

# 法令文の論理式への変換

信岡 俊祐 中村 誠 島津 明  
北陸先端科学技術大学院大学情報科学研究科

## 1 はじめに

安心な電子社会を実現するための研究の一環として、法令工学の研究をしている [4]。法令工学は、法令 (契約書、社内規定等を含む) がその制定目的にそって適切に作られ、論理的矛盾や文書の問題がなく、関連法令との整合性がとられていることを検査・検証し、法令の改定に対しては、矛盾なく変更や追加、削除が行われることを情報科学の手法を用いて支援する学問である。また、法令を実働化している情報システムを設計する技術を研究開発するためのものである。

この法令工学の研究として、法令文を論理演算可能な論理表現に変換する研究を進めている。法令文に関しては、吉野 [12]、岡田 [3] が独自の論理表現を提案している。平松ら [11] が法令文を解析するシステムを作っている。江尻 [2]、北田 [5] は 1 階述語表現に様相を加えた表現を出力するシステムを作っている。

本報告では、江尻、北田のシステムを引き継ぐシステムについて述べる。先行システムは、係り受け解析の後、法令文を条件部・帰結部に分割し、分割された各部分の格解析等を行い、原子式に変換し、全体をまとめるものであったが、埋め込み文や名詞句の解析等が不十分であった。また、文の構造には名詞句と格構造などの関係に再帰的なところがあるが、先行システムは段階的処理の面が強く、再帰性の扱いがソフト的に洗練されていなかった。

本研究では、先行研究で扱われた千代田区条例 53 号 (全 28 条 71 項) と富山県条例 (全 10 条 19 項) に加えて、所得税法 (全 244 条中 100 条 255 項 247 号)、国民年金法 (全 148 条中 100 条) の分析を行い、名詞句や格構造の扱える範囲を広げた。システムは Lisp で再実装することで再帰的意味解析を実現した。さらに法令文の表現の多様性に対して、論理構造への分割と同様の方法で、表現の言い換え処理をするようにした。以下、2 節で関連研究、

3 節で法令文の分析結果、4 節でそれに基づいた提案手法、5 節で比較実験を示す。

## 2 関連研究

### 2.1 法令文の構造

田中 [9] は、法令文は「要件部」と「効果部」から成るとし、以下の「要件・効果構造」を提案している。

$$\begin{aligned} & \text{【要件部】} \Rightarrow \text{【効果部】} \exists \\ & \text{【主題部} \wedge \text{条件部】} \Rightarrow \text{【対象部} \wedge \text{内容部} \wedge \text{規定部】} \quad (1) \end{aligned}$$

式 (1) に示すように要件部と効果部は、各々をさらに細かく細分化すると、「主題部」、「条件部」、「対象部」、「内容部」、「規定部」となる。平松ら [11] はその要件効果構造に基づく法律文統語構造の解析を行っている。

### 2.2 論理式への変換

法令文を論理式に変換する江尻、北田システム [2][5] は、要件・効果構造に基づき、法令文を論理構造に沿って分割し、分割された各部分の格解析等を行い、原子式に変換し、全体をまとめる。江尻 [2] は、法令文書の語順が通常文と比較してかなり定まっているという特徴や固有の言い回しが多いという特徴に着目し、式 (1) に基づいて文の分割ルールを作成し、法令文の表現パターンの特定を行っている。さらに、関係を表す語句や動詞の並列句を考慮して構造の決定を行い、法令文の骨格的な論理構造を決定している。しかし、分析対象の法令文が少なく、パターンが不十分となっている。北田 [5] は、分割された各部分の述語動詞に対して格解析を行い、原子式に変換している。しかし、格関係以外のもの、つまり埋め込み文や名詞句には十分に対応していない。また、格フレーム辞書は分析対象が多くないため、カバー率が低い。

## 2.3 名詞句

名詞句解析のための関連研究として、例えば、以下のものがある。

- 島津ら [6] は、助詞「の」が結ぶ名詞について、名詞 A、B の素性を定義して、「A の B」を意味関係により分類し、その分類とそれらの素性を用いて「A の B」の意味的關係を推定することで意味解析を行っている。
- 鳥澤 [10] は、大量のテキストに統計的教師なし学習である EM 法を適用し、得られた単語の意味クラスを用いて「A の B」の形態をもつ名詞句における AB 間に生じる意味的關係を推定して言い換えを行っている。
- 阿辺川ら [1] は、日本語の連体修飾節と被修飾名詞の關係を決定付ける要因としていくつかの要素を想定し、それらを用いて内の關係、外の關係の判別を行っている。

これらの研究を含め、日本語の名詞句を論理式へ変換する研究はあまりない。そこで、本研究では、文を構成素に分割し、述語動詞に対して格解析を行うことに加えて、埋め込み文、名詞句「A の B」を対象に分析を行い、名詞句も対応できるようにシステムを構築した。

## 3 法令文における名詞句の分析

### 3.1 名詞句に含まれる並列句

法令文に現れる接続詞は、意識的に区別して使い分けられている。法令文ではこのような使い分けは接続詞による段階構造と呼ばれている [7]。そのため、並列句を分割する際、この段階構造を考慮する必要がある。

