

Title	法令文の論理式への変換
Author(s)	信岡, 俊祐
Citation	
Issue Date	2007-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/3599
Rights	
Description	Supervisor: 島津 明, 情報科学研究科, 修士

修 士 論 文

法令文の論理式への変換

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

信岡 俊祐

2007年3月

修 士 論 文

法令文の論理式への変換

指導教官 島津 明 教授

審査委員主査 島津 明 教授
審査委員 白井 清昭 助教授
審査委員 東条 敏 教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

510076 信岡 俊祐

提出年月: 2007 年 2 月

概要

法令文書を検査・検証することを目指し、法令文を論理式に変換する方法の研究をしている。先行研究では、係り受け解析の後、法令文を論理構造に分割し、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。そこで本研究では、名詞句を再帰的に解析するようにし、また、法令文の分析に基づいて解析対象の種類を広げた。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。本稿では、上記の点ならびに、先行研究で実装したシステムの改良について述べる。

目次

第1章	はじめに	1
1.1	研究の背景と目的	1
1.2	本論文の構成	2
第2章	関連研究	3
2.1	法令文の構造	3
2.2	法令文の論理式への変換	4
2.3	名詞句解析	7
第3章	法令文の分析	9
3.1	法令文の分割に関する分析	9
3.1.1	主題部と対象部	9
3.1.2	条件部	9
3.2	名詞句に関する分析	12
3.2.1	名詞句に含まれる並列句	12
3.2.2	埋め込み文	13
3.2.3	名詞句「AのB」	15
3.2.4	言い換えに関する分析	19
第4章	論理式への変換処理	21
4.1	全体の流れ	21
4.2	文の構成素への分割	21
4.3	表現の言い換え	24
4.4	格解析	25
4.4.1	格フレーム辞書	25
4.4.2	格解析手法	26

4.5	名詞句解析	29
4.5.1	名詞句に含まれる並列句の分割	29
4.5.2	埋め込み文	29
4.5.3	名詞句「AのB」	30
4.6	論理式の結合	31
第5章	実験	32
5.1	実験結果	32
第6章	おわりに	35

目 次

2.1 要件・効果構造	3
-------------------	---

表 目 次

3.1	主題部と対象部	10
3.2	条件部の従属節	10
3.3	並列句の分析	13
3.4	埋め込み文の分析	15
3.5	名詞句「A の B」の分析	15
3.6	言い換えの分析と変換後の形	20
4.1	様相表現	24
5.1	実験結果 1	32
5.2	実験結果 2	33

第1章 はじめに

1.1 研究の背景と目的

現在、社会生活の様々な分野でIT技術による電子化が進み、電子社会の時代を迎えている。電子社会は、我々の社会システムのうちで情報システムによって実現されている部分である。電子社会の役割は多様な社会活動の支援であり、安心して生活できる基礎でなければならない。その意味で、電子社会の安心性は電子社会に対する最も重要な要求である。そこで、安心な電子社会を作るための研究の一環として、法令工学という学問が提案されている [5]。

法令工学とは、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、論理的矛盾や文書的問題がなく、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。

本研究では、この法令工学研究の一環として、法令文を論理演算可能な論理表現に変換する研究を行った。本報告では、特に、先行研究で実装したシステムの改良について述べる。先行研究 [2][6] では、係り受け解析の後、法令文を論理構造に分割し、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、元となる法令文の分析の範囲も少なく、埋め込み文や名詞句の解析等が不十分であった。また、論理式への変換のタスクを「法令文全体の論理構造への変換」と「要素の原子式への変換」に分割して、段階的に実行する方式をとっているため、2つのシステムのリンクが正常に働かない場合がある。そのため、分割は出来ていても格解析を行うことが出来ない場合などの問題や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、再帰性の扱いが洗練されていないなどの問題がある。

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行うようにした。

1.2 本論文の構成

本稿は、次の構成をとる。2節では関連研究について述べる。3節において法令文の分析結果を示し、それに基づいた提案手法を4節で述べる。5節で具体的な解析手法を説明し、6節で実験を行う。最後に7節でまとめを行う。

第2章 関連研究

2.1 法令文の構造

一般的に、我が国における法令文は、個々の条例に共通の概念あるいは抽象度の高い概念を規定する条文を前にもってくるといった原則の他に、「要件・効果論」といわれる原則で記述されている [14]。その「要件・効果論」を採用して記述されている法令文は、(1)「…は、…するものとする。」といった、その条例が規定したい権利・義務関係の内容である法律効果と、(2)「…のときは」、「…においては」といった社会関係を示す法律要件が含まれ、(3)一つの要件部に対しては一つの効果部を対応させるという形式をとっている。このような法令文の構造的特徴に着目した研究として田中ら [9] の研究がある。

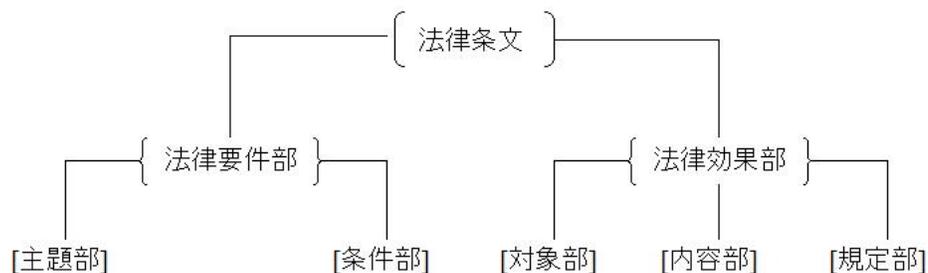


図 2.1: 要件・効果構造

田中ら [9] は、法律条文は「要件・効果論」をもとに目的的に作られたものであり、このため具体的条文の構造もその特殊性に影響された形での構造的特徴を持ち、さらに語順で考えた場合でもそれがかなり定まっているということも法律条文の特徴であると示し、このような特徴を留意して、法令文の構造を、図 2.1 のように与えている。ここで図 2.1 に示すように、法令文は「要件部」と「効果部」に大別できる。さらにそれぞれを細かく分化させると、要件部の主題を示す主題部、その条件を示す条件部、条文の対象を示す対象部、その内容の規定を示す内容部、その内容の規定を示す規定部となる。

