

新たな次元へとシフトする——さらに深化する コミュニケーション科学の取り組み

NTTコミュニケーション科学基礎研究所は、人と人、あるいはコンピュータと人との間の「ここまで伝わる」コミュニケーションの実現をめざして、人間と情報の本質に迫る基礎理論の構築と社会に変革をもたらすコア技術の創出に取り組んでいます。本稿では、音声音響処理、対話処理、人間情報科学、スポーツ脳科学、機械学習と最適化など、一段と深化するコミュニケーション科学の取り組みの一端を紹介します。

やまだ たけし

山田 武士

NTTコミュニケーション科学基礎研究所 所長

AI技術の発展

最近のAI（人工知能）技術の発展にはめざましいものがあります。特に深層学習の発展のおかげで、本来人間が得意で、なかなかコンピュータが追いつけなかった音声認識や画像認識、自然言語処理などに関しても、人間に迫る精度が達成できるようになってきています。NTT研究所としては、これらの最先端の技術を活用し、直面する課題にうまく適用することも重要です。求められてもいます。しかしこれらの技術の普及に伴って、これまでの延長ではない、新たな次元を切り拓く取り組みに常に挑戦し、研究テーマを大胆にシフトさせていくことが必要となります。

NTTコミュニケーション科学基礎研究所（CS研）は、人と人、あるいはコンピュータと人との「ここまで伝わる」コミュニケーションの実現をめざして、人間と情報の本質に迫る基礎理論の構築と社会に変革をもたらすコア技術の創出に取り組んでいます。研究分野としては、NTTグループのAI技術corevo®を構成する4つのAIであるAgent-AI、Heart-Touching-AI、Ambient-AI、Network-AI⁽¹⁾を支

えるべく、主に、「メディア処理」「人間情報科学」「データと機械学習」を中心に取り組んできており、最近では「スポーツ脳科学」にも力を入れています（図1）。本特集では、一段と深化するコミュニケーション科学の取り組みの一端を紹介します。

人間の能力に近づく音声音響処理技術

corevo®を構成するAgent-AIは、人間とコンピュータのインタラクションを支援するためのAIです。CS研ではAgent-AIの基盤となる、音声・映像などのメディア処理や自然言語処理に取り組んでいます。中でも音声に関しては、スマートフォンや最近ではAIスピーカなどの急速な普及によって、接話マイクに向かって1人が話す状況での音声認識がすでに日常的に使われるレベルまで成熟してきました。研究開発の潮流は、テーブルを囲んで複数人がマイクから離れて自由に話す状況下での音声認識へとシフトしつつあります。そのためには音声認識そのものの性能向上も必要ですが、それ以上に騒がしい環境でも周囲の雑音を取り除く雑音抑圧技術や、室内での壁や床などへの反射による残響を除去する、

残響除去技術等の音声強調技術との組合せが重要となります。CS研の音声認識は、これらの組合せにより2015年に行われた国際技術評価（CHiME-3）において、参加25機関中で第1位の精度を実現しました⁽²⁾。これらの技術を用いて、まだ実験段階ではありますが、複数人で行う会議の議事録を自動的に作成することにも挑戦しています。このような会議でのやり取りや、パーティでの歓談などにおいて、複数の人が同時に話したり、背景に音楽がかかっていたりしても、人間は「聞きたい」と思った人の声の特徴を選び分けて、その人が話す内容を聞き取ること、すなわち「選択的聴取」を普段から自然に行っています。本特集記事『SpeakerBeam: 聞きたい人の声に耳を傾けるコンピュータ——深層学習に基づく音声の選択的聴取』では、深層学習を用いて、人間が得意とする選択的聴取をコンピュータで実現する技術について紹介します。さらに人間は、音声を聞いただけでその場の情景をある程度頭の中に思い浮かべることができます。このようなことをコンピュータで実現するのが、クロスメディア情景分析です。深層学習はさまざまなメディアで共通に適用できるので、単一

