

## 英日・日英機械翻訳の実力

成田一

narita@lisa.lang.osaka-u.ac.jp

大阪大学言語文化部

最近、翻訳ソフトへの関心が再び高まっており、雑誌等でも取り上げられている。ここでは最新の英日・日英翻訳ソフトはほぼ全てを対象に、新聞、雑誌などにおける評価の問題点を踏まえ、個別の例文と分野別テキストによって、その処理能力を、(英語、日本語に特有の表現を含め、) 文法解析、構文処理について検証する。

### 翻訳ソフトの実力

文法構造によって分類したデータはできるだけ多くのソフトの翻訳結果を示すものの、テキストについては、文書の種類、分野によって得手、不得手があるため、「実用レベルの翻訳精度」になっているものだけを提示する。(検証しようとする文法事象や構造に関係しない部分、語や代名詞の訳の間違いや冠詞の欠落などは下線で示すが、あえて評価対象とはしない。)

### 英日翻訳ソフトの構造処理能力

ここでは英日翻訳ソフトの構造処理能力を検証するが、テキスト翻訳例に種々の文法構造が含まれるので、最小限の典型例を見るに留める。

I remembered posting the mail.  
I remembered to post the mail.  
Flying planes can be dangerous.  
I don't know how good meat tastes.  
I had my car stolen.

**解説** 第一文は「したことを憶えている」で、第二文は「忘れないでした」という意味が分かれば良い。第三文は Flying planes が「飛行機を飛ばすこと」「飛んでいる飛行機」いずれでも良いが、can be はモーダルな用法なので「できる」ではなく「であり得る」の意味にとれる訳でないで困る。第四文は how good meat tastes が「良い肉がどんな味か」もしくは「肉がどんなにおいしいか」のいずれかに訳されていたら良い。ただし、「よい肉が味わう」の類の訳は good meat tastes の意味がズレている。第五文は使役の had を「された」と訳すのが良いが、「させた」でも仕方がない。知識処理を期待することができないし、わざと盗ませる保険金詐欺もある。

『ATLAS』『翻訳サーフィン』

×私は、メールを提示するのを覚えていた。  
○忘れずに、私はメールを提示した。  
○飛行機を飛ばすのは危険な場合がある。  
○私は良い肉がどう味がするのかわからない。  
○私は自分の車を盗まれた。  
『PowerE/J』  
○私は、メールを投函したことを記憶していた。  
○私は、忘れずにメールを投函した。  
○飛行機を飛ばすことは、危険であるかもしれない。

○私は、肉がいかに良く味がするかわからない。  
△私は、私の自動車を盗まれた状態にした。  
『PC-Transer』『ポケットラ』  
○私は、メールを郵送したのを憶えていた。  
○私は、メールを郵送するのを忘れていた。  
○飛行機を飛ばすことは、危険でありえる。  
○私は、肉がどれくらいおいしい味がするかわからない。  
×私は、盗まれる私の車を持った。  
『翻訳アダプタ』『J・London』『WorldNet』『訳セゴマ』  
○私は、メールを投稿したことを覚えていました。  
○私は忘れずにメールを投稿しました。  
△飛行機を飛ばすのは危険であるかもしれません。  
△私は、どれほどよい肉が味わうかわかりません。  
○私は私の車を盗まれました。  
『LogoVista X』『コリヤ英和』  
○私はメールを公表したことを覚えていた。  
○私はメールを公表することを忘れずにした。  
○飛行機を操縦することは危険であり得る。  
○私は良い肉がどのような味がするかわからない。  
○私は私の自動車を盗まれた。  
『翻訳の王様』  
×私はメールをポストすることを覚えていました。  
○私は、忘れずにメールをポストするようにしました。  
×飛ぶ平面は危ないことができます。  
○私は、どのようによい肉が味がするかわかりません。  
○私は私の車を盗まれました。  
『たちまち翻訳』  
○私は、メールをポストしたことを覚えていた。  
△私は、メールをポストすることを覚えておいた。  
△飛行機を操縦することは、危険なはずだ。  
△私は、どのように良い肉が味がわかるかわからない。  
○私は、私の車を盗まれた。  
『The 翻訳』『辞スバ』『Roboword』  
×私は、メールを記入することを覚えていました。  
○私は忘れずにメールを記入しました。  
○飛行機を飛ばすことは危険かもしれません。  
△私は、よい肉がどう味わうかわかりません。  
×私は自動車を盗みました。  
『Dr.Surf』『翻訳革命』  
○私は、メールを投函したことを覚えていた。  
○私は、メールを投函するのを忘れなかった。  
○平面を飛ばすことは、危険かもしれない。  
△私は、おいしい肉がどのように味が分かるかわからない。  
○私は、自分の自動車を占領された。  
『バンセ』  
○私は、郵便を公示したことを覚えていた。  
○私は、郵便を公示することを忘れずにした。  
○飛行機を操縦することは、危険になり得る。  
○私は、良い肉がどのように味がするかわからない。  
○私は、私の自動車を盗まれた。