また平松ら [13] は、この要件・効果構造に基づく法律統語構造の解析を行うシステムを開発している。

吉野 [14] は、法的知識を表現するのに論理流れ図および複合的述語論理という手法を提案している。

岡田 [4] は、法律文をそのままの自然言語で形式的論理表現に移し得る論理的知識表現の枠組みを与えるとともに、その知識表現から計算処理可能なエキスパートシステム用の法的知識表現への変換が論理的に同じ意味を保っていることを保証しながら得られる知識表現の変換のための論理的方法論の枠組みについて考察し、独自の論理表現を提案している。

2.2 法令文の論理式への変換

法令文を論理式に変換する方法として、江尻、北田ら [2][6] の研究がある。江尻、北田システム [2][6] 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、全体をまとめるものである。

江尻 [2] は、法令文のもつ特徴を利用して条文の構造を解析し、全体的な論理構造を決定する手法を提案している。法令文は、上述にあるように個々の条文に共通の概念等を規定する条文を前にもってくるといった原則の他、「要件・効果論」と呼ばれる原則に基づき記述されている [14]。そのため、法令文は、以下のような特徴を持っている。

- “条件、付帯状況、結論”の形で記述されることが多いという構造的な特徴
- 語順が一般自然言語と比較すると、非常に定まっているという特徴
- 固有の言い回しが多いという特徴

これらの特徴に着目し、分割のルールを作成し、要件・効果構造に合わせて例 2.1 のように表現のパターンの特定を行っている。

例 2.1(千代田区、条例 53 号、23 条 3 項)

区長は、前項により届出のあった協定が、内容等に関し適切なものであると認めるときは、これを認証し、告示するものとする。

分割された語句

主題部 : 区長
条件部 : 内容等に関し適切なものであると認める
内容部 : 前項により届出のあった協定
規定部 : 認証し、告示する

表現のパターンの特定

[主題部] ∧ [条件部] [内容部] ∧ [規定部]

ここで、例 2.1 を図 2.1 と比較すると対象部が欠けていることや出現順が異なることが分かる。構造内の一部の要素が欠ける場合は ϕ (ゼロ)化の問題と呼ばれており [9]、このような構造を考慮して分割するルールとして法令文を 7 つのパターンに分類している。そして、「ため、」、「して」、「により、」等の意味の上で述語動詞の関係を表す語句を考慮した構造の決定を行った後、動詞の並列句を考慮した構造の決定を行い、法令文の全体的な論理構造を決定し、法令文の骨格的な部分を論理表現へと変換している。しかし、分割されるべき箇所では分割されていないことや、限られた法令文を対象に実装されているため表現のパターンが少なく対応できな文もある。

北田 [6] は、江尻のシステムによって分割された各部分の述語動詞に対して格解析を行い、原子式へ変換している。原子式へ変換するためには、述語動詞と名詞との意味的な関係を表す深層格を決定する必要がある。そこで、実際の法令文を基にして、各述語動詞がどのような深層格をとっているのか、その名詞が深層格としてとられた頻度、深層格の名前、付随する表層格を記載した 129 種類の述語動詞に対する格フレーム辞書を構築している。構築した格フレーム辞書を用いて、各述語動詞に対して格解析を行っている。格解析手法としては、格解析の対象となる述語動詞がとりうる深層格を調べ、とりうる深層格と

してのスコア付けを行い、閾値を超え、最もスコアの高い文節を深層格として決定し、原子式に変換している。しかし、格フレーム辞書を分析対象が多くないため、カバー率が低く、格フレーム辞書にないものが多く存在する。また、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ないことなどの問題がある。

2.3 名詞句解析

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 [7] は、助詞「の」が結ぶ名詞の意味関係の解析として、「A の B」の意味関係を決定するための意味的な素性を定義している。素性は以下のように定義している。

- 主素性 語の素性のうち、その語の主要な意味特徴を示す素性
「イス」: thing、「イヌ」: animate、「遊び」: action
- 依存素性 語の素性のうち、他の語との意味的依存関係を示す素性
「日本人」: [belong-to nation]
- 機能素性 他の語との結合の仕方や、結合における役割を示す素性
「人間」: 人間は動作に対し、動作主 (agent) という格役割を持ちうることから、[role agent]

つづいて、これらの意味素性に従って「A の B」を以下のように大きく 5 分類し、さらに意味的な依存関係を示す素性に従ってさらに詳しく 86 分類している。

- B が述語相当語で、A はその述語の格要素である場合 (意味関係が格関係であるとき)
- B が、後続の語に対して、A を基点にして格的な役割を示す場合 (意味関係が場所用の指定のとき)
- B が A の属性である場合 (意味関係が属性の指定のとき)
- A が述語相当語で、B はその述語の格要素である場合 (意味関係が格関係のとき)
- A が B の一種の属性値であるとみなせる場合 (意味関係が数量等による限定のとき)

「A の B」の意味関係の解析に A あるいは B の主素性により、上記の 5 分類のうち、どの意味関係にあるかを予測する。そして、この予測を用いて素性ユニフィケーションを行い、意味関係を求めている。

