題です」

その後、エヤルは素敵な質問で発表を始めた。「みなさん、いかがでしょうか。スザンナは2年の経験で、話すこと、遊ぶこと、話についていく

Is AI Riding a One-Trick Pony?

ことができるようになりました。人間の脳にこれほど優れた学習能力があるのはなぜでしょうか。コンピューターはこれほど速く、これほど自由に学習できるでしょうか。

人間はすでに分かっていることに基づいて新しい現象を理解する。ある領域を小さい部分に分けて、小分けしたものを学習する。エヤルは数学者でありコンピューター・プログラマーなので、タスク（たとえばスフレを料理すること）を非常に複雑なコンピューター・プログラムという形で考える。だが、プログラムの膨大なマイクロ命令（たとえば「肘を30度回転させ、調理台を見る。次に人差し指を伸ばす。それから……」）のすべてを学びながらスフレの作り方を学んでいるわけではない。新しいタスクを学ぶのにいちいちこんなことをしてはたいへんだし、既に知っていることが邪魔になって立ち往生してしまう。そうではなく、プログラムを高レベルのステップとして作成するのだ。たとえば、「卵を割る」や「卵の黄身を取り分けて自身だけにする」などのサブ・プログラムから構成された「卵白を泡立てる」というス

テップだ。

コンピューターは、こうはしない。それこそがコンピューターが脳なしである大きな理由だ。深層学習システムにホットドッグを認識させるには、ホットドッグの写真を4000万枚入力しなければならない。スザンナにホットドッグを認識させるには、ホットドッグを見せればすむ。近いうちにスザンナは、しばしば一緒に使われる複数の単語があることを認識するだけでなく、言語に関するもっと深い理解を得るだろう。スザンナはコンピューターとは違い、世界全体がどう動いているかというモデルを心の中に持っている。「人々がコンピューターに仕事を奪われると言っていますが、私には信じられません」とエヤルは言う。「弁護士の仕事が非常に複雑だからコンピューターが代わりに仕事をできないのではありません。弁護士が書類を読み、人々と話すからできないのです。人間とコンピューターはかけ離れたものなのです」。

本当の知能であれば、解決しようとしている問題の条件を多少変えても混乱は起こさない。エヤ

ルの理論で重要なのは、基本的に言って、コンピューターにわざとそのような混乱を起こさせる方法を実証した点だ。つまりすでに知っていることを新しいタスクに柔軟に適用したり、新しい領域についてほとんど何も知らない状態から専門家の状態に素早くブートストラップ（引き上げ）したりするのだ。

これは基本的にエヤルが「調査・圧縮」アルゴリズムと呼んでいる手順だ。コンピューターにプログラマーに似た機能を持たせて、再利用可能なモジュール式コンポーネントのライブラリを構築させ、それを積み重ねてもっともっと複雑なプログラムを構築させるのである。新しい領域について何も命令を受けなくても、コンピューターは知識を構造化しようとする。それはちょうど、人間の子供が手探りでいろいろ試しながら理解した内容を統合し、さらにいろいろ試すのに似ている。

エヤルの指導教官であるジョシュア・テネンバウム教授は、AI分野で論文を引用されることの多い研究者の一人である。他の研究者たちと話すとき、その会話の半分にはテネンバウム教授の名

前が出てくる。グーグルのAI関連子会社であるディープマインド（DeepMind）の中心的人物の何人かは、博士号取得後にテネンバウム教授の下で研究者として働いた経験がある。ディープマインドは、2016年に複雑なゲームとして知られる碁の世界チャンピオンを負かしてコンピューター科学者たちに衝撃を与えた「アルファ碁（AlphaGo）」を開発した。テネンバウム教授は現在あるスタートアップ企業に関与しており、その企業は自動運転車に基本的な物理知識と他のドライバーたちの動作の意図についてある種の直感を与えようとしている。そうすることで、たとえば、目の前の大型トラックがV字型に折れ曲がったり、誰かが非常に強引に合流してきたなど未知の状況が発生しても、自動運転車がうまく予測して反応できるようにしている。

エヤルの論文はまだこのような実際的な場面に応用されてはいないし、もちろん新聞記事に人間以上の見出しを付けるプログラムにも使われていない。「エヤルが研究している課題は、とにかく極めて難しいのです」とテネンバウム教授は言う。

「解明には何世代も、何世代もかかるでしょう」。

テネンバウム教授の髪は長く、ウェーブがかかっている、白髪になってきている。コーヒーでも飲もうと腰かけたときには、ボタンドアンのシャツを着て黒いズボンを履いていた。テネンバウム教授はバックプロパゲーションの話に注目しており、刺激を得ていると言った。何十年もの間、バックプロパゲーションは何も達成できない、格好がいいだけの数学だった。しかし、コンピューターの処理速度が上がり、コンピューター工学が洗練されるにつれ、突然役に立ち始めた。テネンバウム教授は同じことが、彼や彼の学生の研究にも起こることを望んでいるが、「まだ何十年もかかるかもしれません」と言う。

ヒントンはAIの限界を破るには「コンピューター科学と生物学の間の橋」をかけなければならないと確信している。バックプロパゲーションはこの視点から言うと、生物学にヒントを得たコンピューター科学の勝利である。アイデアは工学分野から出たものではなく、心理学から来ている。そこで今、ヒントンは同じような方法で状況を打

開したいと考えている。

今日のニューラル・ネットワークは大きな平面層でできているが、人間の脳の新皮質にある本物のニューロンは水平の層になっているだけでなく、垂直な柱にもなっている。ヒントンはこの柱が何のためのものか分かっていると言う。たとえば、視覚においては視点が変わっても物を認識するためにこの柱が非常に重要なのだ。このためヒントンは「カプセル」と呼ぶ人工の柱を構築して、その理論を実証しようとしている。今のところ良い結果は出ておらず、カプセルはニューラル・ネットワークの性能を劇的に向上させてはいない。しかしこの状況は30年近く前のバックプロパゲーションと同じ状況だ。

ヒントンは自分の大胆さを笑いながら、カプセル理論について「これは絶対に合っているはずですよ」と言う。「うまくいかないのは一時的な問題です」。

#

ジェイムズ・ソマーズは (James Somers) は

Is AI Riding a One-Trick Pony?

ニューヨーク市に本拠を置く作家兼プログラマーである。MIT テクノロジーレビューへの前回の寄稿は 2015 年 5 ~ 6 月の「Toolkit for the Mind(心のためのツールキット)」。インターネット・スタートアップ企業が使用するプログラミング言語によってどのような影響を受けるか述べたものである。📖

**eムックは、MITテクノロジーレビュー
有料会員限定サービスです。
有料会員はすべてのページ（残り43ページ）を
ダウンロードできます。**

ご購入はこちら



<https://www.technologyreview.jp/insider/pricing/>

No part of this issue may be produced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of KADOKAWA CORPORATION.

本書のいかなる部分も、法令または利用規約に定めのある場合あるいは株式会社 KADOKAWA の書面による許可がある場合を除いて、電子的、光学的、機械的処理によって、あるいは口述記録の形態によっても、製品にしたり、公衆向けか個人用かに関わらず送信したり複製したりすることはできません。