

AI時代の自動翻訳技術 ～ 劇的精度向上と近未来 ～

本稿では、自動翻訳技術の最近起こった様変わりをご紹介します。

英語は、文法・語彙・文字・発音に関して日本語と大きく異なるので、英語母語話者が日本語を習得するのは困難で長時間を要する。逆もまた真である。実際、日本語の母語話者の少数だけが英語を話せる。英語の母語話者で日本語が話せる人数はさらに少ない。

日本語が利用される地域は日本にほぼ限定されている。「世界語」である英語は日本ではほとんど通じないし他の外国語に通じるわけでもない。外国人と日本人とのコミュニケーションには越えがたい高い壁がある。

だからこそ、日本には自動翻訳技術に対する大きなニーズがある。しかしながら、日英間の翻訳は、前記の理由で、人間にとって困難であるだけでなく、コンピュータにとっても困難であり、長年の研究にもかかわらず、日英間の自動翻訳の精度は、英仏や日韓のような翻訳が簡単な言語対に比べて低いレベルにとどまり、誤訳を揶揄されたり非難されたりし続け、広く普及することがなかった。

2010年以降のAI（人工知能）の復活に同期して、自動翻訳の精度が目覚ましく向上した。次節以降で、ここ数年の急速な改善と2020年を区切りに短期・中期の見通しを述べる。

AI時代の自動翻訳技術

近年、AI（人工知能）の成果が巷間を賑わせている。ニューラルネットによって、人間の知的活動のかなりの部分がコンピュータで実現可能になったと考えられている。

写真に映っているモノが花か動物かを識別する課題での成功に始まって、囲碁や将棋など名人・強者がひしめく高度なゲームで、第一級の専門家をコンピュータが打ち負かしたことによって、ある種のブームになった。その影響で、AIの研究者・開発者の市場は極端な人手不足・売り手市場となって、その給与は鰻上りである。

ニューラルネットは、ニューロン（脳の神経細胞）にインスパイアされたものであり、入出力データを模倣できるようにコンピュータに学習させる仕組みである。ニューラルネットの研究は古く1940年代に起源があり、様々な研究を経て1980年代には、計算量をはじめとする限界にぶつかり、一旦急速に研究が下火になり、かえりみる研究者はほとんどいなくなった。その後のハードウェアとアルゴリズムの進化で2010年頃に劇的に復活した。以前と異なり、多層のニューラルネットを対象とすることから、深層

学習と呼ばれている。

自動翻訳も、ニューラルネットで実現可能である。過去の翻訳が大量に電子的に蓄積されており、これを模倣するように動作するもので、ニューラル翻訳（NMT）と呼ばれる。

NMT になって、従来技術と比べて格段に翻訳精度が向上した。表 1 に一例を示すが、従来技術の訳文はまさに機械が翻訳したようなぎこちないものである一方、NMT の訳文は母語話者が翻訳したような流暢なものである。このように素晴らしい訳文を見て、人間の翻訳者は「NMT は馬鹿にできない」と感じたし、一般の人は「NMT が十分役に立つ」として受け入れつつある。

より注意深く見ると、この NMT 訳には『英語への』に相当する英語表現がない。このように、訳が抜けたり、逆に、原文に無いものが訳文に湧き出したり、翻訳としての適格性に抵触する誤訳が時々生じる。

全体としては、その精度は商用化レベルに達しており、現在、熾烈な研究・開発競争が行われている。2020 年までの数年で前記の課題がかなり解決され普及も加速するだろう。

表 1 訳文の比較（従来手法と NMT）

日本語原文	近年の NMT の進展により、従来は自動翻訳が非常に困難だった日本語文章の英語への自動翻訳精度が顕著に向上してきた。
従来技術で英訳	The development of NMT in recent years, conventional automatic translation was very difficult to machine translation accuracy of Japanese sentences in English has improved
NMT で英訳	With the recent development of NMT, the automatic translation accuracy of Japanese sentences that had previously been very difficult to translate has improved significantly.

