

機械翻訳における事例の適用の一手法

武田 紀子

takeda@is.seikei.ac.jp

成蹊大学工学部経営工学科

1. はじめに

機械翻訳の質の向上のため、事例、定型パターンを用いる方法が、提案され、いくつかのシステムで実現されている。¹⁾⁻⁴⁾そして、事例による翻訳における問題の1つとして、入力文に対して、いかに事例を適切に効率よく適用させるかということが挙げられ、これに対し、いくつかの事例を組み合わせて文全体の翻訳を行う手法⁴⁾が、提案されている。しかし、入力文と、複数の事例との適当な対応付けをとるためには、多くの計算量が必要とされる。ここでは、文の構文的特徴を表す構文事例と、対訳文付き事例からなる翻訳事例の2種類の事例を用意し、まず、入力文と構文的に近い事例と、意味的に近い事例から、構文的、意味的に近い事例を作り、次に、この事例を利用して入力文の翻訳を行うことにより、複数の事例の組み合わせを容易にするという手法を提案する。

この手法を用いて、to-不定詞句、that-句を含む、中学の教科書にでてくる程度の英文を対象に、基本的な構文と例文を事例として登録し、英日機械翻訳の実験を行なったので、その結果について報告する。

2. 事例データベース

事例データベースは、構文的構造を表す構文事例と、各語ごとに対応付けられた訳、文のキーワード等からなる翻訳事例からなる。構文事例、翻訳事例は、文全体に対するものばかりではなく、名詞句、形容詞句、前置詞句等の句に関するものも定義されている。

2. 1. 構文事例

構文事例の例として、文に対する事例の構造と、その例を示す。

構文 各構文要素の文法的範疇 文の種類 動詞のパターン この構文からなる翻訳例へのポインタ

図1. 文に対する構文事例の構造

ここで、文の種類は、平叙文、否定文、疑問文、命令文、感嘆文に分類される。また、動詞パター

ン(vp)は、Hornby の動詞パターン⁵⁾を基準に分類される。

[例1] 文に対する構文事例の例

1) There is (np) (prepp). 構文

ここで、かっこで囲まれているのは、非
終端記号、囲まれてないのは終端記号

(@ there @ be np prepp) 各要素の範疇

構文を特徴付ける語は、@ の後に語その
ものを登録する

平叙文 文の種類

vp1a 動詞パターン

15 この構文を持つ翻訳例へのポインタ
翻訳例は、vp により分類されている。

2) Is there (np) (prepp)?

(@ be @ there np pp)

疑問文

vp1a

17

3) I go to (np) to (vp_inf).

(subj verb prep np @ to vp_inf)

平叙文

vp3A

16

4) How many (np) do you need?

(@ how @ many np @ do subj verb)

疑問詞疑問文

vp6A

16

次に文以外の構文事例の例を示す

[例2] 名詞句に対する構文事例の例

det adj_er np @ than np 構文

er_than_np 1 翻訳例へのポインタ

[例3] 形容詞句に対する構文事例の例

as (ap) as (np) 構文

ap_rei 1 翻訳例へのポインタ

(adj_er) than (np)

ap_rei 2

2. 2. 翻訳事例

文に対する翻訳事例の構造と例を示す。

英文
日本語訳
文の要素に分解された語のリスト
キーワードのリスト
キーワードの意味のリスト
文の種類、動詞パターン

図 2. 文に対する翻訳事例の構造

[例 4] 文に対する翻訳事例の例

- 1) There is a TV set in the room. 英文
 ("there" "the" "部屋" room) 訳文
 ("には" in) ("a" "テレビ" TV)
 ("セット" set) ("が" jyoshi)
 ("あり" (masu)) is) ("。" nil syuushi))
 ((is) (there) (a TV set) (in (the room)))
 構文要素に分解された語のリスト
 (there set) キーワードのリスト
 ((there) ((2))) キーワードの意味のリスト
 平叙文 文の種類
 vp1a 動詞パターン
- 2) How many stamps do you need?
 ("あなた" you) ("は、" subj_jyoshi)
 ("何" how many) ("枚" suushi_np)
 ("切手" stamps) ("が" dobj_jyoshi)
 ("い" ("ら" d5)) need)
 ("ますか?" ? renyou))
 ((need) (you) (stamps) (how many))
 (need stamp)
 ((actt) ((2)))
 疑問詞疑問文
 vp6a

構文事例と同様に文以外にも句に対する翻訳事例が定義されている。

[例 5] 名詞句の翻訳事例の例

- a bigger room than mine 英文
 ("わたしの" mine) ("よりも" than) ("a")
 ("大き" kei) bigger) ("部屋" room rentai))
 訳
 ((a) (bigger) (room) than (mine))
 文の要素に分解された語のリスト
 ((7)) 名詞句の意味 (place)

[例 6] 形容詞句の翻訳事例の例

- 1) as tall as Tarou 英文
 ("太郎" Tarou) ("と同じくらい" as)
 ("as) ("背が高" kei) tall)) 訳文
 (tall Tarou) キーワード
 ((3 1 1 1)) 名詞句の意味

