

表層の手がかりによる六法全書法律文での要件部・効果部の抽出手法

角田 達彦 清水 仁 長尾 眞

京都大学 工学研究科 電子通信工学

〒606-01 京都市 左京区 吉田本町

e-mail: tsunoda@kuee.kyoto-u.ac.jp

要旨

本稿では、六法全書法律文の大局的構造の解析と要件の意味推定を表層の手がかりによって行なう手法を提案する。文の構成要素を主題、要件、効果に分け、それらが対比構造をなしているかを調べ、その結果によって各主題や要件の係り先を特定する。そして各要件の機能表現によって要件のさす内容を特定する。同時に主題の連体修飾部や、効果部に入りこんだ要件の抽出を行なう。その結果、六法全書の条文 181 文の学習コーパスに対して 170 文 (93%) が、そして 275 文のテストコーパスに対して 224 文 (81%) が正しく解析できた。また、とりたて助詞「は」と読点の有無が対比構造の生成・認識の鍵となり、それによって係り先が決定されることを明らかにした。

キーワード 法律文, 要件効果, 対比構造, 構文解析, 意味解析, 表層の手がかり

A Method of Extracting Conditional-part and Effect-part from Legal Sentences on a Compendium of Laws Using Surface Clues

Tatsuhiko TSUNODA Hitoshi SHIMIZU Makoto NAGAO

Department of Electronics and Communication, Kyoto University

Yoshida-honmachi, Sakyo, Kyoto 606-01, Japan

e-mail: tsunoda@kuee.kyoto-u.ac.jp

Abstract

We propose a method of automatic detection of global structure and semantical logics in legal sentences. Firstly, the method extracts elements in them and classifies them into three types: subject, condition, and effect. Second, it checks whether they have comparison structures, and, depending on the result, specifies their dependency. Finally, it grasps their contents using surface clues and extracts conditions from the subject and effect parts. Our method achieved 93% correctness for 181 training sentences, and 81% correctness for 275 unseen sentences. We also clarified the importance of particle 'ha' and commas for generating and understanding comparison structure, which decides phrase dependency.

Keywords Legal sentences, Condition and effect, Comparison structure, Syntactic analysis, Semantic analysis, Surface clues

1 はじめに

法律文は、人間の社会的行動の基本的な規範を明文化したものであり、法律を職業とするしなに関わらず、その社会に生きる様々な人間がその理解を求められる。それはあらゆる事態を考慮しているため量が多く、社会の動きとともに常に更新されてもいる。また、言葉の定義のゆれなどによって解釈が問われるため、常に質的に検討され続けている。このため、個人的にも、社会的にも、その内容を調査し適宜解釈することの需要と重要性がある。

だが、常に質的にも量的にも更新され続けているすべての法律文を記憶し、それらすべてを個々の事例に適用して判定をすることは、大変手間がかかることである。また、法律を専門としない人にとっては、専門用語や決まった言い回し、背景となる知識、解釈の複雑さの問題などが障壁となり、その結果法律を十分に検討することができていないのが現状である。

このような問題に対処する一つの方法として、法律文をコンピュータ上に論理構造の形で蓄積すること [1] が考えられる。膨大な量の法律文の中から、解決しようとしている問題に適用できる可能性のある文に絞り、その後人間が詳細に検討する。場合によってはコンピュータに推論させ [2]、人間が見落とす可能性のあるものまで網羅的に調べる。また、あらかじめ論理構造の形で蓄積されていれば、法律の専門家でなくても、法律文を解釈する手助けとなる。

さて、法律文であるが、確かに法律用語には適用範囲などの解釈の曖昧さがあり、社会的な問題にもなっている。しかし、論理的な意味の構造とそれを支える文章構造や文内構造に関しては、解釈のゆれが少ないように立法の段階で入念に検討され作られてきた。例えば日本の法律は、パンデクテン形式というものをとっている [3]。その特徴は、共通の概念や抽象度の高い概念は、条文の前に示していることや、要件・効果論という原則をとっていることなどがあげられる。それらの構造をとらえ、コンピュータ上の論理的表現に変換することは、個々の意味的解釈に大きく踏み込むことなく実装でき、また有意義であると思われる。

