

コグニティブ・コンピューティングへの飽くなき探究心

IBM東京基礎研究所では、近年、コグニティブ・コンピューティングに関する研究開発にさまざまな角度から取り組んでいます。その中から今回は、2人の研究員、浅川智恵子と恐神貴行の取り組みをご紹介します。

実世界で身の回りを認識できる日を夢見て

視覚障がい者は、オンライン・ネットワークの世界では合成音声を通じてさまざまな情報を自ら得られるようになりました。しかし、実世界に身を置いたとたん、身の回りの情報を認識したり、行きたい場所に行ったりすることが困難になります。

実世界では近年、スマートフォンに代表されるモバイル・コンピューティングやIoT(Internet of Things)によるセンサー・ネットワークの広がりなど、身の回りの情報を認識するために必要な技術が着実に進展しています。視覚障がい者であるIBMフェローの浅川智恵子は、こうした世の中の流れを、視覚障がい者の実世界でのアクセシビリティ向上に生かしたいと考えていました。2年程前に浅川は、米国カーネギーメロン大学(CMU)のロボティクス・インスティテュートを訪ねる機会に恵まれ、CMUの研究者と語り合ううちに、

自分の夢である実世界を“再び見る”、つまり技術の力を借りて自力で身の回りを認識できるようになると確信しました。

浅川は昨年秋からIBMフェロー兼CMU客員教授として米国IBMへ赴任し、視覚障がい者が取り戻したいと思っている機能を補完し、自立と生活の質を向上するための支援技術「コグニティブ・アシスタンス」の研究開発にCMUと共に取り組んでいます。

最初の成果[1]は今年10月半ばに、IBM Bluemix上に公開したオープン・プラットフォームとツールキット第1弾(<http://hulop.mybluemix.net/>)、屋内外のナビゲーションをスマートフォン向けパイロット版ナビゲーション・アプリ「NavCog」の無償公開という形で開花しました。浅川がNavCogを試している様子は、YouTubeで公開しています(図1)。

ビデオには、現在CMUが研究を進めている顔認識の試作技術が盛り込まれています。浅川が装着してい

るヘッドセットのカメラが、向こうから歩いて来る人が誰でどんな表情をしているかを認識し、イヤホンを通じてその情報を知らせることで、浅川は相手に先駆けて声をかけることができます。

将来、こうした支援技術によって視覚障がい者が実世界で身の回りを認識できるようになれば、歩道からそれないようにナビゲーション・システムが注意してくれたり、行きたい場所に1人で行けるようになったり、コミュニケーションの幅がより広がったりするでしょう。またこの技術は、病院、工場、工事現場などで使用する用品や資材をナビゲーション技術を応用して管理するなど、さまざまなイノベーションにつながる可能性を秘めています。

今回公開したオープン・プラットフォームを通じて、実世界のアクセシビリティ向上の取り組みに、より多くの方々が参加してくれることを願って、浅川と研究チームはさらに研究を進めています。

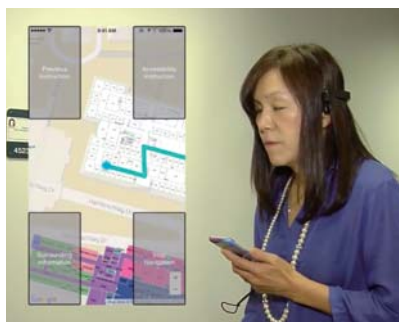


図1. 浅川がNavCogを試している様子
YouTube : <https://youtu.be/w17IS0AJHjU>

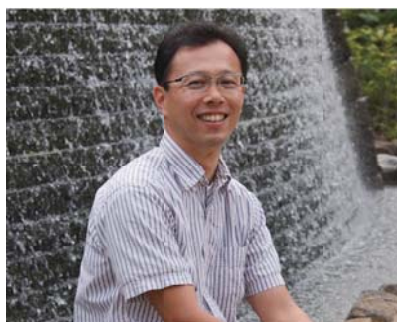


図2. 数理学で神経細胞を模倣した恐神研究員

```
00111001100111011110100010011001111
01000010010010010000100010100101000
010000100000100100001:0010100001000
00110010000010011100101010100001110
00001010000010010000100110100001000
00001010010010010000100010100101000
01110001100111011110100010011001111
```

図3. SCIENCEのビットマップ・イメージ
(1をなぞるとSCIENCEが表される)

人間の五感で捉えられない 概念を理解したい

ハードウェアの研究開発に興味を持つ電子工学科の学生だった恐神貴行研究員(図2)が数理科学の世界に足を踏み入れるようになったきっかけは、IBM東京基礎研究所の採用面接で受けた「計算の複雑さの理論に関する学問を深化させる」ことへのアドバイスでした。

