

AIバブルへの警鐘——人工知能への過剰な期待は禁物、今のAIに出来る事と出来ないことを見極めることがAIビジネスの成否を分ける

執筆者

KDDI総研 特別研究員 小林雅一

🕒 記事のポイント

近年、AI（人工知能）の発達と実用化には目覚ましいものがある。たとえば日々、認識精度を上げているGoogleの音声検索、交通・輸送手段などに革命をもたらすと見られる自動運転車、さらにはコンピュータによる同時通訳の自動化など、AIが私達の社会や産業界に与えるインパクトは広く、そして深い。

が、その一方で最近ではAIの能力が過度に喧伝された結果、それに対する期待が真の実力を遥かに上回るまで達し、一種バブル化の様相を呈しつつある。

サマリー

AIはその長い歴史において、過去に2度、現在と同じようなブームを経験した。そして、それがバブル化して弾けた後には、「AIの冬」と呼ばれる長い低迷期が続いた。その時期には投資が冷え込んだだけでなく、研究者がこの分野に見切りをつけて去ってしまうなど、長い目で見ればマイナスに作用した。当時と同じ過ちを繰り返さないためには、どうしたらいいのか？本レポートでは、現代AIの実力を冷静に評価することによってバブル化を防ぎ、これをもってAIの継続的な研究開発と産業化を促す一助としたい。

主な登場者

Google Microsoft Apple Facebook Andrew Ng Yann LeCun Joshua Bengio
SCHAFT IBM Softbank Geoffrey Hinton KAIST HUGO NEDO 産業総合
研究所（産総研）

キーワード

AI Artificial Intelligence 自動運転車 Watson Siri Pepper Bayes Rule
Deep Learning Deep Neural Net Neural Network Machine Learning Hidden
Markov Model Kalman Filter Sparse Coding Stanley Google Brain
DARPA Robotics Challenge (DRC) fat tail black swan DARPA Grand Challenge

地域

世界、日本、米国

<p>Title</p>	<p>Caution against A.I. Bubble: People should never overestimate Artificial Intelligence (A.I.).</p> <p>--- make your A.I. business succeed, you must know what the current A.I. technologies can and can't achieve.</p>
<p>Author</p>	<p>KOBAYASHI, Masakazu (Research Fellow, KDDI Research Institute)</p>
<p>Abstract</p>	<p>The recent progress and commercialization of A.I. technologies has been quite remarkable. For example, Google's voice search is constantly improving its accuracy, autonomous cars are expected to revolutionize our traffic and transportation systems, advances in the computerization of simultaneous interpretation are commonplace, and many more technologies are constantly appearing. Undoubtedly, these numerous applications of A.I. technologies will have a profound and broad impact on our society and a wide variety of industries in the near future.</p> <p>On the other hand, it is also possible to see the hype surrounding A.I. is clearly surpassing its true capabilities, which this article suggests is creating a sort of A.I. bubble. A.I. has already experienced two large booms over its long history. Unfortunately, unfounded exuberance meant these periods of rising expectations both turned to bubbles, which inevitably burst several years later, and lead to very long periods of stagnation, periods of so-called "A.I. Winter." During the dark periods, not only did investment in A.I. technologies drop significantly, to the point of practically disappearing, but also experts of A.I. gave up and left the field one after another, causing long-term damage to the research and development of the A.I. sector as a whole. This current period of growth can be considered a Third Spring for A.I., but we should also be conscious of what can and should be done so that we do not repeat the same mistakes made in the previous eras of rapid advance. In this report, we would like to evaluate and examine the true capabilities of current A.I. technologies in order to prevent the formation of a potential 3rd A.I. bubble, which, in turn, will help facilitate the development and industrialization of A.I..</p>
<p>Players</p>	<p>Google Microsoft Apple Facebook Andrew Ng Yann LeCun Joshua Bengio SCHAFT IBM Softbank Geoffrey Hinton KAIST HUGO NEDO National Institute of Advanced Industrial Science and Technology (AIST)</p>
<p>Keywords</p>	<p>AI Artificial Intelligence autonomous cars Watson Siri Pepper Bayes Rule Deep Learning Deep Neural Net Neural Network Machine Learning Hidden Markov Model Kalman Filter Sparse Coding Stanley Google Brain DARPA Robotics Challenge (DRC) fat tail black swan DARPA Grand Challenge</p>
<p>Region</p>	<p>World Japan U.S.</p>

1 第3次AIブームの到来とバブル化への懸念

昨年から今年にかけてAI（人工知能）、あるいはそれを搭載した次世代ロボットなどへの一般的関心が急速に高まってきた。自動運転車、Watson、Pepper、ドローン、ディープラーニング等々・・・主要各紙をはじめメディア報道で、これらAIやロボット関連の記事を目にしないう日が珍しいほどだ。

筆者が2012年に「AIの春 UI史的観点から見たコンピューティング・パラダイムの転換 (<http://www.kddi-ri.jp/article/RA2012020>)」を発表した当時は、AIに対する日本社会の反応は鈍かった。特に企業関係者の多くは「AI（人工知能）」という言葉は過去に聞いたことがあるにせよ、それが近い将来、産業各界において、次世代ビジネスを構築するためのキー・テクノロジーになると予想した人は極めて少なかった。

それから僅か3年程の間に、AIを取り巻く情勢は夜と昼ほどに様変わりした。今や各種メーカーをはじめ日本企業はこれらの分野への投資を急拡大し、カスタマー・サポートや経営判断にAIが導入されるケースもある。また銀行の営業フロアやホテルの受付などではヒューマノイド（人型ロボット）が顧客を迎え入れ、投資会社などはAIやロボット専門のファンドまで立ち上げようとしている。さらに政府は人工知能の普及に向けた研究拠点を大急ぎで拡充している。

が、つい数年前までの、ほとんど無関心に近いような状態から、AIに対する関心が短期間にもあまりにも急激に盛り上がったせいも、今はむしろオーバーシュート、つまり行き過ぎたブームや過大評価などが、AIについて目立ち始めたように思われる。

この分野の関係者であれば、よくご存じのように、AIは半世紀以上にも及ぶ長い歴史において、「AIの冬」と呼ばれる低迷期を2度経験した。いずれも、それと前後してAIへの関心が異常に高まり、AIに対する期待や莫大な投資などがバブル化した。やがて、その実力がそうした過度の期待に応えられないことが判明すると、バブルが一気に破裂して投資が冷え込み、AIは長期停滞に陥ったのである。

もちろん当時のAIと現在のそれとでは、その実力に天と地ほどの開きがあることは断っておかねばならない。こう言うと、かつて実際にAI開発に携わった研究者・技術者の方々には失礼かもしれないが、少なくとも1980年代までのAIは（過度な期待の有無によらず）実社会では、ほとんど使い物にならなかったのである。