視点を変え、言語現象の中の法律文の位置付けであるが、上述のように法律文はあらゆる人に読まれることを想定して記述された自然言語文である。したがって、確

かに法律特有の表現 [4] はいくつかあるものの、言語の用法の微妙な違いなどに注意して作られている。一般の文章は複雑に様々な言語現象が絡み合っているためになかなか明確にできない現象で、法律文を解析することによつて的確に取り出せ、かつ一般的に成り立つものも少なくないと思われる。

このような背景から、本研究は法律文の要件・効果構造を特定し、その意味的な関係を抽出することを目的とする。これが的確にできれば、標準的な構造に変換し、検索に供することが容易であると考えられるからである。既存の形態素解析ツール JUMAN [5] と構文解析パーザ KNP [6] を用いて係受け関係を特定し、表層の手がかりにより主題と要件部と効果部を分ける。そしてそれらの対比構造 (図 1) などの大局的構造をとらえ、係り先を決定する。また要件部の意味を表層の手がかりにより推定する。主題とその連体修飾部、効果部内に入りこんだ要件部なども特定し、その意味の推定を行なう。以下、法律文の特徴、対比構造、本研究で提案する解析手法、実験、考察の順で詳しく述べる。

2 六法全書法律文の構造

日本の法律条文は、仏独などの大陸法系の国の法制度を継承し、民法・刑法などの各領域の法典をもとに作られてきた [3]。その記述はパンデクテン形式というものに従っていることが多い。その一つの特徴は、「要件・効果論」という原則に基づくことである。その原則では、法律条文には、その条文が規定する法律効果 (legal effect) と、その条件となる法律要件 (legal condition) とが含まれている。そしてそれぞれの要件は一つの効果のみにかかる¹。田中ら [3] は要件部と効果部の構造をさらに分析し、要件部を主題部と条件部に、効果部を対象部、内容部、そして規定部に分けることを提案している。さらに田中ら [3, 7] は各法律文からこの構造をとらえ、次のような標準構造という形に変換することを提案している。

$$V \left[\begin{array}{l} \text{は} \\ \text{が} \end{array} \right] W \left[\begin{array}{l} \text{ならば} \\ \text{の場合} \end{array} \right], X \left[\begin{array}{l} \text{は} \\ \text{の} \end{array} \right] Y \left[\begin{array}{l} \text{を} \\ \text{に} \\ \text{と} \end{array} \right] Z \left[\begin{array}{l} \text{(する)} \end{array} \right].$$

ここで、V は主題部、W は条件部、X は対象部、Y は内容部、Z は規定部を表す。この構造は様々な条文に共通な基底的な構造を考えただけのものである。各法律文は表面的な表現には様々なパターンがあるとしても、原則的には「要件・効果」を単位とする同一構造を持つと考えら

¹ ただし、後で述べるように、要件が他の要件の一部にかかっている場合もある。

れる。そこで、個々の条文をこのような標準構造に交換しても、その意味・論理構造が保存されると主張している。ただし個々の条文には次のような変化が起こりうることが指摘されている。

1. 主題化と代名詞化: 効果部の名詞句(主語ないし目的語)が主題化操作により文頭に複写されるが、効果部内の元の名詞句が代名詞の形で残ることが多い。
2. 関係節化: 効果部の格要素(主題化された場合も含む)を要件が修飾することがある。
3. 要件の語彙編入: 独立の要件部を欠くが、語彙的に要件となることがある。
4. 代用表現化: 要件を代用表現によって繰り返すことが多い。条文特有の表現。
5. 埋め込み: 要件が他の要件に埋め込まれることがある。
6. 逆転構造: 基本構造の「要件・効果」という配列が逆転した「効果・要件」形式の条文もある。ただし、その文の言語形式は限られている。

法律文を検索するという立場から見ると、上のような標準構造を想定することはきわめて妥当であり、本研究も基本的にはその枠組に交換することを目標とする。具体的にはそのための足がかりとして、個々の条文中で次のような処理を行なう。

1. 主題の特定
2. 主題の連体修飾部の特定
3. 各要件の特定
4. 既存の構文解析パーザと本研究で提案する対比構造解析手法を用い、各要件や効果の間の係受け構造を特定する。

