

特集「深層学習による言語生成」にあたって

鍛治 伸裕

(ヤフー株式会社)

近年、深層学習は人工知能に関連するさまざまな研究分野において大きな革新をもたらしてきました。もちろん自然言語処理もその例外ではなく「過去10年の間に起きた、自然言語処理における最大のブレイクスルーは何か?」と聞かれれば、大半の自然言語処理研究者は深層学習と答えるのではないのでしょうか。この原稿を執筆している間にも「TOEIC 960点レベルのビジネスマンと同等の翻訳精度を実現した」とされる深層学習ベースの翻訳システムのプレスリリースが発表されました。筆者はTOEIC 955点の会社員(=ビジネスマン?)なので、もしプレスリリース*1の内容を字面どおりに受け取るならば、この翻訳システムは筆者と同程度の翻訳能力を有していることになります。翻訳のように高度に知的な作業において、自分と計算機の能力が同程度であるというのは、なかなかショッキングな出来事です*2。

深層学習が自然言語処理にもたらしたブレイクスルーとは、端的に言うてしまうと「深層学習を使うことによって機械翻訳などのタスクの精度が上がった」ということです。しかし、単に精度が上がったというだけの話であれば、これまでもさまざまな取組みが存在していたはずで、では、深層学習が過去のものと同程度であるとされる理由はどこにあるのでしょうか。もちろん、その理由の一つだけではないのですが、あえて一つだけ選ぶとするならば、言語生成*3という問題に対する統計的アプローチを確立させたことであると筆者は考えています。

これまでも、言語生成タスクに統計的アプローチが全く適用されてこなかったわけではありませんが、その方法論が十分に確立されていたとは言いがたい状況でした。例えば機械翻訳においては、IBMモデルを始めとする統計的手法が大きな成功を収めてきました。しかし、それらは基本的に語句単位での翻訳を行おうとするものであり、いわゆる還元主義的アプローチと同種の問題を

抱えていました*4。これに加えて、実際には膨大な数のヒューリスティクスによって高い精度が得られていたという側面があったことも否めません。また、文書要約においてもさまざまな統計的アプローチが研究されてきましたが、抽出型要約手法という名のキーワード抽出が主流であり、残念ながら言語の「生成」を実現しているわけではありませんでした。

深層学習はこうした研究状況を一変させました。深層学習の登場によって任意の言語情報を分散表現に変換して取り扱うことが可能となり、seq2seqモデルに代表される条件付き言語モデルの発明へとつながりました。これによって、単一の統計モデルを用いてさまざまな言語生成の課題を統一的かつ簡便に扱うことが可能になり、言語生成の精度は飛躍的な向上を遂げました。しかも、データをいったん分散表現に変換してしまえば、そこには言語や画像といったメディアの区別はほとんど存在せず、画像に対する説明文の生成やスポーツ動画を鑑賞しながらの対話生成など、マルチモーダルな言語生成も容易に実現可能となります。そのため、言語生成が扱うことのできる研究課題の幅も格段に広がりました。さらには言語処理や画像処理といった既存の分野の壁をも破壊しつつあります。

このように、深層学習の進展によって言語生成技術は劇的に進化しつつあります。そこで本特集では、その言語生成に関する最近の研究動向について、第一線でご活躍されている研究者の方々からご紹介いただくことにしました。機械翻訳や要約といった古典的な言語生成タスクはもちろん、画像説明文生成といった深層学習以後に大きな発展を見せたタスクまで、さまざまな言語生成タスクに関する話題を取りそろえました。本特集が、これから言語生成研究に取り組もうとする若者達にとって良い入門書となるとともに、新たな研究アイデアを生み出す一助になれば幸いに思います。

*1 <https://miraitranslate.com/news/362>

*2 もちろん、このプレスリリースで報告されている実験結果の妥当性や解釈についてはもう少し慎重に議論すべきだと思いますが、本稿は技術論文ではありませんので、そこはあえて触れないことにします。

*3 翻訳や要約や対話など、計算機が自然言語テキストを生成するタスクのことを指します。

*4 「文章を翻訳するという処理は、そこに含まれる語句を個別に翻訳するという処理に還元可能である」という一般的には無理のある仮定が置かれてしまっているという意味です。