# 『国際人』

- ×私は、メールを郵送することを思い出しました。
- △私は、メールを郵送することを覚えていました。
- △飛行機を飛ばすことは危ないことができます。
- 私は、どれほどよい肉が味うかを知りません。
- 私は車を盗んでもらいました。

## 日英翻訳ソフトの構造処理能力

日英翻訳は英日翻訳と比べると全体的にはかなり見劣りするので、紙数の関係もあり、文法構造の処理能力をみるに留める。

- その母親は子供を噛んだ犬を父親に殺させた。
- その教科書を注文したのは先生だ。
- 彼は昨日その図書館に行ったそうだ。
- この子が君の本を破ってしまったらしい。
- 僕の女房は魚が嫌いみたいだ。
- 象は鼻が長い。
- 先生は原稿をお書きになっている。
- 最初に議論すべき問題は外部からの影響である。

**解説** 第一文は使役と関係節を含む。「噛んだ」の時制は過去完了でなく過去でも減点しない。第二文は強調構文である。第三、四、五文ではモーダルな表現「そうだ」「らしい」「みたい」がどう英訳されるかをみる。第五文では「魚が」が主語と解析される例もあるが、これは「嫌い」などの感情表現が目的語の助詞に「が」をとれるという情報の欠落による。第六文の二重主語構文の英訳表現は様ではないが、いずれも関係の解析はクリアしている。第七文は「おVになる」という敬語形式だが、「ている」は進行形と完了形どちらでも良い。「になる」をturned...intoとする例もある。第八文は関係節と主語無しへの対応をみる。

## 『Transland』『LogoVista』『コリヤ英和』『翻訳王様』

- That mother made father kill the dog which bit a child.
- It was a teacher that ordered that textbook.
- It is said that he went to that library yesterday.
- It is said that this child tore your book.
- My wife seems to dislike a fish.
- The nose of the elephant is long.
- A teacher is writing a manuscript.
- A problem to discuss first is an influence from the outside.
- 『ATLAS』『TransLinGo』
- The mother had the father kill the dog which had bitten the child.
- ×It is a teacher to have ordered the textbook.
- I hear that he went to the library yesterday.
- This child seems to have broken your book.
- ×The fish dislikes my wife.
- Elephant's trunk is long.
- The teacher has written the manuscript.
- The problem which should be discussed first is an influence from the outside.
- 『PC-Transer』『ポケトラ』
- The mother let a father kill the dog which bit a child.
- It is a teacher that ordered the textbook.
- He seems to have gone to the library yesterday.
- This child seems to have torn your book.
- ×A fish dislikes my wife and wants to watch it.

- As for the elephant, a nose is long.
- ×It is written a manuscript to a teacher.
- The problem that you should argue in the first place is influence from the outside.
- 『The 翻訳』
- △The mother made father kill the dog who bit the child.
- △The teacher ordered the textbook.
- It is said that he went to the library yesterday.
- Seemingly, this child has torn your book.
- ×my wife .. fish .. a dislike .
- An elephant's trunk is long.
- The teacher is writing the manuscript.
- The problem about which it should argue first is the influence from the exterior.
- 『翻訳アダプタ』『J・London』
- △The mother made a father kill the dog who bit a child.
- It is the teacher that ordered the textbook.
- He is said to have gone to the library yesterday.
- This child seems to have torn your book.
- My wife seems to dislike a fish.
- The nose of the elephant is long.
- ×The teacher becomes writing in the manuscript.
- The problem that it should discuss first is an influence from outside.
- 『訳セゴマ』
- ×The mother caused a/the father the dog which bit a/the child killed.
- ×It is a teacher that I ordered the text book.
- He is said to have gone to the library yesterday.
- ×This child seemed that have broken your book.
- ×My wife is want to observe disagreeable a/the fish.
- As for an/the elephant the nose is long.
- A/the teacher is writing a/the manuscript.
- × The problem that should discuss first is the influence from the outside.
- 『バンセ』
- The mother made father kill the dog which bit a child.
- ×It is a teacher to order the textbook.
- △He was said to go to the library yesterday.
- △This child had seemed to tear your book.
- My wife seems to dislike a fish.
- The nose of an elephant is long.
- ×A teacher has turned a manuscript into writing.
- ×The problem which should argue at first is influence from the outside.