5. 効果部に係る格の特定

さらに、実際に検索を行なう場合にはその意味的關係まで考慮することが重要であると考え、上に加えて、次のことも行なうことにした。

1. 各要件の意味的分類を行ない、それを支えるパターンをとらえる(表1)。
2. 効果部に要件が入りこむことがあるため、その部分とその意味を特定する

その詳しい内容と抽出方法は4章で述べる。

3 対比構造と機能表現

ここでは、法律文の構造を支える表現の中から、日常の言語にもあてはまるものを考察することを試みる。

3.1 法律文の構成要素と文節の依存関係

一つの法律文に現れる文節数は概して多いため、それらの係受けの曖昧さが組合せ爆発を起こすと思われがちであるが、いくつかの表層的な手がかりによってそれが避けられることが多い。そのうち、一般の言語にも法律文にも共通するものは、読点、とりたて助詞「は」による主題化と対比構造、並列構造、機能表現などである。

3.1.1 読点による構成要素の切りだし

法律文における読点の実用性はすでに永井ら[8]によって指摘されている。読点によって各要件の主なものや効果部がまとめられている。これを彼らの呼び方に従い、ここでも構成要素と呼ぶことにする。構成要素中の各格要素は、機能表現を超えて、後ろに位置する構成要素の語に係らず、構成要素内で文構造が閉じている場合が多い。このことは一般の書き言葉にも見られる現象で、読点は文内の大局的なまとまりを良く示していると思われる。

3.1.2 とりたて助詞「は」と対比構造

個々の構成要素が、他のどの構成要素に係っているかを判定する重要な手がかりは、実際の条文を調べた結果、各構成要素の末尾に、とりたて助詞「は」が使われているか否かであることがわかった。とりたて助詞の「は」は、主題化を行なう機能と、対比を行なう機能を持っており、言語学的にはそれらは区別されて論じられることが多い。しかし、実際に利用される時は、それらの意味の両方を担っていることも多く、厳密にどちらかに決めることは難しい。特に法律文では適用する状況を限定する必要性が常にあるために、他の状況から区別する対比の役割が意識される。また、ここで扱っているように読点を伴う場合には、たとえ文の主題とはならなくとも、対比してとりあげ強調する作用が強くなる。

しかし、「は」の機能は単独では決まらず、文全体の構造で決まるものである。

一つの特徴的構造は、「Aの場合は、Bし、Cの場合は、Dする。」という対比構造である。その一例を図1に示す。人間はこのような文は全体を一読した後再び見直し、構造をとらえ直しているものと考えられる。このとき、「Aの場合は、」は「Bし、」のみに、そして「Cの場合は、」は「Dする。」のみに係る。この構造の前に「Eにおいては、」などがついた、「Eのとき

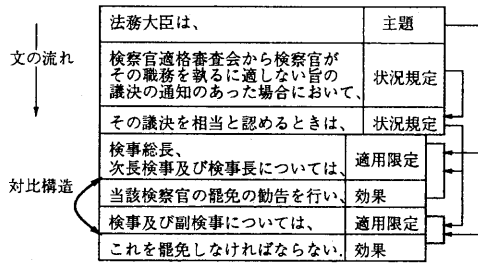


図 1: 法律文での対比構造の例

は、Aの場合は、Bし、Cの場合は、Dする。」という文のようなときは、「Eのときは、」の部分はその以降全体に係るものと認知されることが多い。このときのそれぞれの「は」の機能は対比的意味が強く、主題化の意味は弱くなる。「Eのときは、」の部分が主題になったり、それらが複数ある場合も同様である。法律文の場合には、これ以上複雑な文構造になることはめったにない。

これに対して、対比構造でない場合は、「は、」を持つ構成要素は、主文の述部（効果部全体）を直接修飾する。逆に、「は」を持たない要件は、直後に「は」を持ち述部をもつ要件や主題がある場合には、その述部に係り、そうでない場合は、直接効果部の末尾側の述部に係るように意識される。

このように、構成要素が「は、」を持つか否かが文構造全体を決める重要な指標となる。ただし、「れば」や「ならば」なども、それに類する意味を持つものと考えられるので、それに含めるものとする。これに対し、同じとりたて助詞である「も」の場合には、主題化や対比の要素が大変弱い。ここでは便宜的にこれらの違いを「規定度の強さ」と表現することにする。その他にも個々に見ると、規定度が強いものが個別に存在する。後で述べる学習コーパスを用いて調べた結果を表 1 にまとめた。