この NMT は現在様々な分野に応用されており、特許をはじめとする書き言葉の文章だけでなく、会話文を対象とした音声翻訳にも利用されている。次節で、音声翻訳の状況に目を移そう。

2020 年までの音声翻訳技術

音声翻訳性能は TOEIC を利用した尺度で表せる。同尺度が提唱された 2000 年の時点で、システムは既に約 600 点を達成している。2014 年以来国家プロジェクトとして実施されている音声翻訳の高精度化と普及を目指

すグローバルコミュニケーション計画（GCP）ⁱⁱⁱの最終年度の2020年には、（2000年以降の技術革新による性能向上を踏まえると）非常に高いTOEICスコアが期待できる。高い性能のシステム^{iv}を使って、外国語が苦手な人の能力が拡張されることの意義と波及効果は大きい。

GCPでは、観光・災害・医療等をはじめとする生活の場面で、日本語と英語を含む10言語^vで同様の高性能を実現することを目標にしている。音声翻訳のハードウェアとしては、既に、メガフォン型端末^{vi}、小型端末^{vii}、ハンズフリー端末^{viii}、スタンドアローンのデバイス^{ix}、等が商用化されている。

2020年には、津々浦々に普及した音声翻訳システムによって、日本は世界で最もコミュニケーションが容易な国に脱皮しているだろう。

音声翻訳の研究は1986年の(株)エイ・ティ・アール自動翻訳電話研究所の設立^xとともに日本で始まり、世界に広がり、再び日本に戻って、今、高いピークを迎えつつある。

2020年から先の音声翻訳をはじめとする自動翻訳技術

課題が多々あり2020年には研究開発が未完成と想定できるものに、高精度汎用翻訳と同時通訳と文脈を理解した翻訳の三つがある。

まだ、高精度な汎用翻訳システムを実現していない

現在の技術はコーパスに強く依存している。コーパスの量が多ければ多いほど性能が高くなる。逆に、コーパスが不足している分野では性能が低い。研究開発を進める各機関はコーパスの収集を競っているが、現在の収集方法では不十分なことは、翻訳サービスの精度が分野や言語や提供元で一定でないことが証明している。つまり、全ての分野・言語で高精度を実現した汎用翻訳システムは存在しないのである。

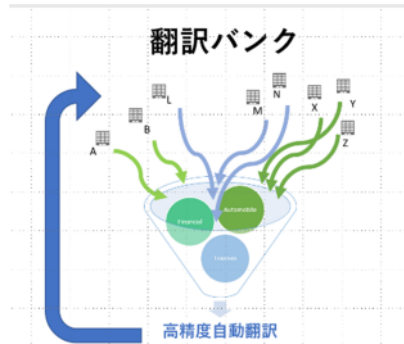


図1 翻訳バンクのコンセプト

対訳コーパスは会社や自治体・省庁等多数の組織に分散している。図1に示したように、全部の対訳コーパスを一つの公的機関に集積して、自動翻訳のための深層学習に投入すれば、翻訳精度が顕著に上がると見込める。

例えば、日本の場合、1年間で2000億円の予算が翻訳に使われているので、年間の翻訳生産量は5億文と推定できる。10年間では50億文にのぼる。翻訳ニーズのある全ての分野・言語で、この大規模な対訳コーパスを用いて構築したシステムは、未体験の高精度を実現できるだろう。

国民の国民による国民のための自動翻訳システムを構築するというビジョンである。日本は、この仕組みを翻訳バンクと命名して注力している。

まだ、コンピュータは『同時通訳』ができない

今の技術は一文全部を聞き終わってから処理する。入力が長いと処理の完了が大きく遅延する。これを避けるためには、同時通訳者のように、文末が来る前に処理し出力する必要がある。