### 3.2 埋め込み文

名詞を修飾する埋め込み文の分析を富山県条例 (全 10 条 34 項 34 文) と所得税法 (全 244 条中 100 条 255 項 247 号約 626 文) で行った。分析の結果、格關係が成り立つ埋め込み文は、延べ 849 例、異なり 646 例存在した。

### 例 3.1 格關係

「食料を販売する事業者」

販売する (e1)  $\wedge$  事業者 (x1)  $\wedge$  食料 (x2)

$\wedge$  agt(e1, x1)  $\wedge$  obj(e1, x2)

格關係がなく被修飾名詞が「こと」、「もの」、「方法」である同格は、延べ 405 例、異なり 327 例存在した。分析範囲では、これらの名詞で格關係を持つものはなかった。

### 例 3.2 同格

「権限を行使することを認められた」

行使する (e)  $\wedge$  権限 (x)  $\wedge$  obj(e, x)

$\wedge$  認める (e')  $\wedge$  obj(e', e)

上記の 2 つの關係で決まらなかった場合はその他の關係としているが、分析を行った範囲には、存在しなかった。

## 3.3 名詞句：「A の B」

名詞句「A の B」については、富山県条例 (全 10 条 34 項)、国民年金法 (全 148 条中 100 条) の分析を行った。分析から法令文に現れる名詞句「A の B」の原子式には様々な場合がある。それらのうち幾つかを以下に示す。

#### 1. 名詞 A がサ変名詞

(例) 提出の申請書

提出 (e)  $\wedge$  申請書 (x)  $\wedge$  obj(e, x)

#### 2. 名詞 B がサ変名詞

(例) 施設の設置

施設 (x)  $\wedge$  設置 (e)  $\wedge$  obj(e, x)

#### 3. 名詞 A の論理式が述語、名詞 B の論理式が関数

(例) 被保険者の氏名

被保険者 (x)  $\wedge$  氏名 (x)=n

#### 4. 名詞 A の論理式が関数、名詞 B の論理式が行為述語、名詞 A と B の關係述語

(例) 年金給付の内払

内払 (e)  $\wedge$  年金給付 (x, t, h)=m  $\wedge$  obj(e, m)

#### 5. 名詞 A の論理式が関数、名詞 B の論理式がその値

(例) 年金給付の額

年金給付 (x, t, h)=m

6. 名詞 A の論理式が 2 つの述語、名詞 B の論理式が述語  
 (例) 二十歳未満の者  
 者(x) ∧ 年齢(x)=t ∧ t < 20

## 4 システムの概要

以下のステップで論理式に変換する。

- (1) 形態素解析 [juman]
- (2) 構文解析 [KNP]
- (3) 文を構成層へ分割
- (4) 表現の言い換え
- (5) 格解析と名詞句解析を再帰的に行う
- (6) 論理式の合成

以下、主な部分を示す。

### 4.1 文の構成素への分割

構成素は、節、主題を表す句である。分析より 84 種類の構造を抽出した。これをパターン化し [2]、文とのパターンマッチを行い、要件部と効果部の構造を決定する。

【要件部】 ⇒ 【効果部】

ここで、要件部が空のこともありうる。構造を決定した後、論理式に付与する様相演算子があるか調べる。様相は「可能」、「義務」、「許可」の 3 種類である。

### 4.2 格解析

#### (1) 格フレーム辞書

全 13 都府県<sup>1</sup>の条例 (全 366 条 818 文) から動詞に係る名詞を抽出し、名詞の意味カテゴリー毎に分けた格フレーム辞書を半自動で作成した。合計 517 種類の述語動詞の格フレーム辞書を構築した。格フレームは以下の形をしている。

格フレーム  
 = (格スロット<sub>1</sub>, ..., 格スロット<sub>l</sub>)  
 格スロット  
 = (深層格 表層格 意味カテゴリー  
 ((例<sub>1</sub> 頻度<sub>1</sub>) ... (例<sub>n</sub> 頻度<sub>n</sub>)))

#### (2) 格解析手法

用言には複数の格フレームがあり、一つの格フレームに対しても複数の格スロットがあり、格スロットの埋め方で、複数の解釈がある。そのため、格要素と格スロットの照合の度合いをスコア付けし、各格スロットのスコアを足して格フレームと格要素列との照合度合いとする。格スロットに対するスコア付けは以下の手順で行う。

1. 格要素の主名詞または被修飾名詞を格スロットの用例と照合する。用例があればスコア 5 点とする。
2. 表層格と格要素の格助詞が一致するかを照合する。一致すればスコア 5 点とする。ここで被修飾名詞であれば 2.5 点とする。また、副助詞であれば 2.5 点とする。
3. 日本語語彙大系のシソーラスを用いて格スロットの用例と格要素の主名詞との類似度のスコアを下記の式 (3) により計算する。