これをまとめると、次のような方法で係り先を決定すればよいことがわかる。

1. 対比構造を持つ場合：

- 対比構造内の要件は直後の効果のみに直接係る。
- 対比構造外の要件は対比構造全体に係る。

2. 対比構造を持たない場合：

- 「は、」で終れば、効果部の末尾側の述部に直接係る。

る。

- 「は、」で終らないときは、

- (a) 直後の要件や主題に述部があるとき、その末尾側の述部に優先的に係る。
- (b) 直後の要件や主題に述部がないとき、効果部の末尾側の述部に係る。

要件の係り先と表現に着目したものととして、平松らの研究 [9] がある。要件を主要要件部、それを修飾する付帯的要件部、要件部の格要素となる制限的要件部に分類し、それらを示す代表的な機能表現をとらえることを示しているが、実際には上のように、要件の文全体の構造によって動的に決まる場合が多い。

3.2 各構成要素の意味を示す表現

田中らの研究 [3, 7] や平松らの研究 [9] にも示されているように、要件を表す機能表現はある程度限られている。ここではさらにその内容を特定することを考える。日常生活でも、このような機能表現とその内容の対応はかなり守られていると考え、その一つの指針として、法律文での対応をとらえる。たとえば、「～場合には、」という表現は状況を規定するものである。その対応は、学習コーパスから取り出した表現を拡張した結果、表 1 のようなものになった。

3.3 連用接続

効果部に入りこんだ要件の典型的なものは、「～して一する」という形をとる。それは手段の意味を持つ。具体的には、効果部の述部に係るタ系連用テ形、基本連用形などの形の用言と、それに係る格すべてをまとめる。ただし、主節と並列構造をなす場合もあり、その判定は難しい。本研究では、KNP の並列構造の推定の結果と、「又」など周囲の語の存在によって判定しているが、一般にはその違いはかなりの意味的推論を要するものと思われる。

4 要件・効果構造の解析手法

要件効果構造の解析では、要件、主題、効果の特定、その内容の推定、主題の修飾部の特定、効果部内に入りこんだ要件の特定、その他の格の特定が必要である。具体的な処理とその順番は、次のようなものになる。

1. 前処理：旧仮名使いから新仮名使いへ変換する。
2. 形態素解析、構文解析：JUMAN, KNP を使用する。
3. 主題の特定：規定度の強い構成要素のうち、要件とな

表 1: 要件部の分類とその手がかりとなる表層的パターン

意味分類	規定度	手がかりとなる表層的パターン
状況規定	強い	ときは、ときには、際しては、あたっては、 当たっては、場合、場合には、 場合においては、場合の外は、限りは、 おいては、上は、の際は、のさいは、 うにについては、るについては、 れば、ならば、
	弱い	とき、ときに、際して、あたって、当たって、 場合、場合に、場合において、場合の外、 限り、際し、おいて、上、の際、のさい、 うについて、るについて
状況付加	弱い	ときも、ときにも、際しても、あたっても、 当たつても、場合も、場合にも、 場合においても、場合であつても、 場合の外も、おいても、上も、の際も、 のさいも、うについても、るについても
別途制定	強い	で定めるところにより、<数>により
状況依存	弱い	事情によって、事情により
手段規定	弱い	によって、により、もつて、以て、よつて
目的規定	弱い	よう、ように、目的で、 <動詞基本形>うため
適用限定	強い	対しては、ついでには、閉しては
	弱い	限度において、限度として
適用付加	弱い	対しても、ついても、閉しても
内容規定	弱い	ついて、閉して
時間規定	強い	までは、までには、以内には、以内は、前は、 前には、後は、後には、までの間は、間には、 間は、中は、期間内は、時は、時には、 後においては、前においては、中においては
	弱い	まで、までに、以内に、内にも、後にも、 後にも、までの間も、間にも、中も、期間内も、 時も、時でも、何時でも、後においても、 前においても、中においても
時間付加	弱い	までも、までも、以内にも、内にも、後も、 後にも、までの間も、間にも、中も、期間内も、 時も、時でも、何時でも、後においても、 前においても、中においても
場所規定	強い	内においては、所においては、
	弱い	内において、所において
場所付加	弱い	内においても、所においても
適用判断	弱い	旨
接続語	強い	但し、ただし
例外規定	強い	除いては、除く外は、除くほかは、 除くの外は、のほかは、の他は、の外は
	弱い	除き、除いて、除く外、除くほか、除くの外、 のほか、の他、の外
原因規定	弱い	ので
適用緩和	弱い	でも、かかわらず、よらないで、を問わず
適用制限	弱い	のうち
異種適用	弱い	ものとして
原則規定	弱い	に従って、に従い、に基づき、に基づいて、 に基づき、に基づいて、に対応じ、対応して
反則規定	弱い	に反して、に反し