日本語から英語の同時通訳は、語順の隔たりに起因して困難である。日本語の語順はSOVで最も重要なVが文末に出てくる。英語の語順はSVOでVが主語の次に出てくる。英語のVを出力するためには日本語の文末まで待つ必要がある。文末まで聞かないと訳出できないはずだが、同時通訳者は文末を待たずに翻訳する。同時通訳者は内容を予想して訳を決めている。誤訳だと分かった時点でリカバリーをしなければならない。

このように同時通訳は人間にとっても大変難しく、コンピュータによる同時通訳の実現には、相当の時間と大きなブレイクスルーとが必要である。一方で、萌芽的な研究はいくつかあり、また、テレビ会議の多言語化として商用化し徐々に機能を高度化していくという開発戦略が可能な技術水準は達成できている。

コンピュータは文脈を理解できていない

現在の自動翻訳技術では、一文単位で処理するので、前文の情報は全く使っていない。そのため、訳文を並べ文章として眺めると一貫性を欠いていることがある。同じ単語の訳語が 2 種類以上あって、文ごとに異なることがよくある。また、「です調」と「である調」の訳文が交互に現れ、意味は通じるのだが違和感を覚えることもままある。

文脈の研究も長期にわたって行われているが適用範囲や処理速度に問題がある。翻訳は周辺の単語を参照して決定可能であることが多い。一文単位の翻訳が実用的であれば、文脈が理解できなくても構わない。

一方、最近、深層学習が文脈理解にも成果を上げ始めている。文脈理解の典型例に、日本語で頻繁に起こる主語の省略を検出し、主語の実体を決定する課題がある。これに関して従来法を大きく上回る性能を実現したという報告ⁱⁱⁱがある。著者は、文脈を理解した翻訳が急速に発展する予兆と期待する。

<プロフィール>

隅田 英一郎 (すみた えいいちろう)

国立研究開発法人 情報通信研究機構 フェロー

1982 年電気通信大学大学院電子計算機学専攻修了。

1999 年京都大学大学院博士 (工学)。

日本アイ・ビー・エム、国際電気通信基礎技術研究所を経て、情報通信研究機構で、2020 年をゴールとする音声翻訳の国家プロジェクト「グローバルコミュニケーション計画」を推進。

—— 受賞歴 ——

情報処理学会喜安記念業績賞

文部科学省科学技術分野の文部科学大臣表彰科学技術賞

内閣府産学官連携功労者表彰総務大臣賞他

2014 年～2015 年 言語処理学会 会長

2015 年～ アジア太平洋機械翻訳協会 理事

2015 年～ 日本翻訳連盟 理事

-
- i <http://www.effectivelearning.com/language-guide/language-difficulty>
 - ii 現時点の性能は、無償公開中のサイト、みんなの自動翻訳@TexTra (<https://mt-auto-minhon-mlt.ucri.jgn-x.jp/>) でご体験いただける。
 - iii http://www.soumu.go.jp/main_content/000285578.pdf
 - iv 現時点の性能は、無償公開中のアプリケーション、VoiceTra (<http://voicetra.nict.go.jp/>) をスマホやタブレットにダウンロードしてご体験いただける。
 - v 日、英、中、韓、尼、泰、越、緬、仏、西語
 - vi <http://news.panasonic.com/global/topics/2016/45751.html>
 - vii http://www.nec.com/en/press/201710/global_20171002_02.html
 - viii <http://www.fujitsu.com/global/about/resources/news/press-releases/2017/0919-01.html>
 - ix <http://logbar.jp/en/index.html>
 - x <http://www.atr.jp/about/history.html>
 - xi Xiaolin Wang, Andrew Finch, Masao Utiyama, Eiichiro Sumita. A Prototype Automatic Simultaneous Interpretation System. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations, pages 30-34, Osaka, Japan, December 11-17 2016.
 - xii Intra-sentential Zero Anaphora Resolution using Subject Sharing Recognition, Ryu Iida, Kentaro Torisawa, Chikara Hashimoto, Jong-Hoon Oh, Julien Kloetzer, In the Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2015), pp.2179-2189, Lisbon, Portugal, September, 2015.