

## If 節を伴わない仮定法の翻訳手法

加藤 鉦三  
信州大学

黒田 航  
情報通信研究機構けいはんな研究所

### 【 1 】 目的

次は、市販されている翻訳ソフトの出力である。(訳文の前の×は誤訳であることを示す)

(1) The company could go bankrupt without government help.

[LV] × 会社は政府の手助けなしで倒産することができました。

[The] × その会社は政府支援なしで破産することができました。

[Atlas] × 会社は政府支援なしで破産できました。

[PC] × 同社は、政府援助なしで破産することができた。

(1)の正解訳は「その会社は政府の支援がなければ倒産することがあり得る」であるが、4つ全ての出力で、正解訳とは全く逆の印象を与えるような訳を与えている。翻訳ソフトはcouldのデフォルトの訳語を「できた」と設定しているように思われるが、それがこの誤訳の原因である。本発表では、助動詞couldとwouldを含む英文を機械翻訳ソフトが誤訳しないようにするには何が必要であるかを考察し、対処法を提案する。

### 【 2 】 方針と動機

(2)の文は、それを単独で見ると、2つの解釈を認めるしかないように見える。

(2) I could swim faster.

a. 私はもっと速く泳ぐことができた。

b. 私ならもっと速く泳ぐことができるだろう。

しかし、次のように文脈を与えると、(2)のような複数の解釈の余地はなくなり、それぞれ単独の解釈しか許さない。なお、ここでは、決定要因は前文の時制である点に注意されたい。

(3) 「できた」のみ可： When I took a lifeguard certification course with a few of my track buddies, though, I did much better at all the swimming drills than they did. I could swim faster and for longer amounts of time.

(4) 「できるだろう」のみ可： Her voice had an edge of hysteria. "Row!" "I am!" "For God's sake, I could swim faster than this!"

人間が解釈する場合は、いわゆる「文脈」により判断している。本発表は、その文脈情報を明示化することにより、「文脈」による判断を機械にさせることを試みる。その際の考察の方針は次である。

(5) 訳語とその訳語が当てはまる環境を明示化し、かつ、それらの環境が相補分布を成すことを前提に環境を設定する

これは、人間が文を解釈する際、意識せずに行っている判断プロセスをモデル化するという方向性を持つ語彙意味論であり、おそらく伝統的な言語学的手法にはなかった取り組みであるとも言えるかもしれない。そして(5)の作業は、言語学者が自然言語処理の発展に貢献できる場所であると考えられる。

### 【 3 】 Would

Wouldの用法は、概略次のものである。

(6) 時制の一致による単純未来のwould

訳語： 「だろう」

環境： 時制の一致によって単純未来のwillがwouldになっている

She believed that her husband would soon get well.

(7) 時制の一致による意志未来の would

訳語： 「しよう」

環境： 時制の一致によって意志未来の will が would になっている

I said I would do it.

時制の一致は、動詞 1 によって導かれる that 節や if/whether 節内の動詞 2 について、動詞 1 が過去時制であるという理由で動詞 2 が過去形になる現象を言う。日本語では時制の一致は起こらないため、この二つの環境では、would を will とみなして訳す。(6)と(7)では、翻訳ソフトは誤訳しない。

(8) 意志未来の過去形

訳語： 「しようとした」

環境： 時制の一致によらず、かつ過去時制環境にあり、かつ否定文である

I asked her several questions, but she would not answer any of them.

(8)は全ての翻訳ソフトが誤訳する。

[LV] ×私は彼女にいくつかの質問をしました、しかし彼女は(彼・それ)らのいずれにも答えないでしょう。

[The] ×私は彼女にいくつかの質問をしました。しかし、彼女は、それらのうちのどれにも答えないでしょう。

[Atlas] ×私は彼女のいくつかの質問をしました、彼女はそれらのいずれにも答えないでしょう。

[PC] ×私は彼女にいくつかの質問をした、しかし、彼女は彼らの誰にも答えない。

誤訳の原因は、翻訳ソフトが(8)を仮定法の would と誤認していることによる。(8)は仮定法ではなく、意志未来の will の過去形である。(8)の例文の would は、but 以下を見る限り、後述の仮定法の would であるかそれとも過去の意志であるのかは区別がつかない。しかし but の前の節の定動詞は過去形である。もし but の前が I will ask her several questions であったとしたら、この would は仮定法である。すでに(3)と(4)でも見たように、前文の時制が現在であるか過去であるかは、could と would の解釈には決定的に有効かつ必要な情報である。以下、本稿では、前節・前文もしくは付随する副詞節の時制が過去/現在である時、当該の文はそれぞれ過去時制環境/現在時制環境にあると言う。なお、(8)の用法は否定文が多いが、それは「期待値に反する」が意味論的な条件であるからであり、それに反しない限り、We tried to keep grandpa from drinking so much, but he would have his way. (ジーニアス英和辞典)のように肯定文も可能である。よって(8)の環境設定は厳密には正確ではない。しかし「期待値」を機械に判定させることはできないため、とりあえず否定文であることを条件としておく。

(9) 単純未来の過去形

訳語： 「したものだった」

環境： 時制の一致によらず、かつ過去時制環境にあり、かつ肯定文である

He would (often) go fishing in the river when he was a child.

(9)は、[Atlas]と[PC]は「だろう」と誤訳する。[LV]と[The]も、would の後の often を削除すると「だろう」と誤訳する。ここでも(8)と同様に、仮定法の would であるかそれとも単純未来の過去形であるか、という問題である。[LV]と[The]は would often で訳語を登録しているようであるが、それよりも would の時制により判定する方が効率的であろう。なお、単純未来の will の過去形がなぜ習慣・反復を表すのかは言語学的にも非常に興味深い問題であるがそれに立ち入る余裕はないため、「人に関する法則」であるとだけ指摘しておく。Oil will floats on water.が法則を表すことを参考にされたい。

(10) 仮定法

訳語： 「だろう」

環境： 時制の一致によらず、かつ現在時制環境にある

If a car hit it, the occupants would be killed on the spot.