鳥澤 [11] は、大量のテキストに統計的教師なし学習である EM 法を適用して、得られた意味クラスを用いて名詞句「A の B」について A と B の間に成立する意味的關係を推定している。例えば、「レストランのビール」という入力に対して、出力は「レストランで飲むビール」というように名詞句「A の B」の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関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/外の関係の判別を行っている。これらの研究では、名詞句の意味的關係や内/外の関係の判別を行っているが論理表現への変換に関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」を対象として分析を行い、名詞句も対応できるようにシステムを構築した。

第3章 法令文の分析

3.1 法令文の分割に関する分析

法令文の全体的な論理構造を明らかにするためには、法令文のどのような特徴に着目して文を構成素に分割するべきかを知る必要がある。そのため、実際の法令文を基に分析を行った。分析結果を以下に示す。

3.1.1 主題部と対象部

文を構成素に分割するための要素の一つとなる主題部と対象部について、先行研究 [2] で行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) の分析に加えて、所得税法 (全 244 条中 100 条 255 項 247 号) を基に分析を行った。ここで主題部と対象部の形態素の一覧を表 3.1 に示す。なお、主題部と対象部の形態素に関しては先行研究 [2] での分析で得られた形態素と同様であったため [2] の表を抜粋した。

表 3.1 より、主題部と対象部の形態素は同様のものであり、省略される場合が多い。そして、形態素も「は、」, 「が、」, 「も、」の 3 種類に限られているため文を分割するためには形態素を考慮するのは有効であると考えられる。

3.1.2 条件部

千代田区条例 53 号 (全 28 条 71 項) と富山県条例 (全 10 条 19 項) の分析に加えて、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) を基に条件部となる部分の分析を行った。条件部の従属節に関する分析結果を表 3.2 に示す。

分析結果より、条件部の従属節には「～するときは、」, 「～においては、」, 「～については、」, 「～においては、」, 「～により、」の 5 種類であることが分かった。そして、上記の 5 種類に加えて、「～するとき、又は～するときは、」のように条件部を表す部分が並列して現れる場合がある。実際の例を例 3.1 に示す。なお、本研究では文の構成素の分割の際に、

要件効果構造で示された内容部と規定部を分けることなく規定部として文を分割する。

表 3.1: 主題部と対象部

	形態素
主題部	は、 が、 も、
対象部	は、 が、 も、

表 3.2: 条件部の従属節

従属節のタイプ	形態素
～するときは、	動詞+とき+は+句読点
～については、	名詞+について+は+句読点
～においては、	動詞+において+は+句読点
～する場合には、	動詞+場合+に+は+句読点
～により、	名詞+に+より+句読点
～において、～するときは、	において+句読点+動詞句+ときは+句読点

例 3.1 (国民年金法、第 13 条 1 項)

社会保険庁長官は、前条第四項の規定により被保険者の資格を取得した旨の報告を受けたとき、又は同条第五項の規定により第三号被保険者の資格の取得に関する届出を受理したときは、当該被保険者について国民年金手帳を作成し、その者にこれを交付するものとする。

主題部 : 社会保険庁長官は
条件部 : 被保険者の資格を取得した旨の報告を受けた
条件部 : 第三号被保険者の資格の取得に関する届出を受理した
帰結部 : 国民年金手帳を作成し、その者にこれを交付するものとする

また、法令文では仮定条件文を記述するときの規則として「～において、～するときは、」の順序で記述するとされている [3]。文の分割にはこのような場合も考慮する必要がある。以下に実際の例を示す。

例 3.2 (国民年金法、第 22 条 1 項)

政府は、障害若しくは死亡又はこれらの直接の原因となつた事故が第三者の行為によつて生じた場合において、給付をしたときは、その給付の価額の限度で、受給権者が第三者に対して有する損害賠償の請求権を取得する。

主題部 : 政府は
条件部 : 障害若しくは死亡又はこれらの直接の原因となつた事故が第三者の行為によつて生じた場合
条件部 : 給付をした
帰結部 : 受給権者が第三者に対して有する損害賠償の請求権を取得する。

上述で示した分析結果をもとに、先行研究 [2] で述べられた ϕ 化の問題、主題部と条件部の出現順序が逆の場合、条件部が並列の場合等を含め、84 種類の文の構造を抽出した。

3.2 名詞句に関する分析

法令文を論理式に変換するシステムにおいて、先行研究 [6] では十分に扱えなかった名詞句に対応するため、法令文における名詞句の分析を行った。名詞句に含まれる並列句、埋め込み文、名詞句「A の B」を対象に行った分析結果を以下に示す。

3.2.1 名詞句に含まれる並列句

法令文に現れる接続詞は、日常用語としては同じような意味で区別せずに使われている接続詞でも法令上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 [8]。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。そこで千代田区条例 53 号 (全 28 条 71 項) と富山県条例 (全 10 条 19 項) をもとに名詞句に含まれる並列句の分析を行った。分析結果を表 3.3 に示す。分析の結果、「及び」と「又は」が頻繁に出現している。これは、同じレベルで並列句を構成するときは「又は」と「及び」を使われるためである。実際の法令文の例を例 3.3 に示す。

例 3.3 (富山県条例、第 10 条 2 項)

前項に規定する者は、その土地、建物又は工作物に廃棄物が捨てられ、又は落書きがされ、貼り札、チラシ等が放置されているため地域の良好な生活環境を損なう状況にあるときは、その廃棄物等を自らの責任で処理しなければならない。

下線部の並列句を分割した場合

(土地 ∨ 建物 ∨ 工作物)

同じレベルで構成されている並列句は例 3.3 のように分割できるが、「その他の」という接続詞のときは特に注意が必要である。この場合は節 4.5 で述べる。

表 3.3: 並列句の分析

接続詞	事例数
1: $NP_1, NP_2 \dots NP_{k-1}$ 及び NP_k	26 例
2: $NP_1, NP_2 \dots NP_{k-1}$ 又は NP_k	18 例
3: $NP_1, NP_2 \dots NP_{k-1}$ その他の NP_k	6 例
4: $NP_1, NP_2 \dots NP_{k-1}$ その他 NP_k	4 例
5: $NP_1, NP_2 \dots NP_{k-1}$ 若しくは NP_k	3 例
合計数	57 例