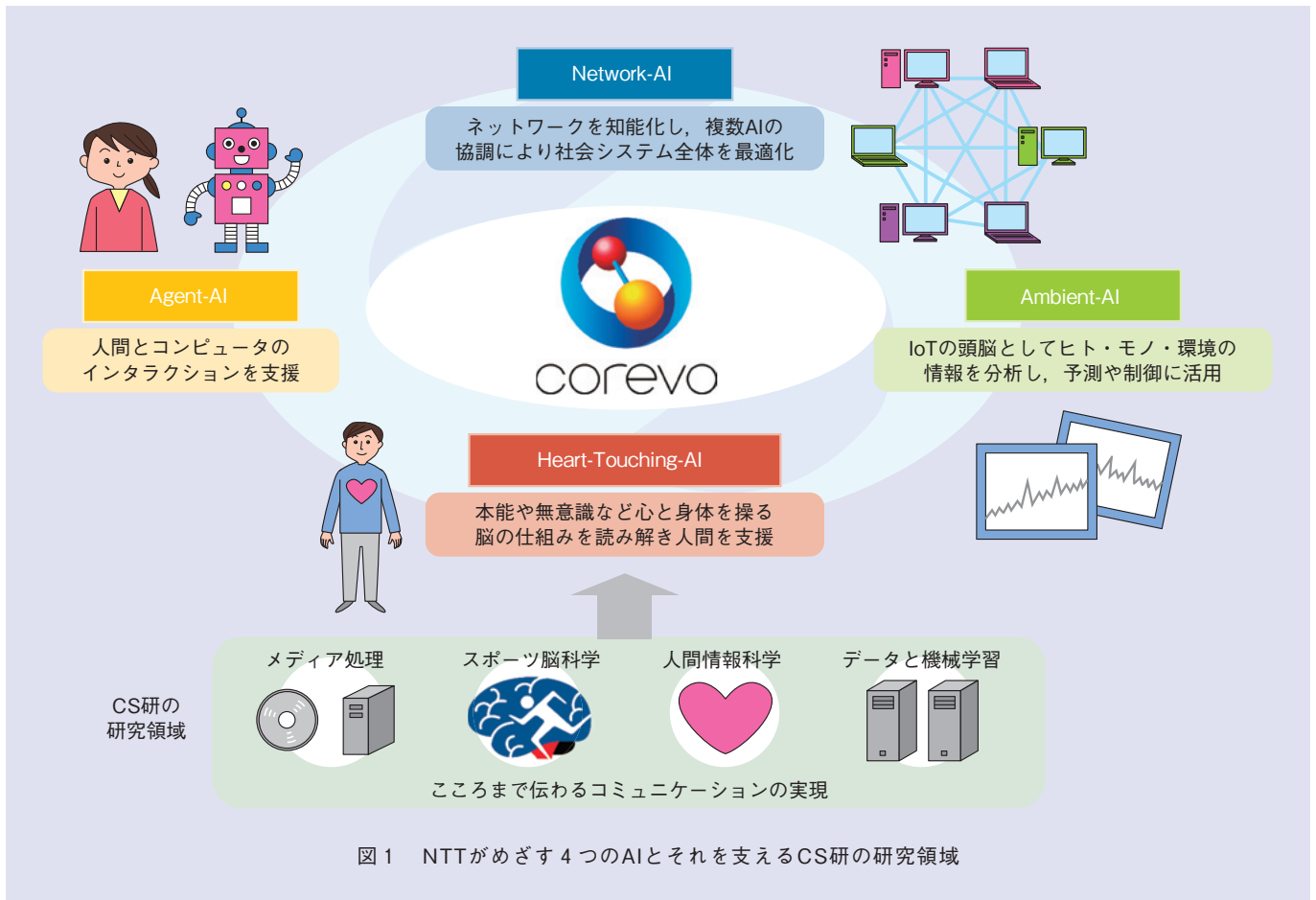


図1 NTTがめざす4つのAIとそれを支えるCS研の研究領域

メディアを超えたクロスモーダルなメディア処理、すなわち、複数のメディアを組み合わせた、相補的に変換したり、などが可能となります。クロスメディア情景分析については本特集記事『音から画像認識結果を予測するクロスメディア情景分析技術』で紹介します。

人間とAIとのギャップの解消

このように特定の場面ではコンピュータの能力は人間に近づきつつあります。しかし、AIの性能が複雑な人間の脳を超えるほどに進歩するにはもっと時間がかかりそうです。一方で、人間はときに「振り込め詐欺」にも簡単に騙されてしまいます。これは人間が持つ「認知バイアス」が関係しています。例えば人間は、自分がすでに持つ