これに対し現在のAIは、実社会や産業界の諸方面で、その確かな実力や将来性の片りんを示している。たとえば騒々しい駅中や雑踏でも驚異的な認識精度を誇るGoogleの音声検索、あるいは走行試験が着々と進み、実用化が数年後に迫った自動運転車など、現在のAI技術は私達が心底「凄い」と思えるような実力を備えている。

つまり今回のAIブームは本物である公算が高い。が、逆にそうであればあるほど、AIに対する期待が日増しに高まり、投資がバブル化する危険性もいつになく増している。特に注意すべきは、現代AIの技術的本質を理解することなく、「これからのAIを使えば何でもできるんだ」とばかりに、実際には実現できそうもないアプリケー

ションへの投資が膨らむことである。

この傾向があまりにも行き過ぎたとき、過去のAIブームやその後のバブル崩壊（AIの冬）と同様の道を辿りかねない。そこで本レポートでは現代AIの技術的ポイントと実力、そして今後の可能性を見極めたい。つまりAIが現在、あるいは近い将来、出来る事と出来ない事を峻別しようと思う。

これによって過去のAIバブルと同じ過ちを繰り返すことを回避できれば、と筆者は考えている。つまり産業各界の関心や投資を、実現可能なAIアプリケーションへと導くこと。これが本レポートの目的である。

2 自然言語処理の現状：Pepper、Siri、Watsonなどの実力

まずコンピュータやロボットなど機械が、人間の言葉を理解するための「自然言語処理」技術を見てみよう。たとえばソフトバンク「Pepper（ペッパー）」のようなヒューマノイド、あるいはApple「Siri（シリ）」のような仮想アシスタント、さらにはIBM「Watson（ワトソン）」のような質疑応答マシンなどが、この技術を搭載している。

自然言語処理はAI（人工知能）を構成する主要要素技術の一つとして、1950年代から現在まで長年に渡って研究が続けられてきた。しかし本当に人間の話す言葉を理解し、それを自由自在に操るような技術はいまだ確立されていない。

少なくとも現時点のSiriやPepperに搭載された自然言語処理の技術は、いずれも「チャットボット」に近いものだ。チャットボットとは、あらかじめ技術者、つまり人間が大量の想定問答集を事前に用意しておき、「ユーザーがこういう質問をしたときには、この答えを返すようにしましょう」、あるいは「こういうキーワードには、こんな反応を返しましょう」というルールを機械に移植した疑似的な会話機能に過ぎない。それは決して、我々人間のように言葉を理解し、操っているわけではないのだ。

それでもSiriの場合、それがiPhone 4Sの標準機能として搭載された当初から「Siriは所詮、チャットボットに過ぎない」という辛口の批評が欧米メディアによって報じられたので、その後の失望感はそれほど大きくはなかった。つまり多くのユーザーはSiriに対し、最初から「どうせ、この程度のモノさ」という低い期待値をもって臨んだので、逆に、このサービスを使ってみた後の失望感も小さくて済んだのである。

これとは対照的にPepperの場合、このロボットが「人間の感情を理解し、人間と自由に会話することによって日々変化し、成長していく」といったソフトバンク側の宣伝文句を、各種メディアが真に受けて、そのまま報じてしまったために、実力以上の期待がPepperに寄せられることになった。

が、そのような能力を（少なくとも現在の）Pepperは持っていないし、そもそもそうした高度なAIは現在、地球上のどこにも存在しない。Pepperに搭載されている

自然言語処理技術は、基本的にはSiriと同様「こういうキーワードにはこういう反応を返しましょう」といったルールが多数移植された疑似的な会話機能である。

2015年に発売が開始されたPepperは当初、銀行の店頭等に置かれ、物珍しさも手伝って顧客からは好感をもって受け入れられたようだ。が、いずれPepperが一般家庭にも普及し、家族の間でこのロボットが本当に人間の言葉を理解していないことが分かった時には、遅かれ早かれ飽きられてしまう恐れがある。

こうした懸念に対し「Pepperは近い将来、IBMの『Watson』や最先端AIの『ディープラーニング (Deep Learning)』を導入することにより、飛躍的に成長するから大丈夫」との見方もある。しかしWatson、あるいはディープラーニングにしても、言葉の意味を本当に理解し、人間と自由に会話するという点については、実は「まだまだ、これから」というレベルにあるのだ。

3 AI開発の歴史：「ルールと論理」から「確率・統計」の重視へと推移

まずWatsonのベースとなっているのは、1980年代に一世を風靡した「ルールベースのAI」や、90年代以降に盛んになった「統計・確率的なパターン解析」などを組み合わせた折衷型のAI技術である。

ここで「ルールベースのAI」とは、たとえば「言語の文法」や「医師の診断ノウハウ」など、人間社会を構成する様々な知的成果物を各種の記号と論理に基づいてルール化し、これをコンピュータやロボットに移植したものだ。こうしたやり方は知的な情報処理を行う最も現実的な方法として、AI研究の黎明期である1950年代後半から研究者の間で支持され、1980年代中盤までAI開発の主流となった方式だ。

しかし1980年代後半から、その限界が見えてくると、この種のAIは急速に廃れた。私達の生きる現実世界は、様々な例外や微妙なニュアンス、そして偶発的な出来事などで満ち溢れているため、杓子定規のルールと固定的な論理に従うAI技術では太刀打ちできなかったのである。

これに代わって1990年代から勢いを増してきたのが、統計・確率的なパターン解析をAIに持ち込むという方式である。そのベースとなったのが「ベイズ理論」と総称される一連の確率理論だ。この名称は18世紀の英国の牧師・数学者Thomas Bayes（トーマス・ベイズ）が考案した「ベイズ定理 (Bayes Rule)」に由来する。ベイズ定理は、最初適当に決めた確率（事前確率）をある種の実験、観測、あるいは測定などによって、より正確な確率（事後確率）へと改良する——その手続きを記した定理だ。

その長所は、大量のデータが集まってくるのに伴い、物事の因果関係や相関関係などを記述する確率がどんどん改良されていくことだ（それは「統計・確率的な誤差を小さくする」と言い換えることもできる）。たとえばGoogleが自社サイト上で提供している機械翻訳サービス「Google翻訳」などが、その恰好の事例だ。

Google翻訳では、ウェブ上に存在する様々な言語の対訳文書をコンピュータが消

化して、異なる言語間での対照作業を行う。それは肝心の言語の意味を全く理解することなく、たとえば「I love you」は確率55%で「私は貴方を愛している」、確率45%で「私は貴方を憎んでいる」といった暫定案（事前確率）を算出する。ここから、さらに多くの対訳文書を読み込んで、その度に前述のベイズ定理を適用して、より正確な確率（事後確率）へと改良する。高速プロセッサによって、この処理を毎秒何度も繰り返すことにより、正しい翻訳文の確率はどんどん高まり、誤った方の確率は逆に低下していく。最終的に、「I love you」は（確率99%で）「私は貴方を愛している」という翻訳結果が導かれる。

