

法令工学の言語処理 Natural Language Processing for Legal Engineering

島津 明

北陸先端科学技術大学院大学
情報科学研究科
安心電子社会教育研究センター

はじめに

- 背景

- 社会は, 税, 年金などの多数の社会制度により支えられている.
- 税や年金などの制度は法令に基づき情報システムとして実現されている.
- 法令遵守, 内部統制, 情報倫理, 技術標準などの問題

- 問題

- タイムリーな法令の制定・改定
- 安心安全な電子社会の実現
- 法令, 規則などの適切な扱い

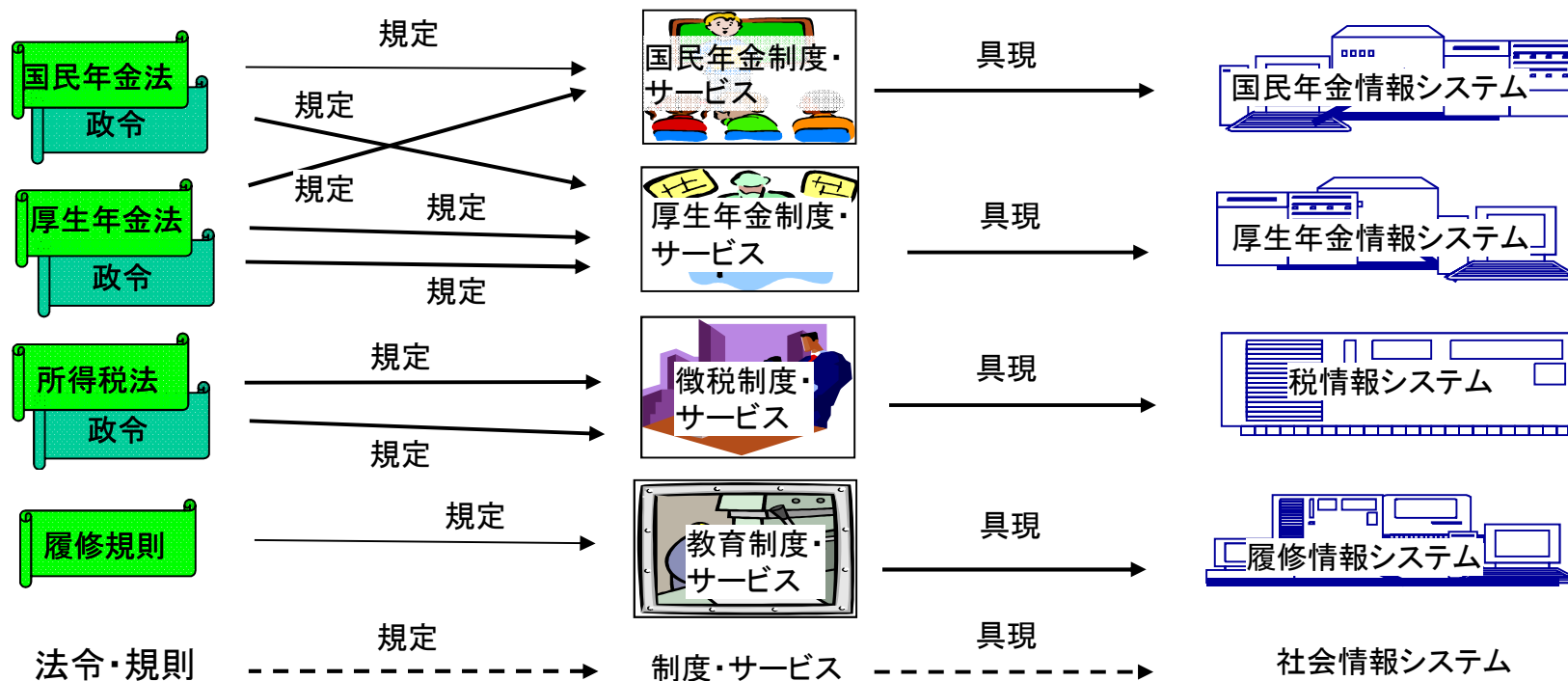
- 法令工学

- 「法令は社会の仕様書」, 「法令は社会を動かすソフトウェア」という立場から, 法令に情報科学的手法を適用し, 以下のことを明らかにする
 - 完全で矛盾のない法令の作成や変更を支援する方法論
 - 法令が規定する情報システムの設計開発を支援する方法論
 - 規則, 仕様なども対象とする方法論