る機能表現を持たないものを主題として特定する。

4. 効果部の特定：

- KNP の構文解析結果を利用し、主節の並列節を特定する。
- 他の構成要素で効果部になるものを特定する。読点で区切られた構成要素のうち、要件の機能表現を持たないもので、動詞・助動詞の連用形 [、] [{{ 又 | また }} [は]] | [{{ 且つ | かつ }}] で終るもの、あるいは動詞の基本形 + 述語接続助詞「とともに」で終るもの。

5. 効果部の主節に対する係り方：動詞の連用形で終るものは従属節とし、動詞の連用形・基本形 + 接続詞で終るものは並列節と判断する。

6. 効果部内に入りこんだ要件の特定：効果部内の動詞の連用形と、それに係るものすべてをまとめて要件とする。その意味は「手段」とする。ただし、並列節全体の場合や、「わたって」、「起算して」、「れて」、「あわせて」などで終る場合は例外とする。

7. 構成要素の意味の推定：要件や格の種類を示す機能表現を各構成要素に適用し、意味を決める。最長一致順に適用する。どの機能表現とも一致しない場合は、次の構成要素と単純に接続する。このような場合の多くは名詞連続だからである。

8. 要件・格を示す構成要素の係り先の特定：文末から決定していく。効果部、規定度の強い要件部、主題の現れ方によって、それぞれの要件の「は」が主題化の意味が強い対比の意味が強いかを決め、それより文頭側の各要件や格がどこに係るかを特定する。

最後の、構成要素の係り先は、次のような、それより文末側の構文のいくつかの類形によって決まる。

1. 効果部と要件部が対になったものが複数あり、それらが対比になっている場合：それより前にある構成要素は、それらすべての対に係る。
2. 上のような対比構造でない場合：規定度の強い次の要件部に用言がある場合²には、それに係るとする。そうでない場合は、すべての要件を超え、次の効果部に係るとする。

このために、効果部と規定度の強い要件部の現れ方を文末側から調べ、同時にそれらに用言があるかどうかを調

² 厳密には、格や外の関係などの制約を意味的に満たすことが必要であるが、ここでは用言があれば係ると近似する

べておく。

5 実験と評価

実験には、三省堂の「模範六法」を用いた。文語文を除く約46,000文の中から、100文おきに抜き出した458文を形態素解析ツールJUMANと構文解析ツールKNPで処理した。このうち2文は文が長過ぎるために解析できなかったため、残りの456文を用いた。このうち181文を学習コーパスとし、275文をテストコーパスとした。学習コーパスをみて表層表現と内容の対を取り出し、その末尾を変化させてみてどのような内容になるかを判断することによって、表層表現と内容の対を増やした。それらの対と、本研究で提案した係り受けの規則を適用した結果、要件抽出、要件の意味特定、要件と効果の係り先決定のすべてを正しく解析できたものは、学習コーパスに関しては181文中170文(93.3%)であった(表2, 3)。また、テストコーパスに関しては、275文中224文(81.4%)が正しく解析された(表4, 5)。

表2を見るとわかるように、機能表現と内容はほぼ一対一に対応し、誤りがほとんどなく要件の内容を特定できる。表3を見ると、係り受けの結果が良いことがわかるが、これは要件や主題、構成要素という大きな単位の間での関係のみを調べたこと、法律文の構文の特徴がある程度限られていることによると思われる。テストコーパスの結果であるが、表4を見ると、意味推定に関し学習コーパスよりも若干悪いものがあることがわかる。

これは学習コーパスでは見られなかった機能表現を持つ要件が存在することが原因であると思われる。そのために、要件と効果の弁別ができないものがあり、表5のように、効果内の要件の特定にも影響があった。その詳細は次の考察の章で述べる。