このような方法は、スマートフォンなどの音声認識に使われている「隠れマルコフ・モデル (Hidden Markov Model)」、あるいは自動運転車の外界認識に使われている「カルマン・フィルター (Kalman Filter)」など、現在注目されているAI製品の多くで採用されている。いずれも、かなり高い性能を示していることから、こうした統計・確率的な手法は現代AIの主流となっている。

さてWatsonはこの「統計・確率型のAI」とそれ以前の主流だった「ルールベースのAI」を組み合わせた折衷型のAIである。80年代後半に一旦は「使い物にならない」との烙印を押された「ルールと論理を重視するAI」でも、その後、技術に磨きをかけると同時に、統計確率的なパターン解析技術などと組み合わせれば有効であることが再認識された（そこには勿論、いわゆる「ムーアの法則」によるコンピュータ・プロセッサの処理速度や記憶容量のアップなど、ハードウェア技術の向上も大きく寄与している）。

ただし、そこには厳しい制約が伴う。それは、こうした「折衷型のAI」はその用途や使用環境を絞り込む必要があるということだ。逆に言うと、どんな目的や状況でも使える「汎用の人工知能」にはなり得ないのである。

実際、Watsonがそもそも開発される切っ掛けになった「米国の人気クイズ番組に出演して、クイズに回答する」といった用途は、そうした制約条件にぴったりマッチしている。またIBMが後にWatsonのビジネス転用を図った際も、具体的なアプリケーションとして選んだのは「医師の診断支援」や「コールセンターにおけるカスタマー・サポート業務」、あるいは「企業の経営診断」など。つまり、それぞれ用途を一つに絞り込んだ上で、それに向けてWatsonをカスタマイズしている。

これとは対照的に今後、家庭に導入されるであろうPepperは、そもそも何のために使われるのか予め分からない。またPepperと家族との間で交わされる会話も、どんな方向に展開していくか予想できない。こうした場合、ロボットは本当に人間の言葉を理解して話す必要があるが、近い将来、Pepperに導入されるかもしれないWatsonは、そこまで高度で汎用的な自然言語処理の能力を持っていないのだ。

さらに付け加えると、実は現時点のPepperや前述のSiriも単なるチャットボットというよりは、折衷型AI的な機能も多かれ少なかれ備えている。が、その技術レベルを比較すると、明らかにWatsonの方が勝っている。これはIBMが長年培ったAI研究の蓄積、また同社がWatsonを基幹ビジネスに育て上げるために同業他社とは桁違いの開発資金を投じたことなどを照らし合わせれば、当然でもある。が、それでも（前

述のように) Watsonの用途を一つの分野に絞り込むなど、厳しい制約条件を課さざるを得なかった。

逆にWatsonよりも性能的に劣るPepperやSiriが、(何の制約条件も課せられない) 一般家庭における自由な会話などに使われた場合、折角用意されたルールベースの自然言語処理技術なども役に立たなくなり、結果的にはチャットボットとほとんど同じモノになってしまうのである。

4 ディープラーニングの限界：「パターン認識」と「意味理解」の間に横たわる広くて深い谷

では、現在最先端のAIである「ディープラーニング」ではどうか？ディープラーニングは1950年代から営々と研究開発が続けられてきた「ニューラルネット」の最新形である。

人間の脳を構成する無数のニューロン（神経細胞）のネットワークを（非常に単純化した形ではあるが）工学的に模倣したとされる「ニューラルネット」は、当初から「論理的な理解力」や「動作速度」などについて性能上の限界が指摘され、長らくAIの世界では傍流の道を歩んできた。

しかし、この分野の大御所として知られるカナダ・トロント大学のGeoffrey Hinton（ジェフリー・ヒントン）教授らを中心に、2006年頃から脳科学の研究成果を本格的に導入することで、ニューラルネットはそれらの限界をある程度まで打破し、ついにブレイクした。これは今、「ディープ・ニューラルネット」あるいは「ディープラーニング」などと呼ばれ、人工知能の世界で一大ブームを巻き起こしている。

ディープラーニングは今までのところ「画像認識」や「音声認識」など、いわゆるパターン認識の分野で極めて高い性能を示している。この大きな理由の一つは、ディープラーニングが本格的な脳科学の研究成果、中でも大脳の視角・聴覚野などに共通する認識メカニズム「スパース・コーディング (Sparse Coding)」を導入したことにある。

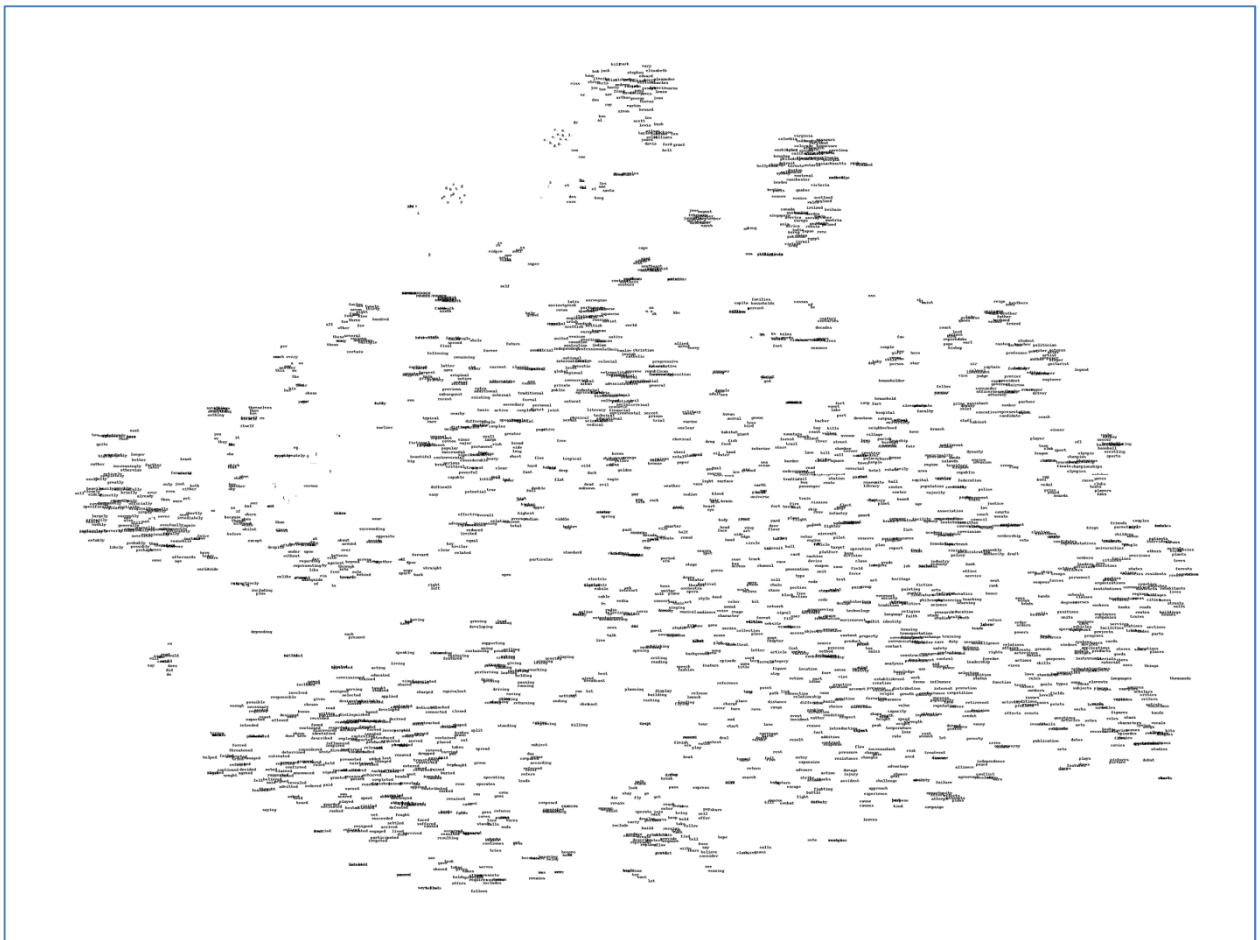
つまりディープラーニングとは「人間がモノを見たり、音を聞き取ったりする際の、脳の認識メカニズムを参考にしたAI」なのだ。そうである以上、それが画像・音声認識で高い性能を示すのは合点がいく。これが今後さらに進化し、いずれは人間の頭脳のように何でもできる「汎用AI」へと進化することを期待する声すら上がっている。