## テキスト評価の問題点

朝日新聞夕刊(平成 11 年 9 月 2 日大阪本社)の「週刊 you・me ジャーナル」の機械翻訳を扱った記事の中で、東大入試問題を翻訳させ、訳文を河合塾の講師が評価しているが、次のような問題がある。まず、合格者の 3 割程度しか満点を取れないとする英文を評価対象とするのは妥当だろうか。それに現在の市販の機械翻訳ソフトは人工知能機能を持っているわけではない。この問題文は文脈に沿って文意を適切に解釈して翻訳することが必要だが、それは期待できないし、設計上守備範囲としていないのである。構造を解析し意味情報を整合的に駆使して翻訳するのが限界だ。

長文の客観的な評価法はないが、対象文を句ごとに区分して、判断の揺れを押えれば、相対的には信頼できる評価になる。それでも「こなれた意識」を重視すると、機械翻訳の設計上の能力を正當に捉えた評価になってこない。語彙、熟語、構文、修飾・主述関係が適切に解析、翻訳できるかといった点に絞り込んで見るのが良い。

例題は全体としてかなり正確に訳されている。コンピュータ処理としては十分に満足できるものだ。決定的なミスと断ずるような文法の間違ひはない。入試の採点でも7、8点は付けたい。

Computers do what they are told to do, whether we mean it or not. Moreover, they cannot turn themselves on, nor can they ever begin something entirely new on their own. (98年度東大入試問題)

河合塾の模範訳:「人がそう意図していようといなかろうと、コンピューターは言われた通りのことをする。それだけでなく、コンピューターは自分でスイッチを入れることもできないし、今まで一度もやったことがないことを自分だけで始めることもできない」

機械翻訳:「我々がそれを意図したか否かに関係なく、コンピューターはそれらがするように言われることをします。そのうえ、彼らはつくことができなくて、それら自身で完全に何か新しいものは始めることはできません」(『ATLAS V6』使用)

講師評:うーん、十点満点で六点ですね。和訳が硬いですねえ。「コンピューターは自分でスイッチを入れることができない」と訳せば自然なのに、「彼らはつくことができない」としているところが気になりますね。雰囲気はなんとなく伝わるのですが、文法を間違えているのが決定的なミス。ここだけで減点三です。...

解説 turn themselves on を「自分でスイッチを入れる」と訳さないことで、三点も減点するのは酷だろう。これは分離動詞 turn on が themselves を目的語とし「自身にスイッチを入れる」すなわち「自律的に起動する」の意なのだが、内部化されても目的語と解析し、しかも再帰代名詞なので「をつける」ではなく「がつく」と自動詞的に訳す配慮は捨てたものではない。on [their own] を「[それら自身の]で」と訳すのも拙いので減点されるだろうが、この「の」は名詞機能を持つ。

『LogoVista X』『コリヤ英和』

コンピューターが(彼・それ)らが、我々がそれを意味するか否かにかかわらず、するように言われることをする。さらに、(彼・それ)らは(彼・それ)ら自身をつけることができない、同様に(彼・それ)らは独力で今までに完全に新しい何かを始めることができない。

『The 翻訳』『辞スバ』『Roboword』

私たちがそれを意味しても意味しなくても、コンピューターは彼らが行うように命じられることを行います。さらに、それらはそれら自身をつけることができません。同様に、それらは常に完全に、独力で新しい何かを始めることができません。

『Dr.Surf』『翻訳革命』

私たちがそれを意味するとしても、そうでないとしてもコンピューターは、彼らが行うように言われることをする。更に、彼らは彼ら自身をつけることができない。彼らは何か全く新しいものをかつてまた自分一人で始めることができない。

『PC-Transer』『ボケトラ』

コンピューターは、彼らが我々がそれを意味するかどうか

に関係なく、するように言われるものをする。さらに、彼らは彼ら自身を駄きつけることができないし、彼らが自分で完全に新たにこれまでに何かを開始することができるというわけでない。

『翻訳の王様』

コンピューターは、我々がそれを意味するかどうかにかかわらず、するために、(彼・それ)らが言われるものをつけます。おまけに、(彼・それ)らは(彼・それ)ら自身をつけることができません、また(彼・それ)らは、独力で今までに何か全く新しいものを始めることができません。