その後、数理科学の面白さに没頭していき、「人間は参照する情報に確率的な要素や不確実性が伴うと、なぜ必ずしも合理的な判断をすることができないのだろうか」「予測など人間の五感で捉えられない抽象的な概念の中に潜んでいる変調を、数理科学で解き明かすことはできないのだろうか」といった好奇心が、恐神を意思決定プロセスの研究へ、そして人工ニューラル・ネットワークの研究へと導いたのです。

1940年代に提唱された生物の神経細胞の学習則であるヘブ則は、1980年代にボルツマンマシンなどの人工ニューラル・ネットワークによってその数理的な基礎付けが与えられました。1990年代にはヘブ則の時間的要素を緻密にしたスパイク時間依存可塑性(STDP)という学習則が確認されましたが、これまでその数理的な基礎付けはまだ明らかになっていませんでした。



図4. 「SCIENCE」の出だしのパターンを入力し、DyBMが2通りの並べ方で「SCIENCE」を再現[4]

恐神は、数年前から共同研究者の大塚博士[2]とともに、生物の脳で行われるSTDP則による学習の仕組みに対して基礎理論を与える研究を進め、このたび、STDP則が自然な学習則として導出される人工ニューラル・ネットワーク「動的ボルツマンマシン(以下、DyBM)」[3]の開発に初めて成功しました。これにより、STDP則の人工ニューラル・ネットワークへの工学的応用の道が切り開かれたのです。

恐神のチームは、DyBMがSTDP則の工学的応用を実証するために、まず7個の人工的なニューロン(神経細胞)からなるDyBMに、「SCIENCE」という文字列をビットマップで表現した画像(図3)を学習させました。

この文字画像を、左から右へ、また右から左へと文字画像をなぞるように時系列パターンとして記憶するよう学習を重ねさせるうちに、例えば私たちがクロスワードパズルを解くように、DyBMは一部の画像情報を手掛かりに全体像を思い出し、どの方向からでも正しく「SCIENCE」という文字画像を再現することに成功しました(図4)。

次に、20個の人工的なニューロンを使って20ビットの人類の進化の画像を学習させると、DyBMはわずか19秒で正しいイメージを認識、再現することができました(図5)。

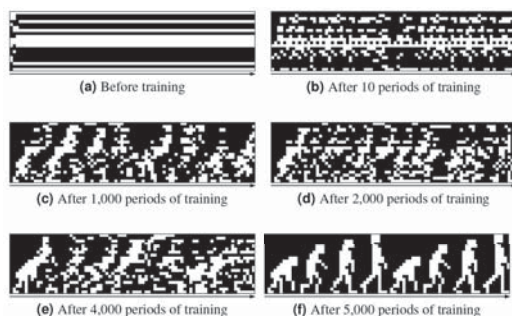


図5. 人類の進化のイメージをDyBMが学習[4]

さらに、12個の人工的なニューロンからなるDyBMが、有名なドイツ民謡「山の音楽家」を90万回の学習をした後に、譜面の再現に成功しました(図6)。ここでは神経細胞はそれぞれ、12音階(ファ、ソ、ラ、シ、ド、レ、ミ、ファ、ソ、ラ、シ、ド)の4分音符の学習を担当するよう割り当てられましたが、ニューロンの数を増やせばさまざまな音階や音符の長さを理解し再現することができます。

この研究は、左脳型の情報処理を行う既存のコンピューターが不得意とするパターン認識や、連想記憶といった右脳型の情報処理を得意とする次世代コグニティブ・システムの実現に寄与すると言われています。

恐神には信念があります。DyBMの特性を応用し、車載カメラでのリスクの予測、文章や動画の中の誤りの発見、ロボットやIoT機器の制御など、「人々の生活に役立つ技術を開発すること」です。恐神はその信念を胸に研究を続けます。

[参考文献]

- [1] <http://www.ibm.com/jp/press/2015/10/2603.html>
- [2] <https://sites.google.com/site/motsukapublic/>
- [3] "Seven neurons memorizing sequences of alphabetical images via spike-timing dependent plasticity" by Takayuki Osogami and Makoto Otsuka appeared in Scientific Reports of Nature Publishing Group on September 16, 2015, DOI: 10.1038/SREP14149.
本研究成果は、国立研究開発法人科学技術振興機構の戦略的創造研究推進事業CRESTにおいて得られたものです。
- [4] Credit: Scientific Reports



図6. DyBMが音楽を学習[4]
YouTube : <https://youtu.be/vCcR3ii9FkA>