ている信念、期待、仮説に都合が良い証拠ばかり探したり解釈したりする傾向があり、かかってきた電話は息子からだと思ってしまう、矛盾点にも気が付きません。これは認知バイアスの1つである「確証バイアス」と呼ばれます⁽³⁾。また人間は、一見知能がありそうな振る舞いをするものに対しては、勝手に擬人化し、知能を感じてしまうこともよく知られています。これは1960年代にマサチューセッツ工科大学で開発されたコンピュータプログラムにちなんで、ELIZA効果と呼ばれています⁽⁴⁾。ペットの犬は「うちの子はなんて賢いんだ！」と飼い主が思うほどには実は賢くないものです(図2)。一方、例えば有名な錯視例として、チェッカーシャドウ錯視というのが知られていますが、これは、人間

はありのままの物理量を「見ている」わけではないことを、文字どおり、目の当たりにさせてくれます⁽⁵⁾。

複雑かつ不完全な人間と今のところ未熟で限定的なAI、この両者のギャップを埋めるためには、認知バイアスや錯覚などをインタフェースやフィードバックにうまく利用することが鍵となります。前者に関連して、CS研ではロボットとの対話処理の研究を進めており、雑談対話から議論対話へと対話処理のレベルも深化していますが、同時に、これまでの1体のロボットとの対話から、複数(2体)のロボットとの対話へとシフトしています。1人の人間と対話する場合、相手も1体のロボットで十分に思えますが、複数のロボットを用いてうまく役割分担させることで人間の認知バイアスを利用し、

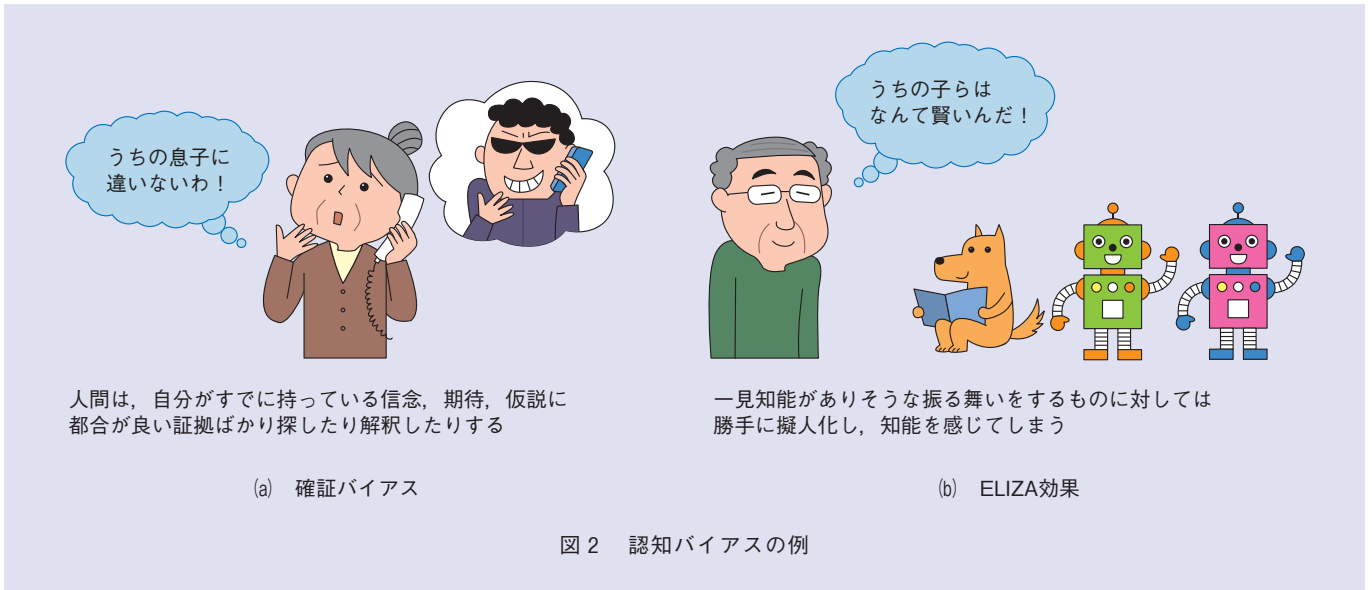


図2 認知バイアスの例

人間にとってより「賢く」見せることが可能となり、たとえば音声認識に失敗したり、対話の文脈が破綻したとしても、人間にとって自然な対話をより長く続けられることが分かってきました⁽⁶⁾。また、後者に関してはCS研ではこれまで人間の錯覚を利用したインタフェースとして、引っ張られる錯覚を生じさせるデバイスである「ぶるなび」や印刷した絵や写真に光を当てるだけで動き出して見える「変幻灯」などを生み出してきました。本特集記事『「浮像（うくぞう）」——影を利用して印刷物に見かけの奥行きを与える光投影技術』では、やはり人間の視覚の錯覚を利用した研究成果を紹介し