6 考察

6.1 学習コーパスでの誤り

学習コーパスでの誤りは、効果部の特定や、その内容の推定に関する誤りがほとんどである。

1. 要件を効果と誤推定(2例)：例えば、「その換価代金につき、」の「につき」は「～とみなし、」という意味の慣用句で、要件を表すが、間違って効果部の連用形としてとらえられた。
2. 効果を要件と誤推定(1例)：「この場合において、同条第一項第四号中「市町村」とあるのは、「事業

表 2: 学習コーパスでの精度の最終結果

意味分類	再現率		適合率	
状況規定	84 / 84	(100%)	84 / 84	(100%)
状況付加	1 / 1	(100%)	1 / 1	(100%)
別途制定	11 / 11	(100%)	11 / 11	(100%)
状況依存	1 / 1	(100%)	1 / 1	(100%)
手段規定	2 / 2	(100%)	2 / 2	(100%)
目的規定	7 / 7	(100%)	7 / 7	(100%)
適用限定	21 / 21	(100%)	21 / 21	(100%)
内容規定	5 / 5	(100%)	5 / 5	(100%)
時間規定	7 / 7	(100%)	7 / 7	(100%)
時間付加	1 / 1	(100%)	1 / 1	(100%)
場所規定	1 / 1	(100%)	1 / 1	(100%)
接続語	16 / 16	(100%)	16 / 16	(100%)
例外規定	7 / 7	(100%)	7 / 7	(100%)
適用緩和	2 / 2	(100%)	2 / 2	(100%)
適用制限	1 / 2	(50%)	1 / 1	(100%)
異種適用	2 / 3	(67%)	2 / 2	(100%)
原則規定	6 / 6	(100%)	6 / 6	(100%)
反則規定	1 / 1	(100%)	1 / 1	(100%)
効果部	220 / 221	(99%)	220 / 223	(99%)
主題	152 / 152	(100%)	152 / 152	(100%)
単純接続	51 / 53	(96%)	51 / 52	(98%)
主格	7 / 7	(100%)	7 / 7	(100%)
目的格	3 / 3	(100%)	3 / 3	(100%)
対象格	6 / 6	(100%)	6 / 6	(100%)
源泉格	4 / 4	(100%)	4 / 4	(100%)
場所格	2 / 2	(100%)	2 / 3	(67%)

表 3: 学習コーパスによる、係受けと主題・効果内の要件の特定の評価

評価内容	正解率	
係受け	618 / 629	(98%)
効果内の要件の特定	35 / 42	(83%)
主題の連体修飾の特定	83 / 84	(99%)

表 4: テストコーパスでの精度

意味分類	再現率		適合率	
状況規定	111 / 112	(99%)	111 / 113	(98%)
状況付加	1 / 2	(50%)	1 / 1	(100%)
別途制定	9 / 9	(100%)	9 / 9	(100%)
手段規定	7 / 11	(64%)	7 / 7	(100%)
目的規定	4 / 5	(80%)	4 / 4	(100%)
適用限定	20 / 27	(74%)	20 / 21	(95%)
内容規定	6 / 6	(100%)	6 / 11	(55%)
時間規定	21 / 29	(72%)	21 / 22	(95%)
場所規定	0 / 1	(0%)	0 / 0	—
接続語	19 / 21	(93%)	19 / 19	(100%)
例外規定	12 / 12	(100%)	12 / 12	(100%)
原因規定	0 / 0	—	0 / 1	(0%)
適用緩和	3 / 3	(100%)	3 / 3	(100%)
適用制限	0 / 6	(0%)	0 / 0	—
原則規定	6 / 8	(75%)	6 / 6	(100%)
反則規定	1 / 1	(100%)	1 / 1	(100%)
効果部	305 / 305	(100%)	305 / 317	(96%)
主題	221 / 223	(99%)	221 / 224	(99%)
遊接接続	0 / 1	(0%)	0 / 0	—
単純接続	102 / 107	(95%)	102 / 107	(95%)
主格	12 / 13	(90%)	12 / 13	(92%)
目的格	1 / 1	(100%)	1 / 1	(100%)
対象格	19 / 20	(95%)	19 / 24	(79%)
源泉格	1 / 1	(100%)	1 / 3	(33%)
場所格	1 / 1	(100%)	1 / 6	(17%)