3.2.2 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) を基に行った。分析結果を表 3.5 に示す。分析の結果、連体修飾語と被修飾名詞との間に格関係が成り立つ場合が 646 例存在した。そして、格関係がなく被修飾名詞が「こと」、「もの」、「方法」である同格は 327 例存在し、今回、分析を行った範囲ではこれらの名詞で格関係を持つものは存在しなかった。

また、格関係がなく、同格でもない場合はその他の関係としているが、分析を行った範囲には、存在しなかった。ここで、格関係を持つ場合と同格の場合を論理式に変換した例を幾つか示す。

例 3.4 (千代田区条例、第 12 条 4 項)

容器入りの飲料又は食料を販売する事業者は、空き缶、空き箱等の容器及び包装若しくは袋の散乱防止について消費者の啓発を行うとともに、その販売する場所にこれらを回収する設備を設けるなど、適正な回収及び資源化に努めなければならない。

下線部の格関係がある場合を論理式に変換すると以下のようなになる。

販売する (e1) \wedge 事業者 (x1) \wedge 食料 (x2) \wedge agt(e1, x1) \wedge obj(e1, x2)

例 3.5 (国民年金法、第 90 条 3 項)

第一項の規定による 処分を受けた被保険者 から当該処分の取消しの申請があつたときは、社会保険庁長官は、当該申請があつた日の属する月の前月以後の各月の保険料について、当該処分を取り消すことができる。

下線部の格関係がある場合を論理式に変換すると以下のようなになる。

受ける (e1) \wedge 処分 (e2) \wedge 被保険者 (x1) \wedge agt(e1, x1) \wedge obj(e1, e2)

例 3.6 (富山県条例、第 2 条 1 項 2 号イ)

アに掲げる機関の職員であつて法律上独立に 権限を行使すること を認められたもの

下線部の格関係がない場合で被修飾名詞が「こと」である場合を論理式に変換すると以下のようなになる。

行使する (e) \wedge 権限 (x) \wedge obj(e, x) \wedge 認める (e') \wedge obj(e', e)

表 3.4: 埋め込み文の分析

	延べ数	異なり数
格関係	849 例	646 例
同格	405 例	327 例

表 3.5: 名詞句「A の B」の分析

	延べ数	異なり数
富山県条例	33 例	21 例
国民年金法	392 例	206 例

3.2.3 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 34 文)、国民年金法(全 148 条中 100 条 396 文)の分析を行った。分析から法令文に現れる名詞句「A の B」の論理式には様々な場合がある。詳しく分析するため、名詞句「A の B」を人手で論理式に変換した。名詞句「A の B」を原子式に変換したものを A、B の種類に分けて以下に示す。

1. 名詞 A がサ変名詞の場合

「提出の申請書」

提出(e) ∧ 申請書(x) ∧ obj(e, x)

2. 名詞 B がサ変名詞の場合

「施設の設置」

施設(x) ∧ 設置(e) ∧ obj(e, x)

3. 名詞 A の論理式が述語、名詞 B の論理式が関数の場合

「第二号被保険者の収入」

第二号被保険者 (x) \wedge 収入 (x, t)=m

「被保険者の氏名」

被保険者 (x) \wedge 氏名 (x)=n

「被保険者の保険料」

被保険者 (x) \wedge 保険料 (x, t, h)=m

4. 名詞 A の論理式が行為述語、名詞 B の論理式が関数で、名詞 A と B の関係述語が加わる場合

「未支給の年金」

\neg 支給 (e) \wedge 年金 (x, t, h)=m \wedge obj(e, m)

5. 名詞 A の論理式が述語、名詞 B の論理式が行為述語で、名詞 A と B の関係述語が加わる場合

「第三者の行為」

第三者 (x) \wedge 行為 (e) \wedge agt(e, x)

6. 名詞 A の論理式が関数、名詞 B の論理式が行為述語で、名詞 A と B の関係述語が加わる場合

「乙年金の支払」

支払 (e) \wedge 乙年金 (x, t, h)=m \wedge obj(e, m)

「甲年金の内払」

甲年金 (x, t, h)=m \wedge 内払 (e) \wedge obj(e, m)

「年金給付の内払」

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

「保険給付の支払」

支払 (e) \wedge 保険給付 (x, t, h)=m \wedge obj(e, m)

「年金給付の支払」

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

「年金の内払」

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

「貯金の払出し」

貯金 (x)=m \wedge 払出し (e) \wedge obj(e, m)

7. 名詞 A の論理式が関数、名詞 B の論理式がその値の場合

「国庫負担の額」

国庫負担 (x, t, h)=m \wedge 政府 (x)

「保険料の額」

保険料 (x, t, h)=m \wedge 被保険者 (x)

「支払金の金額」

支払金 (x, t, h)=m \wedge 被保険者 (x)

「年金の額」

年金 (x, t, h)=m \wedge 被保険者 (x)

「年金給付の額」

年金給付 (x, t, h)=m \wedge 被保険者 (x)

「老齢基礎年金の額」

老齢基礎年金 (x, t)=m \wedge 被保険者 (x)

「遺族基礎年金の額」

遺族基礎年金 (x, t)=m \wedge 被保険者 (x)

「基礎年金拠出金の額」

基礎年金救出金 $(x, t)=m \wedge$ 被保険者 (x)

「障害基礎年金の額」

障害基礎年金 $(x, t)=m \wedge$ 被保険者 (x)

「付加年金の額」

付加年金 $(x, t)=m \wedge$ 被保険者 (x)

「保険料納付済期間の月数」

保険料納付済期間 $(x, t)=m \wedge$ 被保険者 (x)