潜在的な脳の働きの解明

CS研では「ここまで伝わる」コミュニケーションを実現するために、「視覚」「聴覚」「運動感覚」といった人間の基本的な感覚に関する「潜在的な脳の働き」の解明に注力しています。前述の錯覚を利用したインタフェースもこのような研究から派生して生まれたものです。これら人間情報科学の基礎研究はcorevo[®]を構成するHeart-

Touching-AI、すなわち、本能や無意識など心と身体を操る脳の仕組みを読み解き人間を支援するAIの基盤となります。CS研がめざす「ここまで伝わる」コミュニケーションは、まさに「こころ(Heart)に触れる(Touching)AI」に通じています。

最近ではHeart-Touching-AIの新たなテーマとして、スポーツ脳科学への取り組みが進んでいます⁽⁷⁾。これは「潜在的な脳の働き」を解き明かす脳科学の知見と最先端の情報通信技術を活かし、ウェアラブルセンサやVR (Virtual Reality)、機械学習などを活用した分析により、「勝つための脳をきたえる」研究です。すなわち「身体を強くする」ための研究ではなく、あくまでも「身体をいかに最適に操るか」、あるいは「精神状態をいかにうまくコントロールするか」に主眼を置いて研究をしています。例えば、聴覚は視覚より時間分解能が高いことを利用して、体の各部位の筋活動を計測し、それをリアルタイムに音に変換することで、直感的なフィードバックを実現する方法を提案しています。一方で、スポーツに限らず、人間が普段の生活の中で、自ら

の心身の潜在能力を最大限に発揮できるように支援することも重要です。本特集記事『ウェルビーイングを測る、知る、育む』では、このような人間のウェルビーイングという、一見、定性的でとらえどころのない課題を人間科学の立場から定量的に扱い、向上のための設計指針を特定するチャレンジについて紹介します。

機械学習と最適化

機械学習技術を用いると、大量のデータから人間の専門家でも気付かない特徴的なパターンを自動的に見つけ出すことができます。街中を歩く人々の群れなど、集団としての人の振る舞いは複雑ですが、それぞれが目的を持っており、一定のパターンがあります。NTTでは、リアルタイムに観測した人の流れのデータから、近未来の混雑・渋滞リスクを予測し、人々の目的を阻害しない範囲で効率良くリスクを回避する誘導策をオンラインで自動的に導出する技術の研究開発に取り組んでいます。ここでは、マルチエージェントシミュレーションなどとともに、ベイズ最適化が用いられています⁽⁸⁾。