そうした中で今、AI研究者の当面の関心事項は、このディープラーニングを前述の「自然言語処理」、つまり「言葉（単なる単語だけではなく、センテンスや会話まで含む）の意味を理解して、これを自由自在に操る」という目的に転用できないかということだ。たとえばディープラーニングの権威であるFacebookのAI研究所所長のYann LeCun（ヤン・ルカン）氏や、カナダ・モントリオール大学教授のJoshua Bengio（ヨシュア・ベンジオ）氏らが、ここ数年、このテーマに取り組んでいる。

これについて彼ら研究者は、口では「やれる」と強気なのだが、少なくとも、こ

れまでは目覚ましい成果は報告されていない。今のところディープラーニングが自然言語を処理する能力は、同類の単語やセンテンスを確率的に分類することだけだ。その様子を、ディープラーニングの第一人者である、前述のジェフリー・ヒントン氏（トロント大学教授）が、オンライン大学Courseraにおける講義の中で、以下の【図表1】によって示している。

【図表1】 Joseph Turian's map of 2500 English words produced by using t-SNE on the word feature vectors learned by Collobert & Weston, ICML 2008



<http://www.cs.toronto.edu/~hinton/turian.png>

この図はディープラーニングがウィキペディア内の多数の英語文書を読み込み、それを解析して英単語の中で最もよく使われる数千個の単語を2次元マップにまとめたものだ。が、このままでは文字が小さい上に、英語なので分かり難いかもしれない。そこで、これを日本語に訳して説明すると、次のようになる。

たとえば「東京」「パリ」「ニューヨーク」・・・などの名詞は、いずれも世界的に有名な都市を指しているので、高い確率で同一クラスター（集団）へと分類される。あるいは「見る」「目にする」「目撃する」・・・などの動詞も、ほぼ同じ行為を指しているため、やはり高い確率で同じクラスターに分類される。

それはまた、広大な言語空間（言語マップ）において「東京、パリ、ニューヨー

ク・・・」つまり都市名のクラスターが、「日本、フランス、米国・・・」つまり国名のクラスターと、確率的に近い位置関係にあることも示している。

が、今のディープラーニングにやれるのは、ここまでである。つまり肝心の言葉の意味を理解することなく、ビッグデータ解析によって単語やセンテンスの相関関係を確率的に算出するだけなのだ（この言語マップでは単語の相関関係のみを示している）。この程度の能力では、せいぜいチャット・ボット、つまりユーザーが入力した言葉と相関性の高い言葉を返してくる疑似AIシステムにしか使えないだろう。それは現在のSiriやPepperがやれることと大差がないのである。

現在のディープラーニングはどれほど鼻真目に見ても、人間と同じように言語を理解しているとは言えない。私見だが、画像・音声認識のような単なる知覚能力と、「言語を理解して操る」といった高次の知的能力との間には、途方もなく広くて深い谷が横たわっているのではあるまいか。

もちろん「画像認識」や「音声認識」など、すでにディープラーニングの能力が証明された分野では、今後とも認識精度は順調に向上していくだろう。が、その先にある、もっと大きな目標、つまり「自然言語の理解と操作」を端緒とする「汎用AI」への道のりは、そう平坦ではないと予想される。

5 自動運転車にディープ・ニューラルネットを搭載

たとえ実現されるとしても、汎用AIはまだ大分先の話だ。これとは対照的に、現実的な射程圏内にある自動運転車に対してディープラーニングが果たす役割を考えてみよう。

現在、テスト走行の段階にある自動運転車には「カルマン・フィルター」と呼ばれる技術が搭載されている。この技術は前述のベイズ定理に基づく、「統計・確率型のAI」をロボット工学に応用したものだ。カルマン・フィルターは、自動運転車が自らのセンサーから取得したデータを解析するためのアルゴリズムだ。

自動運転車にはミリ波レーダーやレーザー・レンジ・ファインダーをはじめ多彩なセンサー類が搭載されている。自動運転車はこれらセンサーを使って、自らの周囲にある様々なモノや移動体（たとえば障害物や歩行者、あるいは他のクルマなど）の居場所を測定する。

しかし最初に1回測定しただけでは、あるモノの居場所を正確に把握することはできない（つまり測定値の統計的誤差が大きい）。そこでもう一度、センサーで測定し、そこにベイズ定理を適用することで誤差を若干小さくする。この作業を毎秒数百万回も再帰的に繰り返すことで、誤差を徐々に最小値へと収束させる。これによって自動運転車は周囲のモノの居場所を正確に把握し、安全に移動することができる。これがカルマン・フィルターの原理だ。

しかし、このやり方では、走行中のクルマの前に突然何かが飛び出してきたことまでは分かるが、それが「(飼い主の手から偶然離れてしまった) 大型犬か、それと

も（風に流された）大き目のポリ袋か？」までの違いは分からない。大型犬なら急ブレーキをかけてでも停車した方がいいが、ポリ袋なら（急ブレーキをかけて後続車に追突されるよりは）むしろ、そのまま行き過ぎる方がいい。しかしカルマン・フィルター、つまり統計・確率的なAIでは、そのように異なる物体の区別が出来ないので、今後、製品化に向けて重大な危険性が残る。