- 2) taller than I
 (("私" I) ("よりも" than)
 ("背が高" kei) taller))
 (taller I)
 ((3 1 1))

3. 事例の適用

3. 1. 事例の適用の流れ

入力文に対する翻訳は、以下のような流れに沿って事例が適用され行われる。

[1] 入力文と構文的に似ている構文事例を見つける。ここで、類似度の高さは次のような順序で決められる。

1. 入力文が構文事例に登録されている構文と同じ構造を持つ
2. 各構文要素の文法的範疇に現れている "θ" の後の語が入力文に含まれている
3. 動詞パターン、構造が等しい

似た構文事例が見つかったら、この構文からなる翻訳事例を検索する。ここで、入力文と同じ翻訳例が見つければ、翻訳は終わる。入力文と同じ構文事例が見つからない場合は、同じ構文からなる翻訳事例を1つ選ぶ。これを翻訳事例1とする。

[2] 入力文を似た構文事例の構造に沿って解析し、翻訳事例の中から、対応するキーワード間の類似度の計算⁷⁾によって、意味的に最も近い翻訳例を選ぶ。これを翻訳事例2とする。翻訳事例2が、入力文と等しい場合は翻訳は終わる。

[3] 翻訳事例1と、翻訳事例2を組み合わせ、入力文と構文的、意味的に近い事例を作成する。これを翻訳事例3とする。

[4] 翻訳事例3と入力文を比較して、異なる語の部分置き換えて、訳を得る。

3. 2. 事例の適用

入力文に対して、どのように事例が適用され翻訳がなされるかについて、上で示した[1]~[4]の流れに沿って述べる。

[例 7] Are there many temples in Kyoto?

この入力文と似た構造を持つ構文事例として、[例 1] 2) が選ばれ、この構文からなる翻訳事例として、次のような事例が得られる。

- [1] Is there an elephant in the zoo?
 ("there" "the" "動物園" zoo)
 ("には" in) ("an" "象" elephant)
 ("が" jyoshi) ("い" (masu)) is)
 ("か?" ? syuushi))

((is) (there) (an elephant) (in (the zoo)))
 (there elephant)
 ((there) ((3 1 2)))

疑問文

vp1a

次に入力文と意味的に最も近い翻訳事例をキーワード間の類似度の計算⁷⁾によって求める。

[2] There is a TV set in the room.
 ((" there) (" the) ("部屋" room)
 ("には" in) (" a) ("テレビ" TV)
 ("セット" set) ("が" jyoshi)
 ("あり" (masu)) is) ("." nil syuushi))
 ((is) (there) (a TV set) (in (the room)))
 (there set)
 ((there) ((2 1)))

平叙文

vp1a

この2つの事例から、入力文と構文的、意味的に近い事例を作成する。

[3] ((" There) (" the) ("部屋" room)
 ("には" in) (" a) ("テレビ" TV)
 ("セット" set) ("が" jyoshi)
 ("あり" (masu)) is) ("か?" ? syuushi))
 ((is) (there) (a TV set) (in (the room)))
 (there set)
 ((there) ((2 1)))

疑問文

vp1a

[3]で得られた事例と、以下に示す構文事例に沿って解析された入力文とを比較して、異なった語のところのみ、訳を入れ替える。

(((" there))
 (((("たくさんの" many) ("寺" temples))
 ((2 1) (7)))
 (((("京都" Kyoto) ("の中に" in))
 ((prep in) ((5) (2) (3 1 1)))))

ここでは、" a TV set" と "many temples"、" the room" と "Kyoto" の訳が入れ替わり、入力文の訳が求められる。

[4] ((" there) ("京都" Kyoto) ("には" in)
 ("たくさんの" many) ("寺" temples)
 ("が" jyoshi) ("あり" (masu)) is)
 ("か?" ? syuushi)))

" 京都にはたくさんの寺がありますか?"

[例8] I have a lot of work to do.

この文はto-不定詞句を含むことから、まず、似た構文事例として、次のような構造を持つものが

選ばれる。

[1] subj @ be ap @ to vp_inf
 subj verb pp @ to vp_inf
 subj verb @ to vp_inf

しかし、これらの事例のいずれも入力文の構造とは一致しない。そこで、入力文の主動詞の動詞パターンによって似た構文事例を求める。すると、" a lot of work to do" は、名詞句の構文事例

ap noun @ to vp_inf
 により解析され、入力文の構造は、

[1] I play (np).
 (subj verb np)
 vp6a

という構文事例と似ているとみなされる。また、入力文と意味的に近い事例としては、

[2] He has homework.
 (("彼" He) ("には、" subj_jyoshi)
 ("宿題" homework) ("が" dobj_jyoshi)
 ((" (aru)) has) ("ます。" nil renyou))
 ((has) (he) (homework))

が選ばれ、これらを組み合わせることにより、入力文の訳が求められる。([3], [4])

" 私には、するたくさんの仕事があります。"

ここで、to-不定詞句の処理は、不定詞部分に、必要に応じて、" they"、" it" を加え、文の形に変換して、文の翻訳と同様に事例を適用しながら行われる。例えば、" to do" は、" They do it." に変換して、処理される。

[例9] 1) I get up earlier than my mother.
 2) I get up as early as my mother.

