

英文 CISG の格構造

7 B-6

門田 隆史[†] 柴田 誠[‡] 竹田 正幸[†] 松尾 文碩[†]
[†]九州大学工学部[‡]九州大学大学院工学研究科

1. まえがき

国際物品売買に関する国連条約 (CISG) の第 II 部に関する推論システムを作成した¹⁾。このシステムは、基礎原子式書換え系に基づく規則推論 (RBR) 方式を採用し、高速で安定していることから、この方式の有効性を検証することができた。一方、第一階言語に基づく形式言語は、知識表現言語としては問題規模が大きくなるにつれ、人手による取り扱いが困難になることも判明した。そこで、CISG のために知識表現言語として、格文法的構造をもった言語の設計を行うことにした。本稿では、格構造化の方法とこれまでにわかった動詞の性質について報告する。

2. 方法

格構造化する部分は、CISG の本体である第 II 部と第 III 部である。英語で書かれた本文、つまり部や章の見出し等を除いた本文における延べ単語数は、9,621 であり、異なり単語数は 966 である。ここで、単語とは大文字と小文字を区別しない英字、数字、ハイフンを字母とする文字列である。このように文章量が少ないことから格構造化を内省 (introspection) で行うことが可能であるが、この方法は情報処理技術への寄与がほとんど期待できない。そこで、できるだけ内省を少なくするため、次のような作業手順を採用した。

- 1) 英文 CISG のコーパスを作成する。このコーパスのタグは、統語情報の付与を効率的に行えるように簡単なものにした。タグが複雑になると、内省のために作業がはかどらなくなり、誤りが多くなる。

The case structure of CISG

Takashi Kadota[†], Makoto Shibata[‡], Masayuki Takeda[†] and Fumihiko Matsuo[†]

[†]Faculty of Engineering, Kyushu University 36, Hakozaki, Fukuoka, 812-81 Japan

[‡]Division of Engineering, Graduate School, Kyushu University 36, Hakozaki, Fukuoka 812-81, Japan

- 2) コーパスから CISG の格構造を得るため、辞書や文法を作成する。

- 3) 1 と 2 から CISG の格構造化を行うための手続きを開発する。

- 4) 3 を使って、自動的に格構造化を行う。このとき、1, 2, 3 における誤りが見つかるので、それらを修正し、4 の過程を繰り返す。

1 の作業は完了している。コーパスの 1 部を図 1 に示す。次の作業は上記 2 の文法を作成するために、動詞の格フレームを確定することである。その準備として CISG に生起する動詞について調査した。生起した単語が動詞かそうでないかはコーパスから簡単な処理で判定することができる。

Article 3

(1) ((Contracts) for ((the supply) of ((goods) to be manufactured or produced))) < are > to be considered (sales) (unless [((the party) [(who) < orders > (the goods)]) < undertakes > to supply ((a substantial part) of ((the materials) necessary for (such ((manufacture) or (production))))])) .

(2) (This Convention) < does not apply to > ((contracts) [(in (which)] ((the preponderant part) of ((the obligations) of ((the party) [(who) < furnishes > (the goods)]))) < consists in > ((the supply) of ((labour) or (other services)))) .

図 1 CISG のコーパスの一部

3. CISG の動詞

まず、コーパスを用いて主動詞を抽出し、それぞれの動詞がどのような意味で用いられているかを調査した。次に、現在分詞および to 不定詞を含んだ動詞について、それらが動詞としてのみ用いられているかどうかについて調査した。更に、軽動詞についても調査を行った。

3.1 主動詞

CISG に出現する主動詞の語数は 145 であった。それらのうち、cause, be, do, enter, make, take の 6 語を除

いて一つの意味でのみ用いられていることがわかった。(意味が二つ以上あったものの例を図2に示す。)これより、CISGでは、ほとんどの動詞の格フレームが一意的に決まることがわかる。

cause

1) もたらす

Article 34

... he may, up to that time, cure any lack of conformity in the documents, if the exercise of this right does not cause the buyer unreasonable inconvenience or unreasonable ...

2) ... させる (使役)

Article 38

(1) The buyer must examine the goods, or cause them to be examined, within as short a period as is practicable in the circumstances.

図2 意味が2つ以上あったものの例

3.2 現在分詞およびto不定詞を含んだ動詞

CISG に出現する現在分詞およびto不定詞を含んだ動詞の語数は172であった。それらのうち、150語(87.2%)は完全に動詞としてのみ用いられていることがわかった。のこりの22語は、表1に示すように動詞と名詞の綴りが同じものであった。このことから、CISGでは、動詞かどうかの判定は比較的容易であると考えられる。

表1 動詞と名詞の品詞をもつ単語

account	need
act	object
amount	order
claim	place
contract	redespatch
control	remedy
damage	request
deposit	result
dispatch	state
exercise	supply
follow	use

3.3 軽動詞

make, take, have, giveなどの動詞は、例えば‘make a reservation’のように動詞としての意味がないものがある。このような動詞は、light verb²⁾あるいはdelexical verb³⁾などと称されているようである。ここでは、これらを軽動詞ということにする。軽動詞の場合は、後続の名詞が実質的な意味をもつので、この名詞を合わせて格フレームを決定しなければならない。

しかし、軽動詞は表2に示すように、25回生起したにすぎない。

表2 軽動詞の生起頻度

軽動詞句	生起頻度
make a declaration	5
make no declaration	1
make such contracts	1
make the specification	3
make a purchase or resale	1
make deliverey	1
make restitution	2
take effect	3
take measures	2
take steps	2
take possession	1

4. むすび

CISGにおいては、日常英文と異なり、

- 動詞の意味的曖昧さがほとんどない
- 動詞の品詞的曖昧さがほとんどない
- 軽動詞的用法があまりない

という性質があることがわかった。そこで、動詞の格フレームを決定することは、比較的容易であると考えられる。現在、その格フレームを決定する作業を行っている。

なお、本研究は、一部文部省科学研究費補助金(#07204205)の援助により行った。

参考文献

- 1) 松尾文碩, 瀧口昌朗, 竹田正幸: 基礎原子式列書換え系による法的推論, 情報処理学会第51回全国大会講演論文集(3), pp.19-20 (1995).
- 2) Live, A.H.: The Take-Have Phrasal in English, *Linguistics*, 95, pp.31-50 (1973)
- 3) Sinclair, J.: *Collins Cobuild English Grammar*, Collins (1990)