こうした限界を突破する上で、今、最も注目されているのがディープラーニングだ。その最大の長所は、前述の通り、脳科学の成果、特に大脳視角野の認識メカニズムを取り入れることにより、人間と同じような認識方法に近づいたことだ。たとえば独アウディは、自社で開発中の自動運転車に同技術を搭載すべく研究開発を進めている。その主な理由は、前述の「大型犬とポリ袋の違い」など重要な課題を、脳を模倣した認識メカニズムによってクリアできると考えているからだ。

しかし少なくとも現時点のディープラーニング（ディープ・ニューラルネット）では、ドラム缶とポリ袋の本質的な違いを識別することは難しい。できるのは、あくまでも外見上の違いを、より正確に認識することだけだ。その理由は2012年に米スタンフォード大学のアンドリュー・エン（Andrew Ng）准教授らが、Googleと共同開発したディープ・ニューラルネット「Google Brain（グーグル・ブレイン）」を見れば明らかだ。

Google Brainは、動画サイト「YouTube（ユーチューブ）」上にある無数のビデオからキャプチャされた、数千万枚に及ぶ静止画を元に学習することにより、自力で「猫」や「人」「自動車」などの概念を獲得して、コンピュータ画面上に、これらのイメージを（ぼんやりとはあるが）描き出した。

Google Brainのようなニューラルネット、つまりコンピュータが自力で何らかの概念を獲得したことの衝撃は大きかった。2012年にニューヨーク・タイムズなど米メディアがエン氏らの研究成果を報じると、このニュースは瞬く間に世界に広がった。これを契機にディープラーニング（ディープ・ニューラルネット）は世界中のAI研究者の間で大きな注目を浴びるようになった。

しかしニューラルネットが獲得した「猫」や「人」の顔などの概念は、私達人間が抱く概念とは大きく異なる。たとえば「猫」を例にとれば、私達が「猫」について抱いている概念は多面的である。つまり猫の外見（視覚的概念）に加え、その毛並みに触れたときのふわっとした感触（触覚的概念）、その鳴き声（聴覚的概念）、さらには猫の気まぐれな性格（内面的概念）に至るまで、実に多面的な概念を私達は猫について抱いている。これが猫について本当に理解するということだ。

これに対しエン氏らの開発したニューラルネットが獲得した概念とは、あくまで猫の外見、つまり視覚的な概念だけだ。これは彼らのシステムがこの世界について学習する際の教材としたのが、YouTube上の静止画だけであったことを考えれば当然だ。ここから導かれるのは、どう工夫したところで視覚的概念しかあり得ない。

実はアウディが自動運転車でやろうとしていることも、これと同じなのだ。それはクルマに搭載されたビデオカメラで撮影した映像をディープラーニングで解析し

て、周囲にある物体を識別しようとしている。つまり視覚的な概念を形成するということだ。

しかし視覚的概念だけでは、自動運転車が安全に走行するには不十分なのである。それは何故か？再び「大型犬とポリ袋の違い」に話を戻そう。現在のディープラーニングは確かに、両者の形状についてであれば、カルマンフィルター（純粋に統計・確率的なAI）よりは正確にその違いを識別できる。が、それはあくまで視覚的な概念に基づいている。

それ以外の概念、たとえば「大型犬は重い、ポリ袋は軽い」「大型犬は生き物だが、ポリ袋は単なるモノだ」・・・等々の本質的な違いは全く理解していない。このため、「走行中の自動運転車の前に風で流されてきたポリ袋は無視した方がいいが、（重い生き物である）大型犬が飛び出して来た場合、犬は死んでしまうし車も破損するから急停車した方がいい」という判断は、（たとえディープラーニングが搭載されたとしても）クルマが自力ではできないはずだ。そうした判断基準は技術者、つまり人間がクルマに教えてあげねばならないのだ。

6 世界観や一般常識を持たないAIがクルマを運転することは許されるのか？

もっと深刻なケースを考えてみよう。たとえば、繁華街にある細い裏道を走行中のクルマの前に突然、物陰から小さな子供が飛び出してきた。ブレーキを踏んでも、すぐには停車できない。となると、ハンドルを左右いずれかに切らなければ当然、クルマは子供をはねてしまう。しかし逆にハンドルを左右どちらかに切れば、クルマは道路沿いにあるビルに突っ込むことになる。

仮にドライバー（人間）がクルマを運転していたとすれば、この人は咄嗟にハンドルを切ってビルに突っ込むだろう。当然、クルマもビルも破損するし、ドライバーも恐らく軽傷を負うが、子供は助かるからだ。

これに対し自動運転車であれば、正反対の行動を選ぶかもしれない。つまり子供よりもビルの方が大きく危険な物体であると判断し、あえてハンドルを切らずに子供をはねてしまうかもしれないのだ。

両者の違いはどこにあるのか？それを「子供」に関する本質的、かつ多面的な概念を有しているか否かの違いだ。たとえば「目の前にいる子供は私と同じ人間だが、ビルは単なるコンクリートの塊だ」「子供は生きているが、ビルは生命を持たないモノだ」「子供は死んだら生き返らないが、ビルは壊れても、修理したり建て代えれば元に戻る」・・・等々、実に多面的な概念を、私達人間は心の内側に（ほとんど無意識のうちに）抱いて生きている。だからこそ、いざというときにビルよりも子供を優先する決断を咄嗟に下すのである。

これに対し自動運転車（に搭載されたディープラーニングというAI）は前述の通り、あくまでも視覚的概念しか有していない。たとえ視覚的に「子供」と「ビル」の区別ができたとしても、その背後にある多面的かつ本質的な概念を有していない以上、自動運転車が人間のドライバーと同じ高度で倫理的な決断を下せるとは限ら

ないのだ。

要するに自動運転車を今後、実用化する上で最も重要なポイントは「この世界を私達人間のように本質的に理解することなく、単に視覚的な概念しか持たないAIにクルマを運転させていいのか？」という問題に帰結するのである。

どんな国でも、人が自動車の運転免許を取得するには法定年齢がある。たとえば日本の場合、人がクルマを運転するには、18歳に達するのを待たねばならない。それは、その年齢になれば、この人が自分の生きる世界を自分なりの経験に基づいて深く理解しているからだ。18年の人生を通じて培われた多面的で豊かな世界観や一般常識というものが、この人がクルマを安全に運転する上で、無意識のうちに役立っている。

一方、自動運転車にはそれがない。自動運転車がこの世界を理解する方法は、そこに搭載された各種センサーが計測したデータだけが頼りだ。これを元に周囲にある様々なモノの視覚的概念を形成するのが精一杯である。現在、最先端のAIであるディープラーニングをもってしても、(少なくとも現時点では)それが限界なのだ。

