

高校生のページ**コンピュータで言葉を理解する**

情報学研究科 知能情報学専攻 知能メディア講座
黒橋 禎夫

1. はじめに

言葉は、われわれ人間にとってなくてはならないものですが、逆にあまりに身近なために、普段その役割やありがたみに気づくことはあまりないかも知れません。しかし、私はいま言葉を使って何を書くかを考え、それを言葉を使って文章にしています。みなさんも、この文章を読み、みなさんの頭の中の言葉で表現された知識に対応付けながら、その意味を解釈していることと思います。このように、言葉は思考、コミュニケーション、記憶のもっとも重要な手段であり、言葉を持ったことによって人間が生物の進化の中でまったく別次元の存在になったとも言えるのです。

私は、言葉をコンピュータで扱う、最終的には理解させることを目標として、大学で研究をしています。このような研究分野を自然言語処理とよびます。コンピュータの世界にはCやJavaのようなプログラミング言語もありますので、それと区別するために日本語や英語のことを自然言語とよびます。自然言語処理の研究は、電子計算機（いわゆるコンピュータ）が誕生した1945年ごろ、英語とフランス語などの間の翻訳をコンピュータで自動的にできないか、という発想とともに始まったといわれています。しかし、当時はコンピュータの能力が今とは比べ物にならないほど非力でした。コンピュータに少量のデータを入力するだけでも大変でしたし、英語アルファベットを扱うことはできても何千種類もある漢字などはとても扱えないという状況でした。

その後、コンピュータの能力は年々進化し、漢字を含む日本語テキストも扱えるようになりましたが、一方で、コンピュータに言葉を理解させるということが簡単でないこともわかってきました。言葉を理解するために必要な「知識」をコンピュータに与えるということが極めて難しかったのです。これは私が大学院生であった1990年ごろでも同じでした。ところが、この20年間で自然言語処理をとりまく環境は劇的に変化し、自然言語処理の研究そのものも大きく進展しました。この記事では、高校生のみなさんに、そのわくわく感をすこしでもお伝えできればと思います。

2. 言葉のはたらきと解釈

言葉は思考、コミュニケーション、記憶の手段であるといいましたが、具体的に、言葉はどのようなはたらきを持つのでしょうか？ それは、ものごとに名前をつけて、その関係を表現するということです。

(1) 私は今朝ご飯を食べた。

という簡単な文について考えてみましょう。「私」「ご飯」はものの名前、「今朝」も時間についてのある概念の名前、「食べる」はある動作の名前です。そして、これらが「いつ、誰が、何を、どうした」という関係にあることが表現されています。ちなみに（嫌いな人が多いかも知れませんが）文法用語では「は」や「を」は助詞です。関係を表現することを助けているからそう呼ぶのでしょうか。

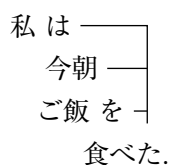
言葉を使えば、このような、ものごとの間の複雑な関係をいくらかでも表現することができます。たとえば次のような感じです。

(2) 寝坊した私は昨日買っておいたパンを食べずに、サブリにした。

このように名前付けと複雑な関係を表現するという意味で言葉を使えるのは地球上の生物の中で人間だけなのです。動物でも単純ないくつかのものごとに記号（動作や鳴き声）を与えたり，チンパンジーが記号のごく簡単な組み合わせを操作・理解できることが知られています。しかし，その複雑さは人間には遠く及びません。

図 1 形態素解析・構文解析の例

語	読み	原形	品詞	活用型：活用形
私	わたし	私	普通名詞	
は	は	は	副助詞	
今朝	けさ	今朝	時相名詞	
ご飯	ごはん	ご飯	普通名詞	
を	を	を	格助詞	
食べた	たべた	食べる	動詞	母音動詞：タ形
.	.	.	句点	



さて，このような言葉をコンピュータに理解させることはすんなりできることでしょうか。まずはコンピュータの腕前をみてもらうために，さきほどの例文 (1) をコンピュータのプログラムで分析してみましょう。その結果が図 1 です。上段は，専門用語でいうと形態素解析，すなわち，文がどのように単語に分割され，それぞれどのような品詞や活用をもっているかを示したものです。どうです，ちゃんと「名詞」「動詞」「助詞」なども理解できています。ここでの文法用語はいわゆる学校文法とは少し違うものですが，おおよそ理解してもらえるものと思います。