When U.S. President Bill Clinton visited China last summer, China and the United States agreed to mutually dismantle strategic nuclear missiles targeted against each other. (『朝日イブニングニュース』平成11年8月4日付け)

英文記事の元になった朝日新聞社説:「去年夏のクリントン大統領の中国訪問で、米中は互いに戦略核ミサイルの照準を外すことで同意した」

解説 dismantle...each other が「互いに対して対象とされる戦略の核ミサイルを解体する」と訳されているが「互いを目標とする戦略の核ミサイルから装備を取り外す(dismantleの辞書訳語)」と訳せば「照準を外す」に理解できる。「取り除く」と訳すソフトもある。

『LogoVista X』『コリヤ英和』

ビル・クリントン合衆国大統領が去年の夏に中国を訪問したとき、中国と合衆国は相互にお互いに対して向けられた戦略上重要な核弾頭ミサイルを取り除くことに同意した。

『ATLAS V6』

米国ビル・クリントン大統領が去年の夏に中国を訪問したとき、中国と合衆国は、互いに互いに対して対象とされる戦略の核ミサイルを解体することに同意しました。

『PowerE/J』

米国大統領ビル・クリントンがこの前の夏中国を訪問したとき、中国、及び、米国は、相互に対してターゲットにされた戦略的核弾頭ミサイルを相互に解体することに同意した。

『PC-Transer』『ボケトラ』

ビル・クリントン米大統領がこの夏中国を訪問したとき、中国とアメリカ合衆国は相互にお互いに対して目標とされる戦略核ミサイルを取り除くことに同意した。

『国際人』

米大統領ビルクリントンが最近の夏中国を訪れる時、中国と米国は、相互に互いに対して目標にされた戦略の核ミサイルを解体することを賛成しました。

『The 翻訳』『辞スバ』『Roboword』

ビル・クリントン米国大統領がこの前の夏中国を訪問した時、中国およびアメリカは対立して目標とされた戦略の核ミサイルを相互に分解することに合意しました。

A skin stimuli generator

This small autonomous robot walks on the human body to generate a pleasing tickling sensation. This robot is equipped with four sensors to be able to detect its inclination relative to a horizontal plane. Left and right track motors enable it to steer left/right and reverse direction. When it encounters a slope that is too steep, it will steer until a safe level is found, and it will continue its way. This behavior is implemented using a hardware finite state machine (ホームページの英文だが、第二文の主語は原文では Tickle である。)

試訳(直訳)皮膚刺激生成機

この小さな自律型ロボットは快いくすぐり感覚を起こすために人間の体を歩く。このロボットは水平面に対する傾きを検知するセンサーを備えている。左右のトラックモーターによって左右と逆方向に舵がとれる。急すぎる坂に遭遇すると、安全なレベルが見つかるまで舵を取り、そのまま進む。この行動はハードウェア有限状態機械によって実現される。

**解説** 第一文では、to generate ... sensation という不定詞句が「...するために」と動詞を修飾するか、「する...」と直前の名詞を修飾するかの可能性があるが、この文では後者は不可。tickling...を「くすぐって」と分詞構文に訳す分析もみられる。また walks on...を「を歩く」ではなく「で歩く」とする訳もあるが、これは「歩く」の目的語を「を」で受けるという情報を盛り込んでいるか否かによる。第二文でも to be able to...plane という不定詞句の係り先が潜在的に曖昧だが、どちらも許容される。また、relative to a horizontal plane「水平面に対する」の訳語とこれがinclination「傾斜」を修飾しているかを見たいが、「に関する傾向」などの訳も多くみられる。ほかにplaneを「飛行機」とする例もある。第三文では、left/rightが「左右」の訳になっているか、さらにrightが優先順位の高い「正しい」になっていないかを見る。leftを「残された」とする構造的おかしな訳もある。この訳語選択は「/」での連結を条件にすることで適正化できる。なお、enable it to steer...は「をすることを可能にする」の意だが、to steer...を「するために」とし、enable itを「それを可能にする」のような訳もみられる。しかし、これはまとまった構造的な表現で不定詞句だけ分けて訳すことはできない。第四文では、that is too steep という関係節の修飾関係と副詞節 until...found and continue its way...「そのまま（その道を）進む」の訳をみる。「方法を続ける」の訳が多いが文脈的にはズレがある。第五文では using a hardware finite state machine「ハードウェア有限状態機械によって」（「有限状態機械」の「機械」は抽象的なモデルであって、これが具体物なら「ハードウェア」を冠する。）だが、stateを「州」とする訳もある。This behavior is implemented「この行動は実現される」の意味になっているかを見る。「実行/実施/遂行される」でも許容できるだろう。