もしも、それ以上のことをさせなければ、我々人間が自動運転車に対し、この世界を構成する多面的な概念や世界観、あるいは常識といったものを教えてあげねばならない。たとえば「子供はビルよりも大切な存在だから、いざというときには子供の方を優先しなさい」というルールを自動運転車に移植する必要がある。

そして、それは単に「子供とビル」の違いにとどまらない。この世界を構成する、ありとあらゆるモノや事象について、そうしたルールを教え込まねばならないのである。そうしない限り、本当に安全な運転はできないはずだが、それらは数え上げれば切りがないので、技術者（人間）にとって事実上、不可能とも言える作業になるのは間違いない。

7 「位置を特定できれば衝突しない」という前提は楽観的に過ぎる

とは言え、周囲の物体や移動体（障害物、他のクルマ、歩行者など）の居場所を正確に把握できさえすれば、あとはそれに衝突しないようにすれば済む。そういう割り切った考え方もあるし、実際、既にテスト運転の段階にこぎ着けた、現在の自動運転車はそうした楽観的な設計思想に基づいて開発されている。そこに搭載されているカルマン・フィルターは、現在のディープラーニングがやっているような「視覚的概念の形成」さえ出来ない。こうした統計確率型のAIにできることは、車載センサーが計測したデータをもとに、周囲にあるモノの位置を正確に把握することだけだ。

が、本当にそれで済むのだろうか？と言うのも、この世界では、普通に考えれば、とても起きそうに無いことが、実は意外に高い確率で起きるからだ。それは「ファットテール (fat tail)」、あるいは「ブラック・スワン (black swan)」などと呼ばれる統計的現象として、よく知られている。たとえば「高速道を走行中に、自分と同じレーンを逆走するクルマに出くわす」といった事態がファットテールに該当する。

ファットテールは統計的な現象なので、通常は見えにくい。たとえばGoogleは自社で開発した自動運転車について、既に100万マイル以上に及ぶ走行試験を行ったが、これまでのところ重大な事故を起こしていないという。が、それは走行試験に使われた自動運転車の台数がせいぜい数十台にとどまっているからだ。つまり、この程度の台数であれば、ファットテールという統計的現象が効いてこないのである。

逆に、今後、自動運転車が実用化され、世界中で数百万台から数千万台もの自動運転車が走り始めたとすれば、そこではファットテール、つまり「常識では起こり得ないような事が、実は意外に高い確率で起きる」という統計的現象が目に見えて効いてくるはずだ。そうした突発事態、ないしは異常事態が発生した場合、自動運転車が周囲の移動体の位置をどれほど正確に把握、あるいは予測したところで、恐らく衝突は避けられないだろう。

つまり「物体の位置を正確に特定できれば衝突しない」という設計思想は楽観的に過ぎるのだ。どれほど技術の粋を凝らしたところで、実際には衝突はある程度の頻度で必ず起きるのである。その際には私達人間が有している豊かな世界観や常識、多面的な概念といったものが、被害を最小限に食い止める上で必要になってくる。ディープラーニングを筆頭とする現在のAIはそれを備えていない。現在、テスト走行までこぎ着けた自動運転車ではあるが、今後の実用化に当たっては、もう一段の技術的進化が必要とされるだろう。

そうした中でGoogleや世界の主要自動車メーカーの多くは、2020年頃に自動運転車を実用化するという目標を掲げている。が、中には米Tesla Motors（テスラ・モーターズ）や米GMのように、製品化の前倒しを表明するメーカーも現れている。

たとえばテスラは「2015年の夏には、（高速道や駐車場など使用環境を限定した上で）自動運転を実用化する」と予告している。が、今年の夏も終わりを迎える中、同社が自動運転を実用化したとのニュースは聞こえてこない。

もしも同社が自動運転の実用化の難しさを改めて痛感し、それをやむなくキャンセルしたのであれば、それはむしろ望ましいことと筆者は考えている。と言うのも、前述の理由から、現時点の技術レベルで自動運転車を実用化するのは危険と思われるからだ。自動運転のような画期的技術の普及に歴史的必然はない。つまり本来なら普及してもおかしくないはずの技術が、功を焦って出足につまずいたばかりに、永久にお蔵入りという悲劇も起り得る。

特に自動運転車の場合、人の命が関わってくる。安全性に疑問符が残る中で製品化を急ぎ、その結果、世界中で重大事故が多発すれば、消費者は最早、自動運転車の安全性を信じてくれないだろう。一旦、そうなってしまえば、後からメーカーがいくら安全対策を強化したところで手遅れである。自動運転車は二度と普及しないだろう。自動運転車は今後の製品化に向けた、最後の詰めが実は最も重要なのである。

8 サービス・ロボット開発の一里塚となる競技会

この自動運転車と並んで、今、日本の産業界で最も高い関心を集めているのは恐らく次世代ロボットだろう。政府は成長戦略の一環にロボット産業を組み入れ、2020年までにその市場を現在の3倍前後に当たる2兆4000億～2兆9000億円程度まで拡大したいとしている。

特に大きな成長を見込んでいるのが、従来の産業用ロボットではなく「サービス・ロボット」と総称される多彩な次世代ロボットである。これは介護・福祉をはじめ、人手不足が深刻なサービス産業などの現場で人に代わって、あるいは人と共同で各種作業をするロボットだ。こうしたロボットの普及戦略を練るために、日本政府は産業界の有識者から構成される「ロボット革命実現会議」も立ち上げている。

が、そうした次世代ロボットを実現するための技術的ハードルは高い。今後登場する各種サービス・ロボットには、従来の産業用ロボットとは比べ物にならないほど高度な能力が必要とされるからだ。たとえば周囲の環境を正確に認識して器用に動いたり、近くにいる人間に危害を与えること無く、安全に共同作業をするといった能力だ。

現在、日本をはじめ世界的な開発が進むサービス・ロボットは、これらの能力をどの程度まで育んでいるのだろうか？それを推定するメルクマールとなるのが、今年6月に米DARPA（国防高等研究計画局）が実施した「DARPA Robotics Challenge（DRC）」だ。

DRCは元々、2011年3月11日に発生した東日本大震災と福島第一原発の事故を契機に計画されたとされる。原発の事故現場のような危険な環境下で、人に代わって復旧・解体作業などにあたるヒューマノイドの開発を促すことを目的としている。

DRCの予選競技会は2013年12月に米フロリダ州で開催され、出場した世界各国のロボットの中から、日本のロボット・ベンチャー「シャフト（SCHAFT）」の開発したヒューマノイドが見事、予選を首位で通過した。この直前に同社が米Googleに買収されたことも相まって、シャフトやDRC、ひいては次世代ロボットに対する日本国内の関心が一気に高まった。

ただし、このシャフトをはじめDRC予選競技会に出場したロボット達の性能は、とても人に代わって作業を行えるようなレベルにはなかった。彼らは事故現場を想定した8種類の競技——「壁にドリルで穴を開ける」「梯子を上る」「バルブを閉める」等々——をこなしたが、いずれの動きも非常に鈍く、競技の途中でリタイヤするロボットも多かった。