次に図 1 下段に示したものが構文解析といわれる処理の結果です。小学校・中学校の国語で「どの文節がどの文節を修飾しているか」という問題をやったと思います（これも多くの人が嫌いだったでしょうね），まさにそれをコンピュータで分析した結果です。これも正しく解析されているようです。

ということで，さすがにこれぐらい単純な文であればコンピュータでも正しく解析することができます。では，例文 (2) はどうでしょうか。例文 (1) に比べれば少し複雑ですが，みなさんにとってはどうということのない文でしょう。ところが，コンピュータの分析はこのあたりからあやしくなってきます。その原因は言葉の「あいまい性」にあります。先入観をとりのぞいてよく読むと，「昨日」というのは「パンを買っておいた」のが昨日なのか，「パンを食べずにサブリにした」のが昨日なのか，2つの解釈が成り立つことがわかります。

さらに，言葉には単語や句の意味のあいまい性もあります。これは，「一つの表現が複数の意味を持つ」ことがあるからです。例文 (2) でも，「パン」はフライパンなどの鍋の「パン」かも知れませんし，カメラ操作のパン・チルドの「パン」かも知れません。さらには，文末の「した」は名詞の「下」または「舌」がひらがなで書かれたものかも知れません。

ここまでくると、「そんなバカな」「屁理屈だ」と思うかも知れませんが、そう思うのは、人間が膨大な常識によって、その文が述べられる状況に応じて整合する（つじつまがあう）解釈を無意識のうちに優先しているからなのです。たとえば、「パン」を「食べる」なら「パン」は食べ物のパンであって、鍋やカメラ操作の可能性には気づくことすらないでしょう。しかし、そのようなことをコンピュータで行うことは難しいことなのです。

もう一つ、「一つの表現が複数の意味を表す」ことのちょうど逆の現象として、「複数の表現が同じ意味を表す」ことも多々あります。「私」は「僕」や「俺」でもいいですし、「食べた」は「食った（くった）」といってもいいのです。さらには、「サプリにした」は「サプリを食べた」と同じことを意味しています。本当に言葉を理解するためにはそのような「同じ意味」であることもきっちりとわかる必要があります。

言葉はある意味で人間にとって非常に効率的にできています。いちいち「昨日私が買った食べ物のパンを今朝は私は食べず、今朝は私はサプリを食べた」という必要はなく、人間の方が効率的に表現された文を柔軟に解釈します。しかし、これを常識をもたないコンピュータに行わせることは簡単ではなく、これが私が研究の世界に入った20年前の状況でした。

3. ウェブという超大規模テキストの出現

われわれはほとんど意識することなく言葉を使いこなしていますが、それは常識に基づく整合的な解釈が行われているからだ、ということのみました。逆にいえばコンピュータによって言葉を理解するためには、コンピュータに常識を与えることが必要だということになります。

しかし、コンピュータに常識をひとつひとつ人手で与えるということはほとんど不可能です。そういう無謀なことをやってみようという人も過去にはいたのですが、やはり無理なことがわかりました。ところが、この20年の間にこの問題の解決に大きな糸口が現れました。ウェブ（World Wide Web）という超巨大な電子テキストの出現です。

いま高校生のみなさんは生まれた時からウェブがあり、ウェブがある世界が当たり前かも知れませんが、少しだけ歴史をみておきましょう。そもそもアメリカの国防総省の研究でコンピュータを相互にネットワーク接続したことがインターネットのはじまりでした。1980年ごろのことです。当初はその上でファイルの転送や電子メールのやりとりが行われていましたが、1990年に、欧州の原子核研究所のティム・バーナーズ・リーという人がインターネット上の電子テキストをリンクで接続することを考え、これがウェブのはじまりとなりました。

その後のウェブの急速な発展は目をみはるものがあり、いまやあたりまえの存在になりました。まさに電気・水道・ガス・ウェブです。みなさんも日常的に検索エンジンで検索を行い、リンクをたどりながら情報を探索していることでしょう。また、（これももはや新しい言葉ではありませんが）ウェブ2.0といわれるように、マスコミだけでなく、一般の人々がウェブを通して情報を発信し、それをみなで共有するという時代になりました。