この2つの文は、" get up" 以降は、それぞれ副詞句の構文規則

adv_er @ than np
 @ as adv @ as np

と似た構造を持つと判断され、対応する翻訳事例との組み合わせにより、副詞句の部分は処理される。従って、この2つの文と似た構文事例としてはともに

[1] You walk (advp).
 (subj verb advp)
 vp2c

が選ばれる。次に、意味的に近い翻訳事例を探すと、1)と同じ文が、翻訳事例として登録されているので、1)は、事例から、直接、訳が求められる。2)の処理は、意味的に近い翻訳事例として、1)

タイプ	文	名詞句	形容詞句	副詞句	その他	計
数	129	17	9	6	4	165

表 1. 登録されている事例データの数

タイプ	文	名詞句	形容詞句	副詞句	前置詞句	その他	計
数	185	37	13	10	27	4	276

表 2. 登録されている翻訳データの数

と等しいものが、選ばれるので、これらを組み合わせることにより、訳が求められる。([2] - [4])

- 1) " 私は、母より早く起きます。"
- 2) " 私は、母と同じくらい早く起きます。"

4. 結果と考察

現在、中学程度の基本的な構文例、翻訳例を事例データベースに人手によって登録し(表1、表2)、中学の参考書にある埋め込み文を含む単文、324文(2-10語からなる)に対し、翻訳実験を行った。ここでは、文レベル、句レベルいずれの解析段階でも、適当な事例が見つからないと、規則ベースの翻訳システムによって、解析、翻訳がなされる。その結果、302文が、1つ以上の事例の組み合わせにより、正しい訳が得られた。

入力文の翻訳は、構文的に近い事例と意味的に近い事例の両者を組合わせて、構文的、意味的に近い事例を作成したのち、この事例を利用して行われるので、

- 1) 構文に依存した言い回し等のための事例と、語の意味の曖昧性を解消するための事例を独立に定義して、容易にこれらを組み合わせることができるので、登録する事例の数を減少させることができる。
- 2) 似た構文事例の決定が最初になされるので、翻訳事例の検索は、構文事例によって決められた動詞パターン、動詞となる語やその意味によって検索範囲に制限を加えてから行える。
- 3) 構文事例の検索は、文の種類、文と構文例に現れる語の語レベルの比較で、候補を絞ってから、似た事例が決定されるため、効率よくなされる。
- 4) 事例の組み合わせの過程で生成された翻訳事例を登録して、事例を増やすことができる。

また、文レベルのみではなく、名詞句、形容詞句等の句レベルでの事例の適用を可能にしたため、

- 1) 文全体に対して、複数の事例の組み合わせが容易になされる。
- 2) 句レベルの事例にも、構文事例、翻訳事例を用

意したため、句レベルでも、文レベル同様、複数の事例の組み合わせが可能である。

- 3) 埋め込み文の処理は、文レベルの事例を適用させて処理できる。

このように、複数の事例の組み合わせが、容易になされるため、特殊な語の用法、独自の言い回し等が、それに関する事例の追加登録によって、簡単に扱えるようになった。しかし、現時点では、句レベルの事例の適用は、文全体の処理とは独立してなされるため、文全体の意味の整合性が、とれない場合があるという問題が残った。

ここでは、中学程度の英語に関して、事例を登録し翻訳実験を行っているが、分野を限って、そこで使われる典型的な構文、言い回し等を事例として登録すれば、実用性のあるシステムになるであろう。

[参考文献]

- 1) 佐藤：MBT2：実例に基づく翻訳における複数例の組み合わせ利用，人工知能学会誌，Vol. 6, No. 6, pp. 861-871 (1991).
- 2) 古瀬，隅田，飯田：経験的知識を活用する変換主導型機械翻訳，情報処理学会論文誌，Vol. 35, No. 3, pp. 414-425 (1994).
- 3) 相沢，加藤，鎌田：外電経済ニュースの英日機械翻訳，情報処理学会論文誌，Vol. 37, No. 6, pp. 1041-1048 (1996).
- 4) Sato：MBT2: a method for combining fragments of examples in example-based translation, Artificial Intelligence, Vol. 75, pp. 31-49 (1995).
- 5) A. S. Hornby：Guide to patterns and Usage in English, second edition, Oxford University Press (1977).
- 6) 好田 監修：チャート式シリーズ、中1英語、中2英語、中3英語、研数出版 (1995).
- 7) 武田：基本的な事例構文と翻訳例を利用した英日機械翻訳，情報処理学会研究会報告，Vol. 96, No. 56, pp. 117-122 (1996).