『PowerE/J』

この小さい自律型ロボットは、楽しくすぐつたいセンサーションを生成するために、人間のボディ上で歩く。このロボットは、水平面と比較してその傾きを検出することができるために、4台のセンサーを装備している。左の、そして右のトラックモーターによって、それは左の/正しい、そして逆の方向に進むことが可能になる。それがあまりにも急であるスロープに遭遇するとき、安全なレベルが発見されるまで、それは、操縦し、そして、それは、その方法を続けるであろう。この行動は、ハードウェアの有限の状態のマシンを用いて実行される。

『PC-Transer』『ポケットラ』

この小さい自主的なロボットは、心地よいむずむずしているセンサーションを起こすために人体の上に歩く。このロボットは、水平面と関連しているその傾斜を見つけることができる4台のセンサーを備えている。左右のトラック・モーターは、それに左で/適切で逆の方向を一定の方向に向けるのを可能にする。それがあまりにも険しい傾斜に遭遇するとき、安全なレベルが見つけれられるまで、それは向かう、そして、それはその道を続ける。この挙動は、ハードウェア有限の状態機械を使って実行される。『翻訳アダプタ』『J・London』『WorldNet』『訳せゴマ』この小さな自治ロボットは、喜びくすぐつたい感覚を生

成するために、人体を歩きます。このロボットは、水平の面と関連するその傾向を検出することができるように、4つのセンサーを装備されます。左右のトラックモーターは、それが、出発/右および逆の方向を操縦することを可能にします。それが急勾配すぎる斜面に遭遇する時には、それは、安全なレベルが見つけれられるまで、操縦し、それはその方法を続けます。この行動は、ハードウェア限定された国マシンを使って実施されます。

『LogoVista X』『コリヤ英和』

この小さい自立したロボットは楽しくすぐつている感覚を生み出すために人間の組織体の上に歩く。このロボットは4つのセンサーで水平な飛行機と比較してその好みを検出することが可能であることができるように装備される。左と右のトラックモーターが左/権利と逆の方角の舵を取るためにそれを可能にする。それがあまりにも急な勾配に遭遇するとき、安全なレベルが達せられるまで、それは舵を取るであろう、そしてそれはその方法を続けるであろう。この行動はハードウェアの有限の州機械を使って実行される。

『The 翻訳』『辞スバ』『Roboword』

この小さな自律なロボットは愉快なくすぐる感覚を生成するために人体上で歩きます。水平面に関するその傾向を検知することができるためにこのロボットは4つのセンサーを装備しています。それは、残された軌跡モーターおよび正しい軌跡モーターによって左/右および逆方向を操縦することができます。険しい傾斜に遭遇した場合、安全なレベルが見つかるまで、それは操縦するでしょう。また、それはその方法を継続するでしょう。この振る舞いはインプリメントされます、ハードウェアの使用、有限の州機械。

テキストの翻訳について載せた例をみても分かるように、文章の種類によって「実用レベルの翻訳精度」を持つ翻訳ソフトの顔ぶれ・順位が多少入れ替わる。文ごとに含まれる構造、句、語彙の違いにより、各ソフトの処理能力に違いが生ずるのだ。しかし、これは標準的な設定のままで処理した結果であって、専門辞書を使うとか磨きをかけたユーザー辞書を使えばまた違ってくるだろう。用例翻訳機能も予め翻訳パターンとして用意されているもの以外に、ユーザーが自ら翻訳パターンを設定することのできるソフトもある。良質な対訳用例を蓄積すれば、ファジーマッチによってその分翻訳品質が改善する機能を備えたものもある。ユーザーが鍛えてやれば翻訳ソフトもどンドン育つ可能性を秘めているのだ。

以上見たように、翻訳ソフトはかなり実用的なレベルになっていることが納得できたのではないかと思うが、文芸作品にこだわる人の中には、「詩や俳句の翻訳はできないだろう」とか、「ジョークが分かるというのはとりわけ人間的な営みであって」といった視点に立って、機械翻訳にはその能力がないことを批判する向きがあるが、翻訳は文化の説明や解釈ではない。機械翻訳が守備範囲とも目的ともしていないことをもって批判するのは見当違いだが、巷にはこうした高度な知識処理までも翻訳の役割だと見る過大な期待や思い込みがあるのも事実だ。ジョークは読み手が理解するものである。分からない人は分からない。