- 法令工学の言語処理

- 言語処理は, 法令がテキストであることから, 法令工学の基盤となる.
- 法令構造を捉え, 法令の矛盾検出, 可読性向上などを支援

法令は社会を動かすソフトウェアである



- 社会には様々な制度があり、様々なサービスが行われている。
- 今後の電子社会では、多くの制度やサービスは社会情報システムにより具現される。
- 社会情報システムは、制度やサービスを規定する法令に基づいて実現される。

従来、法令を仕様とみるシステムの側面は研究対象でなかった。

法令はソフトウェア

第二十七条 老齢基礎年金の額は、七十八万九百円に改定率(次条第一項の規定により設定し、同条(第一項を除く。)から第二十七条の五までの規定により改定した率をいう。以下同じ。)を乗じて得た額(その額に五十円未満の端数が生じたときは、これを切り捨て、五十円以上百円未満の端数が生じたときは、これを百円に切り上げるものとする。)とする。ただし、保険料納付済期間の月数が四百八十に満たない者に支給する場合は、当該額に、次の各号に掲げる月数を合算した月数(四百八十を限度とする。)を四百八十で除して得た数を乗じて得た額とする。

- 一 保険料納付済期間の月数
- 二 保険料四分の一免除期間の月数(四百八十から保険料納付済期間の月数を控除して得た月数を限度とする。)の八分の七に相当する月数
- 三 保険料四分の一免除期間の月数から前号に規定する保険料四分の一免除期間の月数を控除して得た月数の八分の三に相当する月数
- 四 保険料半額免除期間の月数(四百八十から保険料納付済期間の月数及び保険料四分の一免除期間の月数を合算した月数を控除して得た月数を限度とする。)の四分の三に相当する月数
- 五 保険料半額免除期間の月数から前号に規定する保険料半額免除期間の月数を控除して得た月数の四分の一に相当する月数
- 六 保険料四分の三免除期間の月数(四百八十から保険料納付済期間の月数、保険料四分の一免除期間の月数及び保険料半額免除期間の月数を合算した月数を控除して得た月数を限度とする。)の八分の五に相当する月数
- 七 保険料四分の三免除期間の月数から前号に規定する保険料四分の三免除期間の月数を控除して得た月数の八分の一に相当する月数
- 八 保険料全額免除期間(第九十条の三第一項の規定により納付することを要しないものとされた保険料に係るものを除く。)の月数四百八十から保険料納付済期間の月数、保険料四分の一免除期間の月数、保険料半額免除期間の月数及び保険料四分の三免除期間の月数を合算した月数を控除して得た月数を限度とする。)の二分の一に相当する月数

$$\text{老齢基礎年金の額} = 780,900\text{円} \times \text{改定率} \times (\text{保険料納付済期間の月数} + \text{保険料4分の1免除期間の月数} \times 7/8 + \dots) / 480$$

(端数処理は省略)

フライング減税 来春のはずが...法律に書き忘れ 財務省

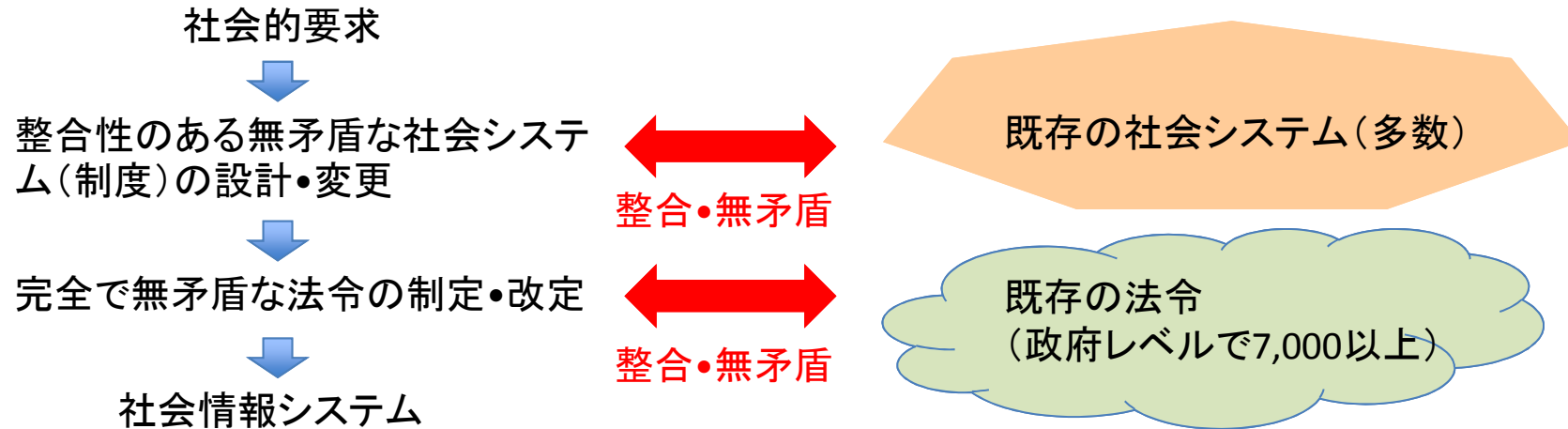
朝日新聞デジタルの記事(2013年5月31日)