中には、動いたと思ったら、すぐに止まってしまい、そのまま何時まで経っても動かないロボットも少なくなかった。たとえ動いたとしても彼らの動きが余りに遅いので、競技を観戦するギャラリーの中には「草が育つ様子を見守るようだ」という感想を漏らす人もいたほどだ。

ロボット達の動きが鈍いのは、外界の認識能力が低いことに起因する。DRCに参

戦するロボットは基本的に無線でリモコン操作されるが、彼らが競技中、主催者であるDARPAが意図的にリモコン用の電波を妨害する。この際にはロボットが自律的に動くことが求められる。ここで必要とされるのは、ロボットが自身に搭載されたビデオカメラなどを通して得た外界の映像を解析して、周囲の状況を正確に認識する能力だ。これも一種のAIだが、こうした外界認識に膨大な時間を消費してしまうため、ロボット達の動作が極端に遅くなってしまふのだ。

外界認識能力の低さはロボットの動きの不正確さ、ひいては転び易さなどにもつながる。実際、DRC予選競技会に出場したロボット達は、いずれも背中に転倒防止用の「命綱」がつけられていた。ロボット達はどれも重量が数百キロ・グラムと重いので、転ぶと壊れてしまうからだ。以上のことから見て、少なくとも一昨年の予選の段階では、事故や災害の現場で縦横無尽に活躍する次世代ロボットなど、夢のまた夢というのが偽らざる印象だった。

9 サービス・ロボットの開発は自動運転車よりも難しい

それから1年半後となる2015年6月5日（金）～6日にかけて、DRCの本戦競技会「DARPA Robotics Challenge Finals 2015」が米カリフォルニア州のイベント会場「フェアプレックス（Fairplex）」で開催された。この本戦には、2013年の予選を通過したチームを中心に、日本や欧米、韓国など世界各国から25チームが出場した（ただしGoogleに買収された日本のシャフトは、何らかの理由で本戦への出場を見送った）。この1年半の間にロボット達の性能がどこまで改善されたか、という点に関係者の関心が集まった。

その競技結果を報告する前に、予選と本戦との違いを指摘しておく必要がある。予選に比べて、本戦ではロボットに課せられる競技の条件が厳しくなった。まずロボットが8種類の競技全体に費やすことのできる総合時間が、大幅に短縮されて1時間となった（2013年の予選では、各々の競技に最大30分を費やすことが許されていた）。また本戦では、予選時のようにロボットの背中に命綱をつけることが許されなくなった。

こうした条件下で実施された本戦の結果だが、競技中に時間切れで失格となったり、ボタンと転んで壊れてしまうロボットが続出した。またロボット達が動くスピードは「草が育つよう」と揶揄された予選会よりは若干改善されたが、それでもやはり遅いと言わざるを得なかった。

この本戦で優勝を飾ったのは、韓国KAIST（Korea Advanced Institute of Science and Technology）が開発したロボット「HUGO」だが、これとて全8種目の競技をやり遂げるまでに44分を要した。日本からは5チームが参戦したが、NEDO・産総研チームのロボットが10位に入ったのが最高の成績だった。このため「韓国にロボット技術で差をつけられた」と懸念する向きもあるが、それは実は、あまり気にする必要がないだろう。

と言うのも、NEDO・産総研チームのロボットが全競技をやり遂げるのに要した時

間は52分。これと韓国HUGOが費やした44分との違いをどう見るかだが、所詮は「ドングリの背比べ」だろう。人間が同じ8種類の競技をこなしたとすれば、全部で10分もかからないからだ。

そもそもHUGOが好成績を収めることができた主な理由は、ロボット制作上のちょっとした工夫によるところが大きい。人型ロボットのHUGOが平坦な地面を移動するときには、両脚の膝を折り曲げて、ちょうど地面に跪くような格好になる。その際、曲げた膝と爪先の両方から車輪が飛び出す仕組みになっている（【図表2】）。

【図表2】 韓国KAISTのHUGO

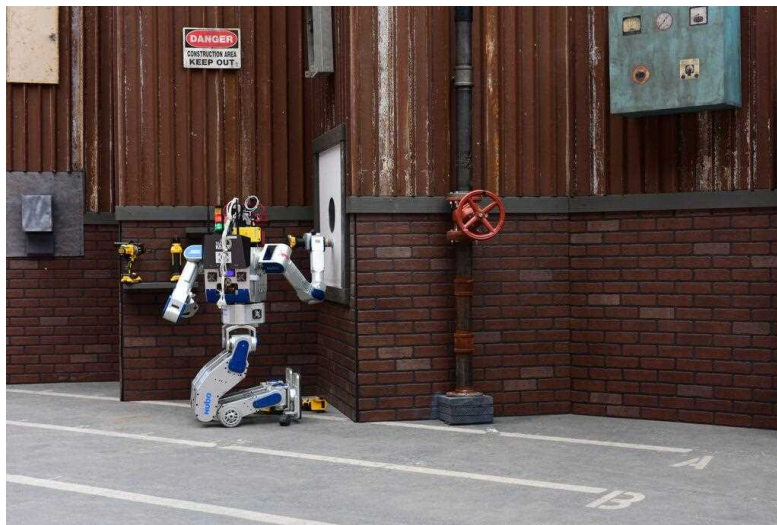


写真 : DARPA <http://www.theroboticschallenge.org/media>

これによって、通常の人型ロボットにとって難しい二足歩行ではなく、4輪車のような形で進むことができる。この方法だと滅多に転ぶことがないので、素早くスムーズに移動できるのだ。勿論、これはこれで意味のある工夫だが、逆に言うと「ロボットにディープラーニングのような高度なAIを搭載することによって、その外界認識能力を飛躍的に高めて転ばないようにする」といった本質的な技術革新が為されたわけではないのだ。

つまり今回のロボット競技会で本当に注目すべき点は、「日本が韓国に差をつけられた」ということではなく、むしろ2013年末の予選競技会と2015年の本戦との間で、出場したロボット達の性能に、それほど劇的な改善が見られなかったことだ（現在、こうした次世代ロボットの外界認識能力を高めるため、ロボットにディープラーニング技術を搭載する動きが米UCLAなどで始まっているが、まだ基礎研究の段階だ）。

実は今回のDRCには、かつて同じくDARPAが主催した「DARPA Grand Challenge」という前例がある。2004～2005年にかけて開催された、このイベントは自動運転車のレースであった。2004年の予選レースでは、米国南西部の砂漠に設けられた約100キロのコースを完走した自動運転車は皆無だったが、2005年の本戦レースでは約210キロに延長されたコースを5台が完走。そこで優勝したスタンフォード大学チームの「Stanley（スタンレイ）」は、現在Googleが自社で開発中の自動運転車のベースとなっている。