表 5: テストコーパスによる、係受けと主題・効果内の要件の特定の評価

評価内容	正解率
係受け	907 / 942 (96%)
効果内の要件の特定	30 / 43 (70%)
主題の連体修飾の特定	124 / 132 (94%)

者」と読み替えるほか、同条の規定に関し必要な技術的読替えは、政令で定める。」という条文中、「～のほか、」は独立の効果部であるが、例外規定の要件に誤推定した。さらに、この文の中の2つの効果と「は、」を持つ構成要素による対は、対比構造ではなく独立の文を合わせたもので、それを推定するには意味的な推論を要する。これらのため「～とあるのは、」の係り先が誤り、さらに「この場合において、」の係り先が誤っている。

- 効果部の並列・従属の誤り(3例): 例えば、「教授は、学生を教授し、その研究を指導し、又は研究に従事する。」という条文中の「学生を教授し、」が従属節であって要件の混入と推定された。さらにこのために主題「教授は、」の係り先が誤っている。
- 効果部内の要件推定誤り: 例えば、「当該出納員をして～委任させる」の「して」の部分の連用形をとらえ、効果部内に入りこんだ要件と間違えて推定している。

また、条文の大局的な構造を誤ってとらえたものがある。「前項に規定する職員は、その事務所以外の場所において登録証明書の提示を求める場合には、その身分を示す証票を携帯し、請求があるときは、これを提示しなければならない。」という条文は、最初の「～求める場合には、」はそれ以降全体に係っている。その中で、「請求があるときは、」はさらに特別の場合として、「これを提示しなければならない。」に係っている。しかし、一見すると、「～場合は、～し、～場合は、する。」という対比構造に見える。その結果、最初の要件が両方の効果に係ると誤った解答をした。このような構造の違いを厳密にとらえるには、条文全体の意味から推論せざるを得ないが、現在のところそれは困難である。

表現の曖昧さのために要件・格要素の意味推定を誤ったものがある。典型的なものは「～者で、～ものは、」の「で」が「で格」と推定されるが、これは判定詞で

あり、「～であり、かつ～」の意味を持つ。解決策として、表層的解析にとどめず、KNPの格解析機能などをより積極的に利用して意味的に推定することが考えられる。

最後に、読点ごとに区切って構成要素の単位としたため、主題の範囲を誤ったものがある。「この場合主任者は、」全体が主題と推定されたが、これは誤りであり、このため主題の連体修飾の推定も誤った。

6.2 テストコーパスでの誤り

テストコーパスに適用した結果、要件を示す機能表現の登録が足りないため、効果の特定そのものの失敗することが多かった。このため、他の主題や要件の係り先も誤ることがあった。

- 効果部の特定の失敗: 原因の主なもの、意味的に次の要件に係る連用接続の用言をとらえてしまい、効果部としてしまったこと。「あらかじめ」などの機能表現を登録したり、「次に」など次の要件への接続の手がかりをとらえて解決できる場合もあるが、構文的には対応できない場合も多い。例えば、「～に伴い、」、「～提出して、」、「～し、」などを持つ節は、次の要件に係るか、効果部かを特定することは難しい。
- 効果部内の要件特定の失敗: 特に「～て」の扱いが難しい。「公務として」や「～を限度として～」、「～基準として」などは「～として」によって区別できるが、「その本来の目的を逸脱して」などは意味的に判断するよりない。
- 主題の修飾範囲の誤り: 条文列举や、括弧などの記号処理など、特殊な形によって、主題の修飾部全体を正しく推定できないものがある。法律文特有のものとして、個別に対応せざるをえない。
- 機能表現の曖昧さ: 「について」は内容規定と、適用限定の意味がある。また、「に対して」は適用限定、対象格、「対抗して」の意味がある。「において」は、状況規定と場所規定の曖昧さがある。
- 格の曖昧さ: 「が、で、から、まで、に」はそれぞれ自身意味的曖昧さがあるし、逆接の接続詞「が」、判定詞「で(だ)」などとの区別が難しい。
- 機能表現の登録不足: 「ときも、また」(状況付加)、「直ちに」(時間規定)、「当たっては」(状況規定)、「するには」(目的規定)などの新たな機