「保険料半額免除期間の月数」

保険料半額免除期間 $(x, t)=m \wedge$ 被保険者 (x)

8. 名詞 A の論理式が 2 つの述語、名詞 B の論理式が述語の場合

「二十歳未満の者」

者 $(x) \wedge$ 年齢 $(x)=t \wedge t < 20$

「六十歳以上の者」

者 $(x) \wedge$ 年齢 $(x)=t \wedge t \geq 60$

9. 名詞 A の論理式が述語、名詞 B の論理式が関係を示す述語の場合

「特別区の区長」

特別区 $(y) \wedge$ 区長 (x, y)

10. 名詞 A の論理式が行為述語、名詞 B の論理式が行為述語で、名詞 A と B の関係述語
が加わる場合

「婚姻の届出」

婚姻 $(e1) \wedge$ 届出 $(e2) \wedge$ obj $(e2, e1)$

3.2.4 言い換えに関する分析

異なった表現でも同じ意味を表す言語表現に対して同じ論理表現を出力することで表現の多様性を吸収することを目的として、富山県条例第 54 号 (全 10 条 34 項)、千代田区条例 53 号 (全 28 条 81 項)、所得税法 (全 244 条中 100 条 255 項 247 号) をもとに分析を行い、言い換えを行う候補を抽出した。分析の結果を、表 3.6 に示す。名詞句を中心に言い換えを行うこととしたが一部、動詞句も含まれている。また、表中に含まれる SN とはサ変名詞を表す。例 3.7 に言い換えを行う例を示す。

例 3.7 (所得税法、第六十三条 1 項)

居住者の営む不動産所得、事業所得又は山林所得を生ずべき事業について、その事業の遂行上生じた売掛金、貸付金、前渡金その他これらに準ずる債権の貸倒れその他政令で定める事由により生じた損失の金額は、その者のその損失の生じた日の属する年分の不動産所得の金額、事業所得の金額又は 山林所得の金額の計算上、必要経費に算入する。

山林所得の金額の計算上

↓

山林所得の金額を計算するとき

例 3.7 のような「NP の SN 上」という形をしたものは所得税法に頻繁に出現した。「NP を SN するとき」とは表現は異なっているが意味は同じであると考えられる。

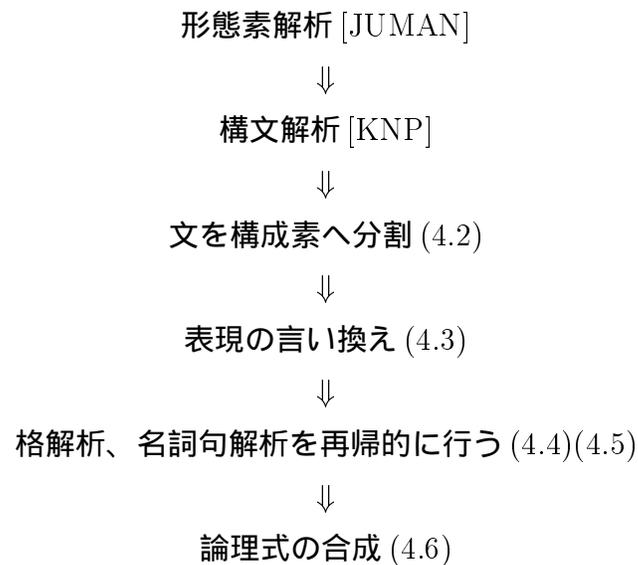
表 3.6: 言い換えの分析と変換後の形

変換前		変換後	事例数
1: NP の SN による NP	→	NP に SN する NP	16 例
2: NP の SN を行う NP	→	NP を SN する	8 例
3: NP の SN をする場合	→	NP を SN するとき	6 例
4: NP にあっては	→	NP の場合	5 例
5: NP の SN 上	→	NP を SN するとき	4 例
6: NP の SN に当たっては	→	NP を SN するとき	3 例
7: NP をもって	→	NP により	3 例
8: NP の SN のために	→	NP を SN するため	3 例
9: NP への SN がされた	→	NP へ SN された	2 例
		合計	60 例

第4章 論理式への変換処理

4.1 全体の流れ

本研究では以下に示すようなタスクにより、法令文を論理式に変換する。



4.2 文の構成素への分割

法令文の骨格的な論理構造を決定するため、文を構成素に分割する。ここで構成素は、節、主題を表す句である。節 3.1 で示した分析により、抽出した 84 種類の構造をパターン化した。これらのパターンを用いて、法令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。このとき要件部は空のときもありうる。文にパターンを適用する順

を以下に示す。なお、記号に関しては江尻 [2] と同様に主題部を SU、条件部を CO、対象部を OB、内容部を含む規定部を RE として論理構造を示す。

条件部が 2 つ存在するパターン

条件部の左に主題部が存在する
条件部が 2 つ存在する場合かつそれらが並列でないパターン

$$SU \wedge CO \wedge CO \quad RE$$

条件部の右に主題部・対象部が存在する
条件部が 2 つ存在する場合、かつそれらが並列でないパターン

$$CO \wedge CO \quad OB \wedge RE$$

条件部の左に主題部が存在する
条件部が 2 つ存在する場合、かつそれらが並列のパターン

$$SU \wedge CO \wedge CO \quad RE$$

条件部の右に主題部・対象部が存在する
条件部が 2 つ存在する場合、かつそれらが並列のパターン

$$CO \wedge CO \quad OB \wedge RE$$

条件部が1つ存在するパターン

条件部の左に主題部が存在する
条件部が1つ存在するパターン

$SU \wedge CO \quad RE$

条件部の右に主題部・対象部が存在する
条件部が1つ存在するパターン

$CO \quad OB \wedge RE$

主題部がなく、(対象部の有無も含む)
条件部が存在するパターン

$CO \quad RE$

条件部が存在しないパターン

主題部があり(対象部の有無も含む)
条件部が存在しないパターン

$SU \quad RE$

対象部があり
条件部が存在しないパターン

$OB \wedge RE$

表 4.1: 様相表現

言語表現	様相
1: なければならない	O
2: ものとする	O
3: できる	M
4: してはならない	$\neg P$

上述したパターンと文のマッチングを行い、要件部と効果部の論理構造を決定する。ここまでの処理を終えた後、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種類であり、それぞれ「可能」を記号M、「義務」を記号O、「許可」を記号Pで表す。また、様相表現と様相の対応を表 4.1 に示す。本研究で使用した様相は [2] を参考にした。実際に例 4.1 で法令文を様相を用いて論理式で表す。