このように予選から本戦にかけて驚異的に進化する可能性も残されていたため、今回のDRCにも相応の期待が集まっていた。が、結果は前述の通り、期待外れに終わった。原発の事故現場などで作業する災害対策ロボット、あるいはそこから派生する各種サービス・ロボットなどは、自動運転車よりも遥かに実現が難しいと見た方がいい。日本が注力する次世代ロボットの開発や実用化には、この先、相応の時間がかかることを覚悟しておくべきだろう。

10 まとめと考察

AI（人工知能）の研究開発は1950年代に始まるが、近年、大脳感覚野に共通する認識アルゴリズム「スパース・コーディング」をはじめ、脳科学（神経科学）の研究成果を本格的に導入。これによって画像や音声など、いわゆるパターン認識の分野でブレークスルーを成し遂げた。これは今、ディープ・ニューラルネット、あるいはディープラーニングなどと呼ばれ、AIの世界で一大ブームを巻き起こしている。

が、一方でこの技術を自然言語処理、つまり「言葉の意味を理解して操作する」といった目的に応用するのは容易ではない。恐らくパターン認識のような単なる知覚能力から、さらに一步踏み込んで「意味を理解し、考えて、話す」といった、より高次の知的領域に踏み込もうとすると、そこにはかなり厚い壁が立ちふさがっているようだ。私見だが、この壁を突破するには「大脳言語野のメカニズム解明」など理論的な脳科学のさらなる発達、そしてAIとの融合が求められてくるのではなかろうか。

一方、ビジネス、つまり産業としてのAIを見た場合、既に確立された技術をどう効果的に生かすかがポイントになるだろう。たとえば自然言語処理の場合、「人間のようにならば言葉の意味を理解して、自由自在に相手と会話する」といったAIは（少なくとも近い将来では）望めそうもない。が、この技術の用途を狭い範囲に限定すれば、従来の「ルールベースAI」と「統計・確率的なパターン解析」を組み合わせる方法でも十分役に立つと見られている。

たとえばIBMの質疑応答マシン「Watson」はそうした折衷型AIの代表として知られる。これは今、日本の金融機関などでコールセンターの顧客サポート等に導入が図られているが、このように技術の用途や使用環境を思い切って絞り込むことがAIビジネス成功の鍵となるだろう。もちろん「Watson」という技術自体は「顧客サポート」以外にも、「医師の診断支援」や「企業の経営判断」など多方面への応用が図られている。が、一旦、それら各々の分野への導入が決まった後は、その分野に限定して使用されることになり、そのためのカスタム化が予め為される。これが最大のポイントだ。

これに対しソフトバンク「Pepper」のように一般家庭に導入され、そこで家族と自由な会話が想定されるようなAIロボットは（少なくとも今のままでは）若干苦戦することが予想される。それは確かに当初は物珍しさも手伝って、世間の関心を惹くかもしれない。が、やがてこうしたロボットに人間と会話するような本物の知能が備わっていないことが判明したときに、ユーザーが興奮するマイナス要因とし

て跳ね返ってくるだろう。特にPepperのように、実用性よりも「楽しさ」や「面白さ」を前面に押し出したマーケティングが行われた場合、そうした失望や反動はより大きくなる。

Pepperのような次世代ロボットは、マーケティングの段階でAIに対して膨れ上がった一般消費者の期待値を、むしろ本格発売を前にいかに下げて、実際のレベルにマッチさせるかがビジネスの成否を分けることになるだろう。そこでは「インターネットによる各種アプリのダウンロード」などに焦点を移行し、数は少なくてもいいから「これだけは使い物になる便利な機能（アプリ）だ」といった実用性をアピールすることの方がむしろ得策なのではないか。

Pepperの、さらに先にある次世代ロボット、たとえば一般家庭で家事や雑用を引き受けたり、過酷で危険な作業現場などで人間の重労働を引き受けてくれるヒューマノイドなどは、実用化までに大分時間がかかりそうだ。確かに一時、ディープラーニングのようなパターン認識、つまり知覚能力に優れたAIを搭載すれば、そうした汎用ヒューマノイドの実現が一気に近づくと期待されたこともあった。が、前述のロボット競技会「DARPA Robotics Challenge」に出場した次世代ロボットのパフォーマンスを見る限り、そこへと至る道はまだ相当長いようだ。

むしろ現時点で確立されたAI技術をどう既存のロボットに活用するか。ここから次世代ロボット実用化への道を探る方が賢明だろう。たとえば画像認識に長けたディープラーニング技術を既存の産業用ロボットに搭載すれば、ロボットが相応の外界認識能力を持つようになる。これによって、産業ロボットと従業員との接触による人身事故などの件数を減らすことができる。それはまた産業用ロボットの応用範囲を広げることにもつながるだろう。

1 1 結論

絶え間ない脳科学の発達、特に近年の理論面の研究成果によって、現在の最先端AIは人間のように「見て、聞く」という知覚能力を得ることができた。しかし、ここから、さらに一步踏み込んで、AIが「言葉を理解し、自由に操る」、あるいは「考えて、行動する」という能力を獲得するまでには、もう一段の科学的ブレークスルーと相応の時間を必要とする。このように、今の先端AIに出来ることと出来ないことを正確に見極め、出来ることを最大限、今後の製品開発に活用することが、AIの事業化を成功に導く最大のポイントとなる。

また、もしも現時点で、「人との会話」のように一際高度な知的機能を製品に実装したいのであれば、「記号処理型のAI」や「統計確率的なパターン解析」など従来型のAI技術を組み合わせて使う必要がある。が、この種の人工知能は人間に備わっているような本物の知能ではなく、あくまで疑似的な知能に過ぎない。ここに能力的な限界があるため、その用途や使用環境を思い切って絞り込む必要がある。また、そうした疑似的かつ折衷型のAIでは将来的な発展性も限られている。産業各界の関係者は、これらの点に留意してAI関連事業の構築を進めるべきだろう。

【執筆者プロフィール】

氏名：小林 雅一（こばやし まさかず）

所属：KDDI総研

専門：メディア・IT・コンテンツ産業の調査研究

経歴：東京大学大学院理学系研究科を終了後、雑誌記者などを経てアメリカに留学。ボストン大学でマスコミ論を専攻し、ニューヨークで新聞社勤務。慶應義塾大学メディア・コミュニケーション研究所などで教鞭をとった後、現職。

主な著書：

『クラウドからAIへ アップル、グーグル、フェイスブックの次なる主戦場』（朝日新書）

『日本企業復活へのHTML5戦略』（光文社）

『スマートフォンのすすめ—手のひらのクラウドで未来を生きる』（ぱる出版）

『ウェブ進化 最終形 「HTML5」が世界を変える』（朝日新書）

『モバイル・コンピューティング』（PHP研究所）

『社員監視時代』（光文社ペーパーバックス）

『欧米メディア・知日派の日本論』（光文社ペーパーバックス）

ほか多数。