財務省は30日、来年4月の消費増税に合わせて始めるはずだった住宅関連減税の一つを、誤って今年1月から始めていたと発表した。税制改正法で「来年4月以降」という規定を記し忘れたためだ。財務省は税制に関する法律のうっかりミスは「過去に例がない」としている。

ミスがあったのは、住宅のバリアフリー改修をすると所得税の一部が戻る減税。3月に成立した今年度税制改正法に含まれていた。予定では、減税対象の工事の限度額を来年4月に今の150万円から200万円に引き上げ、これに合わせて、所得税から差し引く減税額を今の最大15万円から最大20万円に拡大するはずだった。

だが、財務省が法案をつくる際、所得税の規定に「来年3月末までは150万円のまま」という記述を書き漏らした。法案は国会提出前に内閣法制局のチェックを受け、国会で審議されたが、だれも気づかずに3月末に成立した。規定がないと今年1月1日の入居分から減税が拡大されてしまう。財務省はそのまま減税拡大を適用することにした。ミスは最近になって財務省が関連する政令をつくらうとして気づいたという。麻生太郎財務相は「ツメが甘かったということだ」と謝罪し、近く関係者を処分する考えを明らかにした。(鯨岡仁)

法令工学



- ✓ 法令の作成・変更作業の困難さのために、必要な制度・サービスの実施が遅れ、社会がダイナミズムを失っている。
- ✓ 法務担当者は多大な労力を使っている。
- ✓ 社会情報システム構築・維持のために莫大なコストがかかっている。

- ✓ 科学的方法論やソフトウェア開発のような方法論の欠如
- ✓ 従来、以下は研究対象でなかった。
 - 整合性の高い法令の作成・変更などを系統的に行う方法論
 - 法令に基づく社会情報システムの設計方法論

➔ 法令工学

法令工学に関する経過

- 21世紀COEプログラム「検証進化可能電子社会」(2005～2009)
 - 代表: 片山卓也, 島津 明(2009)
 - 目的: 安心安全な電子社会システムを設計・構築する方法論を情報科学の形式的手法を用いて確立するとともに新学問分野を創出すること
 - 成果: 法令工学の提案, 形式検証技術, 世界最大ネットワークシミュレータのOS
- 法令工学の研究(2005～)
 - 法令文書の言語処理(島津研)
 - 法令の矛盾検出・除去法(東条研)
 - 法施行情報システムの設計開発法(落水研)
 - 法令対象ドメインの形式記述と検証(二木研)
- 法令の矛盾検出法の提案と富山県条例への適用(人工知能学会誌, 2009)
- 国民年金法の構造的書換え(JAIST Press, 2009)

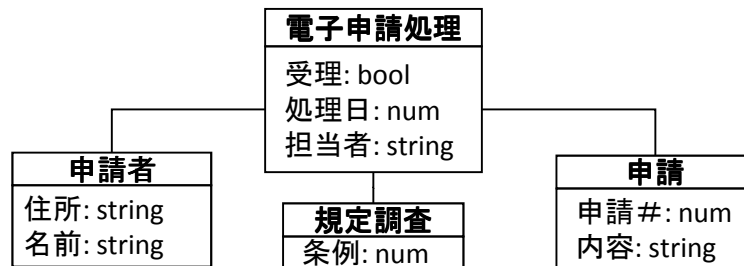
法令工学による方法の例

県の機関は、申請等のうち当該申請等に関する他の条例等の規定により書面等により行うこととしているものについては、当該条例等の規定に関わらず、規則で定めるところにより、電子情報処理組織を使用して行わせることができる。（富山県条例54号3条）