例 4.1 (千代田区条例、第 13 条 2 項)

公共の場所において、チラシ等を配布し、又は配布させた者は、そのチラシ等が散乱した場合においては、速やかにこれを回収し、当該公共の場所の清掃を行わなければならない。

$O(\text{回収する}(e1) \wedge \text{agt}(e1, x1) \wedge \text{obj}(e1, x2) \wedge \text{者}(x1) \wedge \text{チラシ等}(x2) \wedge \text{清掃する}(e2) \wedge \text{agt}(e2, x1) \wedge \text{obj}(e2, x3) \wedge \text{公共の場所等}(x3))$

4.3 表現の言い換え

同じ意味を持つ言語表現でも、その表現が異なれば出力される論理式も異なってしまふ。このような法令文における表現の多様性を吸収するため、表現の言い換えを行う。節 3.2.4 において、表現の言い換えを行う候補を抽出したこと、及び言い換えを行ったあとの言語表現を示した。これらの分析をもとに言い換え前のパターンと言い換え後のパターンを作成し、文を構成素へ分割する処理で述べた手法と同様の方法で作成したパターンとマッチングを行い、表現の言い換えを行う。

4.4 格解析

節 4.1 で法令文を文の構成素に分割することを述べた。この分割した各部分を原子式に変換する。原子式は、動詞と名詞の意味的な関係を示す深層格を表現するものである。そのため、それぞれの動詞、またはサ変名詞に対して格解析を行う必要がある。以下では、格解析に用いた格フレーム辞書と格解析の手法について述べる。

4.4.1 格フレーム辞書

本研究では、青森県条例「空き缶等散乱防止条例」全 20 条、栃木県条例第 40 号「生活環境の保全等に関する条例」全 72 条、奈良県条例第 31 条「ポイ捨て防止に関する条例」全 15 条、兵庫県条例第 8 号「たばこの吸い殻及び空き缶等の投げ捨ての防止等に関する条例」全 12 条、山口市環境基本条例第 128 号全 33 条、さいたま市条例第 189 号「空き缶等のポイ捨て及び飼い犬のふんの放置の防止に関する条例」全 22 条、千代田区生活環境条例全 28 条、岡山県条例第 84 号「議会情報公開条例」全 30 条、新潟県条例第 57 号「情報公開条例」全 33 条、岩手県条例「議会情報公開条例」全 37 条、千葉県条例「情報公開条例」全 34 条、京都府条例第 17 号「情報公開条例」全 20 条、富山県条例第 54 号「情報通信技術の利用に関する条例」全 10 条の条例 (全 366 条 818 文) から各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。その結果、計 517 個、217 種類の述語動詞の格フレーム辞書を構築した。辞書には、深層格、表層格、意味カテゴリー、用例、その用例の出現頻度が記録されている。また、格フレーム辞書を半自動で作成したというのは深層格を判断する必要があるため、その部分は人手で行った。以下に格フレームの形を示す。

格フレーム

= (格スロット₁, 格スロット₂, ..., 格スロット_l)

格スロット

= (深層格 表層格 意味カテゴリー ((用例₁ 頻度₁)... (用例_n 頻度_n)))

上述した格フレームの形にそって構築した格フレーム辞書の実際の例を「管理する」の

辞書を用いて例 4.2 に示す。なお、KNP の出力を格フレーム辞書を構築するシステムの入力としているため、用例を示す部分にその名詞の読みや品詞等が含まれている。そのため上記の形とは多少異なって見えるが内容は同じであり、格解析にも読みや品詞等は用いらず、用例を用いる。

例 4.2 「管理する」

```
((OBJ (を) ("0533")) ((公文書 こうぶんしょ 公文書 名詞 普通名詞 1)))  
(AGT (は) ("0003")) ((議長 ぎちょう 議長 名詞 普通名詞 1)))  
((OBJ (を) ("1001"))  
((公文書 こうぶんしょ 公文書 名詞 普通名詞 1) (文書 ぶんしょ 文書 名詞 普通名詞 1)))  
((OBJ (を) ("0533"))  
((公文書 こうぶんしょ 公文書 名詞 普通名詞 1) (文書 ぶんしょ 文書 名詞 普通名詞 1)))  
((OBJ (を) ("1001")) ((公文書 こうぶんしょ 公文書 名詞 普通名詞 1)))  
(AGT (は) ("0003")) ((議長 ぎちょう 議長 名詞 普通名詞 1)))
```

4.4.2 格解析手法

格解析における手順は以下の通りである。

1. 格解析の対象とする動詞の格フレームのリストを参照する。
2. 解析対象の格助詞の数に着目し、格フレームのリストから、以下の方法で候補を抽出する。
 - 2(a). 格の数が一致しているかどうかを判別する。解析対象が埋め込み文である場合、被修飾名詞も格とみなしてカウントする。格の数が一致していないものは候補から取り除く。
 - 2(b). 格フレームの格と格要素の格が一致するかどうかを判別する。このとき、名詞句解析 (4.5 参照) で埋め込み文の被修飾名詞に付与した hirentaishushoku マーカーは格フレームのどの格にも一致するようにして、対応する格フレームを抽出する。また、格フレーム辞書にある表層格、又は格要素の格が副助詞「は」のときも埋め込み文と同様にどの格フレームの格とも一致するようにした。