能表現の登録が必要であることがわかった。

- その他の曖昧さ：「～もので、」と「ので」の混同。
- 係り先の誤り：要件の係り先が、意味的な要因によって、提案手法から逸脱することがある。例えば、ある条文では、効果に係るとされる「当分の間、」が主題の述部に係ってしまった。このような例に対応するには、全体の意味的な強調部分を推定する、より複雑な方法が必要と思われる。
- 「～も、」の扱い：とりたて助詞「～も、」の意味を特定する必要がある。テストコーパスでは否定の場合の主格に限られていたが、その前につく助詞によってどのような意味もとらうる。

6.3 機能表現について

各要件の意味を推定するとき、連語による機能表現などの表層の手がかりはかなり有効である。その量は多いが、「で」、「が」、「に」、などの助詞や、連用形の末尾「し」、「て」などと意味的に区別するためには、ある程度の量を登録することは避けられないと思われる。

今回は対比構造という法律文の大局的構造と、主題や要件の意味を解析することに重きを置いたが、文末表現などを考慮することによって[10, 11]より精緻な解析ができるものと思われる。また、法律文での独特な名詞並列の扱われ方[12]などを積極的に利用していくことも、精度向上の一助となると考えられる。

7 おわりに

六法全書法律文の大局的構造の解析と要件の意味推定を表層の手がかりによって行なう手法を提案した。文の構成要素が対比構造をなしているかを検出し、それによって各主題や要件の係り先を特定する。そして各要件の機能表現によって要件のさす内容を特定する。同時に主題の連体修飾部や、効果部に入りこんだ要件の抽出を行なう。六法全書の条文で実験を行なった結果、学習コーパスに対して93%の正解率が、テストコーパスに対して81%の正解率が得られた。また、とりたて助詞「は」と読点の有無が対比構造の生成・認識の鍵となり、それによって係り先が決定されることを明らかにした。他の文末表現などによる標準構造への変換と、対比構造の一般性の検討を今後行なう予定である。

参考文献

- (1) 吉野一(編), 平成7年度科研費重点領域研究「法律エキスパート」研究成果報告書, (1996).
- (2) 原口誠, 法的推論, 情報処理学会学会誌, Vol. 37, No. 5, (1996), pp. 425-435.
- (3) 田中 規久雄, 川添 一郎, 成田 一, 法律条文の標準構造, 情報処理学会研究報告 NL-93-97, (1993), pp. 79-86.
- (4) 岩本 秀明, 野村 浩郷, 法律文の自然言語処理について, 情報処理学会研究報告 NL-91-83, (1991), pp. 7-14.
- (5) 松本裕治, 黒橋禎夫, 山地 治, 妙木 裕, 長尾 真, 日本語形態素解析システム JUMAN version 3.1 使用説明書(京都大学工学部 長尾研究室, 奈良先端科学技術大学院大学 松本研究室), (1996).
- (6) Kurohashi, S. and Nagao, M., A syntactic analysis method of long japanese sentences based on the detection of conjunctive structures., *Computational Linguistics*, Vol. 20, No. 4, (1994), pp. 507-534.
- (7) 川添 一郎, 牧 隆史, 田中 規久雄, 法律条文の標準構造(2), 情報処理学会研究報告 NL-95-107, (1995), pp. 97-104.
- (8) Nagai, H., Nakamura, T., Nomura, H., Skelton structure acquisition of japanese law sentences based on linguistic characteristics., *NLPRS'95*, (1995).
- (9) 平松 寛司, 永井 秀利, 中村 貞吾, 野村 浩郷, 要件効果構造に基づく法律文制限言語モデルと法律文解析, 情報処理学会研究報告 NL-96-115, (1996), pp. 21-27.
- (10) 長野 馨, 岩本 秀明, 永井 秀利, 野村 浩郷, 文末表現から見た法律文の制限言語モデルについて, 情報処理学会研究報告 NL-92-89, (1992), pp. 75-82.
- (11) 長野 馨, 永井 秀利, 中村 貞吾, 野村 浩郷, 動詞の機能に基づく法律文の制限言語モデル, 情報処理学会研究報告 NL-93-95, (1993), pp. 25-32.
- (12) 岩本 秀明, 長野 馨, 永井 秀利, 中村 貞吾, 野村 浩郷, 法律文における並列構造の特徴とそれに基づく制限言語モデルについて, 情報処理学会研究報告 NL-93-98, (1993), pp. 17-23.