だいぶ脱線しましたが、コンピュータで言葉を理解するための常識の話にもどりましょう。ウェブは一義的には人間がその情報を利用するものですが、実は、その大規模テキストをコンピュータ自身も利用することができるのです。すなわち、コンピュータが利用するための常識を、コンピュータ自身がウェブのテキストから獲得するのです。

これは、若干不思議に聞こえるかも知れませんが、コンピュータが言葉を理解できないから常識がいるわけで、それが無いのに、どうしてコンピュータがテキストから知識を取り出すことができるのでしょうか。ニワトリとタマゴ？

しかし、これは魔法でも何でもありません。膨大なテキストがあれば、その中にはさきほどの例文（1）のように比較的単純なものもあります。また、個々の文がちゃんと解釈できなくても、語の使われ方の

全体の傾向をみることでわかることもあるのです。重要な点は、そのテキスト量が本当に膨大であるという点にあります。ウェブには1兆ページを超えるともいわれるページがあり、そのほんの一部を取り出して整理するだけでも簡単に数十億文というテキストを収集することができます。ちなみに、大手新聞社の新聞1年分のテキスト量がおおよそ100万文です。ということは、たとえば10億文のテキストを集めれば新聞1000年分のテキスト量に相当するのです！ 次の節では、このような超大規模テキストから常識を獲得する方法を紹介します。

4. 超大規模テキストからの常識の獲得

日本語文の修飾関係のあいまい性に関する有名な例文として、次の2文があります。

- (3) a. クロールで泳ぐ少女を見た。
b. 望遠鏡で泳ぐ少女を見た。

この2文は同じ文体ですが、「クロールで/望遠鏡で」の修飾先が異なります。このような文を正しく解釈するには、「クロール」と「泳ぐ」、「望遠鏡」と「見る」の間の関係をコンピュータが知識として知っておく必要があるわけです。この問題を解決するために、われわれの研究グループでは、ウェブから10億文規模の日本語文を収集し、形態素解析・構文解析・クラスタリング（類似したものをまとめる処理）を行うことによって「誰が何をどうした」という格フレームとよばれるパターンを大規模に学習しました。

まず、大規模テキスト中の各文を形態素・構文解析します。その際、たとえば「綺麗だ」「奇麗だ」「きれいだ」などは同じ語として扱います。これは、コンピュータ用の辞書にそのような同義語をきちっと与えておくことで可能となります。次に、解析結果から確からしい解析結果を取り出します。確からしい解析結果とは、構文解析を行うシステムの文法規則において、修飾先の候補がただ1つしかないものです。たとえば、次にあげる例では、○下線部が修飾先のあいまい性がない要素として抽出されます。

- (4) a. 今日は 石窯で○ パンを○ 焼いています。
b. 80種類もの パンを○ 焼いていますが、…
c. その後、パンを× 焼いた余熱を利用し…
d. 直径が× 15センチのケーキを焼きます。

(4a)は文末であり、(4b)では「～いますが」が文中の強い区切りと認識されるので、「焼く」をあいまい性なく修飾する「石窯でパンを焼く」と「パンを焼く」がそれぞれ抽出されます。

一方、(4c)、(4d)の×下線部は、修飾先のあいまい性があるので抽出されません。(4c)では、下線部「パンを」は「焼いた」と「利用し」の2つの修飾先の候補があるため抽出しません。(4d)では、下線部「直径が」が「15センチ」と「焼きます」の2つの修飾先があるため同様に抽出しません。このような判断は、「～が～センチの」のように動詞ではないのですが修飾先として振る舞う表現を文法規則としてきちんと把握することで可能になります。

格フレームを学習する際の大きな問題は動詞の意味のあいまい性です。「焼く」の場合、「パンを焼く」「手を焼く」「データを焼く」のような意味・用法の異なる表現をクラスタリングによって区別する必要があります。このとき、動詞の直前の名詞が動詞の用法の決定に強く影響していると考え、動詞とその直前の名詞の組が同一のものをまずまとめます。その上で、他の名詞が同じであったり似ているものをまとめていきます。たとえば「焼く」の場合、まず「パンを焼く」「肉を焼く」「手を焼く」などの組