3. 用言には複数の格フレームがあり、また一つの格フレームに対しても複数の格スロットが登録されている。それぞれの格スロットに対して、複数の解釈があるため、抽出した格フレームの候補に対してスコア付けを行うことで、最も高いスコアを持つ候補を一つ選択することにした。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。
- (a) 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。
 - (b) 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。
 - (c) 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 (4.2) により計算する。

格解析の対象が埋め込み文であるときは例 4.3 に示すように埋め込み文の文構造を変えて格解析を行う。ここで、被修飾名詞に付随している (hirentaishushoku) は、被連体修飾名詞であることを示し、格解析におけるスコア付けの対象となるため名詞句解析で埋め込み文であると判断した場合に付与する。

例 4.3 (埋め込み文の文構造の変形)



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。

$$Sim(w_1, w_2) = \frac{2L}{l_1 + l_2} \quad (4.1)$$

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, f_2, \dots, f_n とすると、格スロット中のすべての用例 $w_{2_1}, w_{2_2}, \dots, w_{2_n}$ による類似度は、以下のように求める。

$$\frac{\sum_{i=1}^n Sim(w_1, w_{2_i}) \times f_i}{\sum_{i=1}^n f_i} \times 10 \tag{4.2}$$

4.5 名詞句解析

4.5.1 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必要があることは節3.2で述べた通りである。しかし、前述でもあるように接続詞が「その他の」であり、「A、Bその他のC」という場合は、CはA、Bの上位概念を表しているため、分割する際に注意する必要がある。実際の例を以下に示す。

例 4.4 (富山県条例、第1条)

この条例は、県の機関等に係る申請、届出その他の手続等に関し、電子情報処理組織を使用する方法その他の情報通信の技術を利用する方法により行うことができるようにするための共通する事項を定めることにより、県民の利便性の向上を図るとともに、行政運営の簡素化及び効率化に資することを目的とする。

下線部の並列句を分割した場合

(機関に係る申請 届出) { (手続等) → (申請) → (届出)

例4.4は、「手続等」が「申請」、「届出」の上位概念を表しているため、論理表現の「手続等」には「申請」、「届出」を含まないという表現とした。

4.5.2 埋め込み文

名詞句が埋め込み文であるときは、被修飾名詞も連体修飾語の格要素となりうるため格解析を行う必要がある。埋め込み文の連体修飾語に対して格解析を行うとき、格要素が被修飾名詞であることが判別でき、かつ被修飾名詞に対するスコア付けを行うため、マーカーを付与する。マーカーを付与した埋め込み文を例4.5に示す。また、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して処理を行う。

例 4.5 (被修飾名詞に対するマーカー付与)



4.5.3 名詞句「AのB」

名詞句が「AのB」という形をしているときの処理方法を以下に示す。

- 名詞Bがサ変名詞であるときは名詞Aは名詞Bの格要素になりうるため格解析を行い、深層格を決定する。
- 名詞Aがサ変名詞であるときも同様に名詞Bが格要素となりうるため格解析を行い、深層格を決定する。
- 名詞A、Bが共にサ変名詞でない場合は、名詞句「AのB」の辞書を用いて論理式を導く。

名詞句「AのB」の辞書の例を以下に示す。

「額」

(rel (Aの_) (@A) (保険料) (@A))

「保険料」

(rel (_のB) (額) (額) (保険料(x, t, h) = m))

辞書には、名詞A、Bどちらの辞書であるかを示す部分、意味カテゴリー、用例、及び分析により得た論理表現を記載する。

4.6 論理式の結合

ここまで、法令文を構成素に分割して、その全体的な構造を決定し、その分割した各部分に対して格解析、名詞句解析を行って原子式を生成してきた。

そしてこれらの処理を終えた後、各解析で得られた原子式を合成する。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。以下に例を示す。

区長は、重点地区を指定し、変更し、又は解除するときは、規則で定める事項を告示するものとする。



- (1) 区長は、
- (2) 重点地区を指定し、変更し、又は解除する
- (3) 規則で定める事項を告示するものとする。



区長 (x1) \wedge (指定する (e1) \vee 変更する (e1) \vee 解除する (e1))
 \wedge agt(e1, x1) \wedge obj(e1, x2) \wedge 重点地区 (x2)
 \wedge agt(e2, x1) \wedge obj(e2, x3) \wedge 事項 (x3)
 \wedge O(告示する (e2) \wedge obj(e3, x3) \wedge loc(e3, x4)
 \wedge 規則 (x4) \wedge 定める (e3))

第5章 実験

前述してきた内容に基づくシステムを作成し、JUMANによる形態素解析、KNPによる構文解析を行い、その出力結果をシステムの入力として実験を行った。実験は、節4.4で構築した格フレーム辞書を用いて、各述語動詞に係る名詞の深層格を決定するため、埋め込み文の主動詞が修飾する名詞の深層格を決定するため、名詞句「AのB」でAあるいはBがサ変名詞のときAあるいはBの深層格を決定するため、格解析を行った。以下に実験の結果を述べる。

5.1 実験結果

実験には、「広島市ぼい捨て等の防止に関する条例」全20条に出現する61個の述語動詞に対してそれに係る名詞の深層格を決定するため格解析を行った。ここで格フレームを参照せず、表層格をもとにして深層格を割り当てたものをベースラインとした。そして本システムでは格フレーム辞書を参照して格解析を行い、格フレーム辞書にない動詞や格が存在したときはベースラインで割り当てた深層格とした。その結果を表5.1に示す。