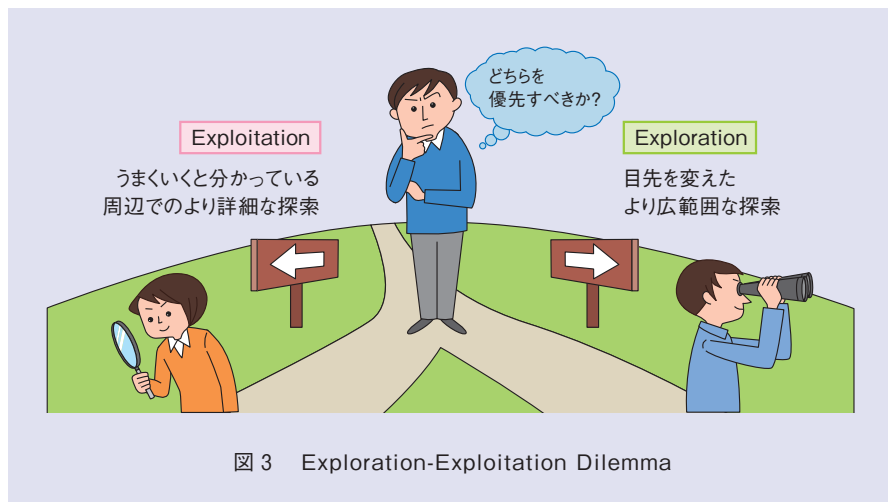


図3 Exploration-Exploitation Dilemma

計算リソースが限られている中、膨大な組合せの中から最適解をしらみ潰しに探し出すことは多くの場合困難です。そこで、うまくいくと分かっている周辺でのより詳細な探索であるExploitationを行うべきか、目先を変えたより広範囲な探索であるExplorationを優先すべきかの選択を迫られます。これはExploration-Exploitation Dilemmaと呼ばれます(図3)。ベイズ最適化とは、ExploitationとExplorationのバランスを考慮して次に探すべき解の候補(例えば、より良い誘導策)を絞り込む手法の1つです⁽⁹⁾。NTTでは、IoT(Internet of Things)の頭脳としてヒト・モノ・環境の情報を分析し、予測や制御に活用するAIであるAmbient-AIの基盤としてこれらの研究開発に取り組んでいます。

一方、単純なしらみ潰しが困難な場合でも、データ構造を工夫することで効率的にすべての組合せを数え上げ、最適解を厳密に求めることができる場合もあります。本特集記事『二分決定グラフを用いたネットワーク信頼性最適化法』では、二分決定グラフ(BDD)と呼ばれるデータ構造をうまく用いることで、これまで不可能だった大規模な問題が驚くほど効率良く解けるよう

になる事例を紹介します。

Ambient-AIがさらに進化するとともにこのようなネットワークへのAI技術の活用が進むと、複数のAIが協調し社会システム全体を最適化するNetwork-AIとなります。

今後の展開

本稿ではCS研によるコミュニケーション科学の新たな取り組みの一端を紹介しました。新たな取り組みにはリスクが伴います。すぐに期待どおりの学術的な成果が得られるとは限りません。研究開発においても、最先端の技術を活用し、直面する課題にうまく適用することを模索するExploitationを行うべきか、これまでの延長ではない、新たな次元を切り拓く取り組みに挑戦するExplorationを優先すべきか、まさに選択を迫られる中⁽¹⁰⁾、私たちはExplorationを優先してチャレンジな課題にこれからも挑戦していきます。ただし、研究を進めるうえでは、自分の仮説に都合が良い証拠ばかり探して解釈することのないよう、認知バイアスには十分注意が必要です⁽¹¹⁾。

参考文献

(1) 山田・高橋・納谷・池邊・古川: “NTTグループにおけるAI研究の取り組みと方向性,” NTT技術ジャーナル, Vol.28, No.2,

pp.8-13, 2016.
 (2) <http://www.ntt.co.jp/news2015/1512/151214a.html>
 (3) 秋山: “だまされる心理からみる高齢者トラブル,” 国民生活, Vol.13, pp.1-4, 2013.
 (4) J. Weizenbaum: “ELIZA - A Computer Program for the Study of Natural Language Communication between Man and Machine,” Communications of the ACM, Vol.9, No.1, pp.36-45, 1966.
 (5) <http://www.kecl.ntt.co.jp/IllusionForum/v/checkerShadow/ja/index.html>
 (6) <http://www.ntt.co.jp/news2018/1801/180131b.html>
 (7) 柏野: “アスリートの脳を解明し鍛える——NTTスポーツ脳科学プロジェクト,” NTT技術ジャーナル, Vol.30, No.1, pp.10-13, 2018.
 (8) 納谷・宮本・上田: “時空間多次元集合データ解析技術による集団最適誘導,” NTT技術ジャーナル, Vol.29, No.7, pp.24-28, 2017.
 (9) E.Brochu, V.M. Cora, and N. de Freitas: “A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning,” arXiv:1012.2599, 2010.
 (10) J.G.March: “Exploration and Exploitation in Organizational Learning,” Organization Science, Vol.2, No.1, pp.71-87, 1991.
 (11) R.Nuzzo: “How scientists fool themselves — and how they can stop,” Nature, Vol.526, No.7572, pp.182-185, 2015.



山田 武士

今後ますます技術の進歩のスピードが速くなり、競争が厳しくなる中で、CS研は得意分野を活かしつつ、これからも新たな次元を切り拓くチャレンジに大胆かつ粘り強く取り組んでいきます。

◆問い合わせ先

NTTコミュニケーション科学基礎研究所
 TEL 0774-93-5000
 FAX 0774-93-5015
 E-mail takeshi.yamada.bc@hco.ntt.co.jp