で例文を集約し、その上で、語の意味の近さに基づいて「パンを焼く」「肉を焼く」のような似ているものをまとめます。ここで、「パン」と「肉」が似ているということも大規模テキストから自動的に学習することができるのですが、その説明はここでは省略します。

このような方法で、ウェブから収集した約 10 億日本語文を用いて、約 4 万種類の動詞・形容詞について格フレームを自動的に構築しました。その一部を表 1 に示します。このような知識をコンピュータに与えておくことで、それを参照することにより「クロールで / 望遠鏡で泳ぐ少女を見た」という文の修飾関係を正しく解釈することが可能となったのです。

このように「誰が何をどうした」という格フレームはコンピュータに与えるべき常識の第一歩であり、文の構造の解釈に寄与するという意味でもっとも重要な知識の一つです。今後、コンピュータに与えることが必要な知識としては、「酒を飲む→酔う」「レストランに行く→料理を注文する」というような因果関係や出来事の連鎖の知識、また、「椅子：目的：座る」「自動車：部品：エンジン」「富士山：高さ：3776 メートル」などの語の特徴や属性に関する知識などがあります。このような知識も大規模テキストの中にさまざまな形で記述されていますので、うまく工夫することによって今後徐々に自動獲得され、コンピュータに与えられていくことでしょう。

なお、ここで紹介したような大規模な計算を行うためには数百～数千台規模の並列コンピュータが必要となります。電子工学、情報工学の進歩がコンピュータの長足の進歩を支え、それがさらにこのような自然言語処理の研究を支えていることも忘れないで下さい。

表 1 自動獲得した格フレームの例（名詞の後ろの数字は学習テキスト中の頻度）

	格	用例（数字は頻度を表す）
焼く (1)	ガ	私：18, 人：15, 職人：10…
	ヲ	パン：2484, 肉：1521, ケーキ：1283, …
	デ	オープン：1630, フライパン：1311, …
焼く (2)	ガ	先生：3, 政府：3, 人：3, …
	ヲ	手：2950
	ニ	攻撃：18, 行動：15, 息子：15, …
焼く (3)	ガ	メーカー：1, ディストリビューター：1, …
	ヲ	データ：178, ファイル：107, コピー：9, …
	ニ	R：1583, CD：664, CDR：3, …
∴	∴	∴
泳ぐ (1)	ガ	イルカ：142, 生：50, 魚：28, …
	ヲ	海：1188, 水中：281, 海中：101, …
	デ	クロール：86, 平泳ぎ：49, 泳法：24, …
∴	∴	∴
磨く (1)	ガ	私：4, 男性：4, 人：4, おれ：4, …
	ヲ	歯：5959, 奥歯：27, 前歯：12
	デ	ブラシ：38, トイレ：15, 塩：13, …
∴	∴	∴

5. おわりに

この記事では、言葉がどのようなはたらきや特徴をもっているか、人はどのようにそれを使いこなしているのか、コンピュータで言葉を理解するためには何が必要かということの説明してきました。また、ウェブという超大規模テキストがコンピュータによる知識獲得の対象であり、そこからさまざまな知識が自動獲得されつつあることをみました。

SFの世界や近未来を描く映画の中では言葉を理解するロボットやソフトウェアがあたりまえのように出てきます。それはそんなに簡単なことではないのですが、もしかしたら本当に近未来にそれに近いものが実現するかも知れません。

世界はどんどんグローバル化されていますが、その一方で紛争も絶えません。その根底に人々の中の相互理解の不足、文化差、そして言葉の壁があります。はじめに書いたように異なる言語間の自動翻訳（まさにドラえもんの「翻訳こんにゃく」です）は自然言語処理の重要な研究テーマの一つであり、その研究も近年大きく進展しています。すでに特定の旅行シーンでの対話に限った音声での翻訳システムもではじめていますが、今後のこのようなシステムの発展が地球規模の共生社会の実現に貢献することは間違いないでしょう。2022年サッカーW杯の日本招致のコンセプトビデオ（残念ながら招致は成功しませんでした）では、未来のすばらしい音声翻訳システムが出てきますので是非見てみて下さい（<http://youtu.be/8hZyCiR0QJU>）。

この記事を読んでくれた高校生諸君の中から、このような研究に興味を持ち、将来、自然言語処理の研究開発にたずさわる人が出てきたとすれば、望外の喜びです。