表 5.1: 実験結果 1

		正解率
述語動詞	ベースライン	61.22 %
	本システム	78.57 %

また、ベースラインの手法を用いらず、格フレーム辞書を参照して格解析を行う方法で61個の述語動詞、49個の埋め込み文の主動詞、25個の名詞句「AのB」のAかBのサ変名詞に対して格解析を行った。解析結果を表5.2に示す。

実験は、動詞またはサ変名詞とそれに係る文節の係り受け関係を正確に抽出できているかどうかを判定し、係り受け関係が取れているものに対して格解析の結果、正しい原子式が得られたかどうかで評価を行った。また辞書の格不足とは、格要素の数に対して格フ

表 5.2: 実験結果 2

		正解	誤り			
			KNP による 構文解析ミス	動詞が 辞書にない	辞書の 格不足	その他
広島市 ばい捨て防止条例 (3条～20条)	普通動詞	55/98	3	16/61	13	4
	埋め込み文	18/90	1	14/49	31	2
	A の B	5/25	0	19/25	1	0

レームに登録されている格の数が足りないことを表す。

実験を行った結果、普通動詞に関しては予想以上に低い正解率となった。その原因としては、格フレーム辞書を構築する際に用いた条例が生活環境に関する条例と情報公開に関する条例の2種類であったため格フレーム辞書に登録した動詞やその動詞に対応する格に偏りが生じたのも原因の一つだと考える。そのため格フレーム辞書を拡張させても登録されていない動詞や格があった。しかし、先行研究のシステムでは格解析の精度が64%だったのに対して、本システムでは78%の精度を得た。また、埋め込み文とサ変名詞を含む名詞句「AのB」の実験結果については以下で考察する。

まず、連体修飾に関しては格フレーム辞書の格不足が多い。これは、格フレーム辞書を構築するときに埋め込み文の被修飾名詞を考慮していなかったため格フレーム辞書に登録されていないことが原因で格不足が起こったと考える。しかし、出力されたもの、つまり格フレーム辞書に用言連体修飾語が登録されており、格不足もない場合は、高い確率で解析できているため格フレーム辞書を構築する際に埋め込み文の被修飾名詞を考慮して作ることにより、埋め込み文に対する格解析の精度は向上すると考える。以下に正しく解析できた例を示す。

(広島市条例、第7条2項)

飲料を自動販売機により販売する者は、当該飲料に係る容器の回収用の箱等を当該自動販売機に附置し、及び当該箱等を適正に管理するよう努めなければならない。

↓

者 (x1) \wedge 飲料 (x2) \wedge 自動販売機 (x3)
 \wedge 販売する (e1) \wedge agt(e1, x1)
 \wedge obj(e1, x2) \wedge int(e1, x3)

容器 (x1) \wedge 飲料 (x2) \wedge 係る (e1)
 \wedge obj(e1, x1) \wedge loc(e1, x2)

そして、名詞句「AのB」では格解析の対象となるサ変名詞に対応する述語動詞が辞書に登録されていない例、及び格不足が目立つ。しかし、「AのB」に関しても格解析を行い、深層格を決定できたものはほぼ正しく解析できている。そのため、この場合もまた考慮して格フレーム辞書を構築する必要がある。以下に正しく解析できた例を示す。

(広島市条例、第15条3項)

喫煙制限区域の指定は、その区域を告示することにより行う。

↓

喫煙制限区域 (x1) \wedge 指定 (e1) \wedge obj(e1, x1)

第6章 おわりに

本論文では、法令文を論理式に変換する方法として特に、先行システムを引き継ぐシステムについて述べた。

具体的には、文を構成素に分割するパターンや名詞句、格構造の解析範囲を広げるため、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) の分析を行った。これらの分析に基づいて、文を構成素に分割するための 84 種類のパターンを作成し、パターンマッチで分割した各部分に対して格解析や名詞句解析を再帰的に行えるようにシステムを構築した。

そして、格解析に用いる格フレーム辞書を半自動で作成して、広島県条例の各動詞、埋め込み文、サ変名詞を含む名詞句「A の B」に対して格解析の実験を行うことでシステムの動作を確認した。

実験結果では格解析の対象となる動詞が格フレーム辞書に登録されていないことや格不足が目立つが埋め込み文の連体修飾語や名詞句「A の B」に含まれるサ変名詞が辞書に登録されており、格不足がない場合は正しい結果が得られているものも多い。

今回、格フレーム辞書を構築するときに用いた条例は、生活環境に関する条例と情報公開に関する条例の大きく分けて 2 種類であったため格フレーム辞書に登録されていない動詞がまだ多く存在する。今後は多種多様な条例を用いて辞書を作成する必要がある。さらに、埋め込み文の格解析に対する格不足が多いため被修飾名詞を考慮して格フレーム辞書を作成しなければならない。また、さらに多くの名詞句「A の B」の解析範囲を広げる必要があることも今後の課題である。

謝辞

本研究を進めるにあたり、多大なご指導、ご支援を頂いた島津明教授、白井清昭助教授、中村誠助手、山田寛康元助手に心から感謝致します。また、同期の皆様にも深く感謝致します。

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] 上田章, 笠井真一「条例規則の読み方・つくり方 全訂新版」, 学陽書房, 2000.
- [4] 岡田光弘「法的知識の論理による表現」, 法律人工知能, 2000.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 田島信威, 「法律の読解法」, ぎょうせい
- [9] 田中喜久雄, 「法律条文の標準構造」, 自然言語処理 1993.
- [10] 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.
- [11] 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.
- [12] 信岡俊祐「法令文の論理式への変換」, 言語処理学会, 第13回年次大会 2007.

- [13] 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.
- [14] 吉野一, 「法的知識の表現方法」, 法律エキスパートシステムの基礎, ぎょうせい, 1986.