格要素の主名詞  $w_1$  とスロット中の用例  $w_2$  の類似度を日本語語彙大系から以下の式で求める。

$$Sim(w_1, w_2) = \frac{2L}{l_1 + l_2} \quad (2)$$

ここで、 $l_1$ 、 $l_2$  はそれぞれ単語  $w_1$ 、 $w_2$  のシソーラスの根からの深さ、 $L$  は  $w_1$ 、 $w_2$  の上位意味属性で一致している階層の深さを表す。次に用例のそれぞれの頻度を  $f_1, \dots, f_n$  とすると、格スロット中のすべての用例  $w_{2_1}, \dots, w_{2_n}$  による類似度は、以下のように求める。

$$\frac{\sum_{i=1}^n Sim(w_1, w_{2_i}) \times f_i}{\sum_{i=1}^n f_i} \times 10 \quad (3)$$

### 4.3 名詞句解析

#### (1) 名詞句に含まれる並列句の分割

名詞句に接続詞が含まれるとき、接続詞の段階構造を考慮して名詞句を分割する。

#### (2) 埋め込み文の解析

名詞句が埋め込み文であるときは、埋め込み文に対して格解析を行う。このとき被修飾名詞にもスコアを付け、深層格を決定する。格関係がなく被修飾名詞が「こと」、「もの」、「方法」であるときは同格とする。

<sup>1</sup> 青森県、栃木県、奈良県、兵庫県、山口県、埼玉県、千代田区、岡山県、新潟県、岩手県、千葉県、京都府、富山県

表 1: 実験結果

		正解	誤り	正解率
本システム	動詞	40	18	69.0%
	連体修飾	38	23	62.3%
	サ変名詞	14	11	56.0%
既システム	動詞	42	24	63.6%
	連体修飾	10	22	31.3%
	サ変名詞	7	18	28.0%

### (3) 名詞句「A の B」

名詞 B がサ変名詞であるときは名詞 A は名詞 B の格要素となりうるので格解析を行い、深層格を決定する。名詞 A がサ変名詞である場合も同様に名詞 B が格要素となりうるため格解析を行う。名詞 A、B 共にサ変名詞でない場合は、名詞句「A の B」の辞書を参照して論理式を導く。辞書の例を以下に示す。

「額」

$(rel(A \text{ の } \_) (@A)(\text{保険料})(@A))$

「保険料」

$(rel(\_ \text{ の } B)(\text{額})(\text{額})(\text{保険料}(x, t, h) = m))$

## 4.4 論理式の合成

以上の処理を終えた後、各解析で得られた論理式を合成する。このとき文を構成素に分割した際に決めた構造に当てはめて全体の論理式を決定する。

## 5 比較実験

「広島県ばい捨て等の防止に関する条例」全 20 条に対して、本研究で実装したシステムと既システムとの比較実験を行った。実験には、先行研究で構築された格フレーム辞書にある用言に対する格フレームを用いた。実験結果を表 (1) に示す。なお、既システムは連体修飾をすべて扱っておらず、一部を動詞として扱っているため、総数が異なると思われる。

## 6 おわりに

本稿では、特に、先行研究で実装したシステムの改良について述べた。今回、法令文の分析に基

づいて解析対象を広げ、名詞句一般が扱えるようにし、論理構造への分割パターンを増やし、格解析の手法も改良した。法令文の表現の多様性に対しては、論理構造の分割と同様の方法で、表現の言い換えを行った。名詞句「A の B」の解析範囲をさらに広げる必要があり、今後の問題である。

## 参考文献

- [1] 阿辺川武, 奥村学, 「日本語連体修飾節と被修飾名詞間の関係の解析」, 情報処理学会, 2004-NL-160, 2004.
- [2] 江尻暁, 「法令文の論理式への変換—論理構造の生成—」, 言語処理学会, 第 12 回年次大会 2006.
- [3] 岡田光弘 「法的知識の論理による表現」, 法律人工知能, 2000.
- [4] 片山卓也, 検証進化可能電子社会 -情報科学による安心な電子社会の実現- 情報処理, Vol146, N05, pp515-521, 2005.
- [5] 北田安希雄, 「法令文の論理式への変換—原始文の生成—」, 言語処理学会, 第 12 回年次大会 2006.
- [6] 島津, 内藤, 野村, 「助詞「の」が結ぶ名詞の意味関係の解析」, 計量国語学会, 第 15 巻 7 号, 1986.
- [7] 田島信威, 「法律の読解法」, ぎょうせい
- [8] 田中喜久雄, 「法律条文の標準構造」, 自然言語処理 1993.
- [9] 田中喜久雄, 「法律効果規定部の意味機能について」, 自然言語処理, 1998.
- [10] Kentaro Torisawa, *A Nearly Unsupervised Learning Method for Automatic Paraphrasing of Japanese Noun Phrase*, Proceeding of the Workshop on Automatic Paraphrasing, pp.63-72, Dec, 2001.
- [11] 平松, 永井, 中村, 野村, 「要件効果構造に基づく法律文統語構造解析」, 自然言語処理, 1997.
- [12] 吉野一, 「法律エキスパートシステムの基礎」, ぎょうせい