If I were rich enough, I would buy it.

4つの翻訳ソフトはいずれも仮定法の would はほぼ正しく訳す。仮定法は Anyone would have thought that. のように if 節が表面に現れないものもあるが、それらも問題なく訳す。これは、翻訳ソフトが would の訳語として「だろう」をデフォルトとして設定していることによる。よって、would では、(8)(9)や次の(11)のように、「だろう」で訳せない時に問題となる。

(11) 条件節内の仮定法

訳語： 主語が you では「していただけるなら」

主語が you 以外では「しようと思えば」

環境： if 節内にあり、主節の定動詞が仮定法である

I should be most obliged if you would grant my request.

I could do so if I would.

やはりここでも翻訳ソフトは「だろう」と誤訳する。なお、この用法は丁寧な依頼である Would you ...? と同じものである。その条件は、疑問文であり、主語が you であること、となる。意外なことに、[LV] と [Atlas] は Would you ...? においても「だろう」と誤訳する。

以上見たように、would については、「だろう」という訳語が使えない環境を正確に記述する必要がある。

【4】 Could

Could の用法は、概略次のものである。Could の場合は、「できた」という訳語が使えない環境を正確に記述する必要がある。

(12) 能力の過去形

訳語： 「できた」

環境： 時制の一致によらず、かつ過去時制環境にある

When I lived by the subway station I could reach the office in ten minutes.

翻訳ソフトはいずれも could のデフォルトの訳語として「できた」を設定しているようであり、誤訳しない。

(13) 時制の一致による能力の could

訳語： 「できる」

環境： 時制の一致による could である

He thought he could swim across the river.

[LV] は一貫して「できた」と誤訳するようであるが、他のソフトは概ね正しく「できる」と訳す。

[LV] ×彼は彼が川を横切って泳ぐことができたと思いました。

(14) 能力の仮定法

訳語： 「できるだろう」

環境： 時制の一致によらず、かつ現在時制環境にあり、かつ might に置き換えたものとは生起数に大きな差がある

We could win the game if we tried harder.

I could swim across the river.

いずれの翻訳ソフトも、たいていの場合、if 節があれば仮定法として訳出するが、if 節のないものは「できた」と訳す。「できた」は正しい訳である可能性もあることは冒頭の(2)ですでに確認したところである。ここで問題にしたいのは、「できるだろう」という訳を出さない点である。

[LV] ×もし我々がもっと一生懸命試みたなら、我々はゲームで勝つことができました。

× / 私は川を横切って泳ぐことができました。

[The] 私たちは、もしより一生懸命努力すれば、ゲームに勝つことができるでしょう。

× / 私は川を泳いで渡ることができました。

[Atlas] 私たちが一層努力するなら、ゲームに勝つことができるでしょうに。

× / 私は川を横切って泳ぐことができました。

[PC] 我々がより一生懸命に試みるならば、我々はゲームに勝つことができる。

× / 私は、川を泳いで渡ることができた。

#### (15) 可能性の仮定法

訳語： 「あり得るだろう」

環境： 時制の一致によらず、かつ現在時制環境にあり、かつ might に置き換えたものと生起数に大きな差がない

Health care could become America's No.1 industry.

We could die.

いずれの翻訳ソフトも if 節を追加すると正しい訳をするが、if 節がないままでは「できた」と誤訳する。次に再掲する冒頭の(1)もこの例である。なお、(1)では without 句が条件提示句として機能している。

(1) The company could go bankrupt without government help.

(14)(15)の仮定法の用法の見極めで重要なのは、それが現在時制環境にあるかどうかである。しかし、仮定法であることはそれで判別できるが、能力なのか可能性であるかの判別は別の手がかりが必要である。それについては、本発表では might との置き換えを提案する。これは、may の二つの意味である許可と可能性のうち、might では後者の分布の方が強いことを利用したものである。

(16) might との置き換え（数字は Google 検索の件数）

	[could]	[might]	[解釈]
(14) "we ___ win"	268,000	38,300	能力
(14) "I ___ swim"	104,000	1,380	能力
(15) "it ___ become"	346,000	344,000	可能性
(15) "we ___ die"	23,200	24,400	可能性

(16)に見るように、能力の場合には could と might では桁が違うが、可能性の場合には大差がない。仮定法の could が能力であるか可能性であるのかは、意味論的には<意図>の有無が決定するものと思われるが、機械に<意図>を判定させることは現実的ではないため、統計的に処理できる方法を取りあえず提案しておく。これには改良の余地が大きいことは認識している。

#### 【5】前文情報

機械翻訳的には、本発表の最大の主張は前文情報の参照を提案しているところにあるだろう。would や could が過去形であるのか仮定法であるのかの判別は、前文の時制を見なければ分らない場合が多い。これは、前文の構文解析の際、たとえば格フレームを構成する時に時制情報も表示し、後続文の解析ではそれを参照する、といったやり方が考えられる。このような前文情報の参照は、たとえば日英翻訳で日本語文の省略された主語を復元するというような作業でどの道必要な機構であると思われる。この場合も、前文の格フレームの主語を参照するといったやり方が考えられるだろう。

#### 使用した翻訳ソフト

[LV]： 『LogoVista 2007 PRO』, LogoVista .

[The]： 『The 翻訳 2007 プレミアム』, 東芝 .

[Atlas]： 『ATLAS V14 翻訳スタンダード』, 富士通 .

[PC]： 『PC-Transer 翻訳 Studio 2008』, クロスランゲージ .