言語解析

電子申請行為(x,y,v,c) ←
申請行為(x,y,v,a) ∧ 申請者(x) ∧ 県の機関(y) ∧ 申請等(v) ∧ 書面等(a) ∧ 申請手段(c) ∧
規定(z,v) ∧ 関連条例等(z) ∧ 電子処理組織(c)

社会情報
システム的设计



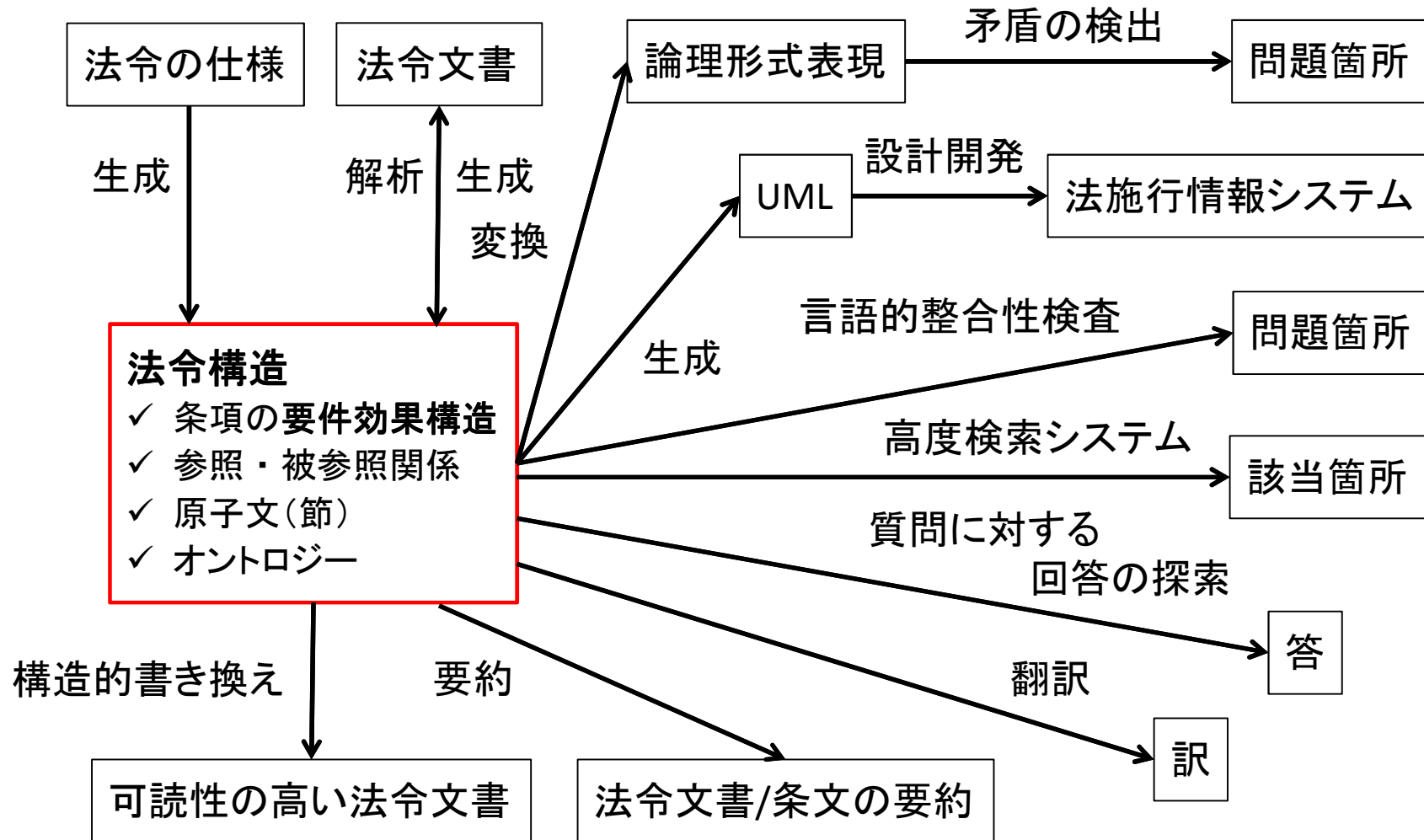
法令変更・改善, 関連法令
へ整合的変更伝播

形式
推論

矛盾・不完全
性の検出除去,
内容確認

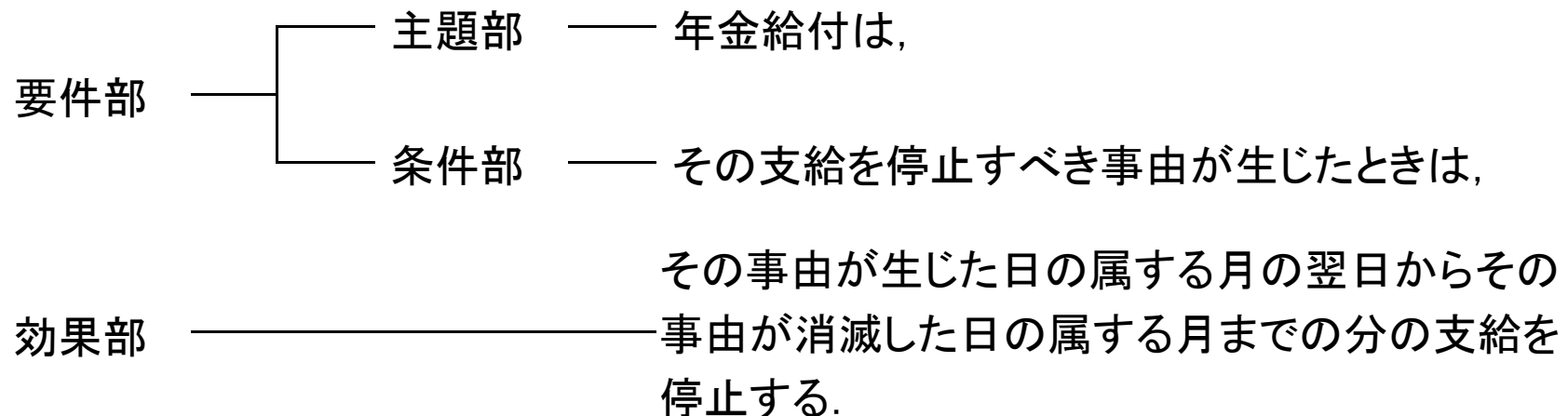
法令文に基づく説明機構,
法令変更時の情報システム
の変更管理

法令工学の言語処理



法令文の論理構造：要件効果構造

年金給付は、その支給を停止すべき事由が生じたときは、その事由が生じた日の属する月の翌日からその事由が消滅した日の属する月までの分の支給を停止する。
(国民年金法第18条第2項)



(年金給付の支給を停止すべき事由が生じた)

→ (その事由が生じた日の属する月の翌日から

その事由が消滅した日の属する月までの分の

{年金給付の}支給を停止する)

法令文の論理・形式表現の様式

- 1階述語論理表現＋様相記号，ダビドソニアン様式（イベント変数）
 - 事象参照表現，ゼロ代名詞，任意格，文末様相表現

- 例

毎月の保険料は，翌日末日までに納付しなければならない。

（国民年金法第91条）

$\forall x1, t1, \exists e1, x2, x3, x4, t2, t3, t4$

被保険者(x1) ∧ 年金(x2) ∧ 月(t1) ∧ 保険料(x3, x1, x2, t1)

⇒ □(納付(e1, x1, x3, x4) ∧ time(e1, t2) ∧

翌月(t3, t1) ∧ 末日(t4, t3) ∧ ≤(t2, t4))

- 具体的に法令を分析，記述
 - 富山県の電子化条例
 - 千代田区の環境条例
 - 国民年金法（第10章などは除く）
 - 所得税法の一部

条文の論理・形式表現への変換の試み

- 法令文の論理構造
 - 要件効果構造
- 論理・形式表現の様式
 - 1階述語論理表現＋様相記号, ダビドソニアン様式
- 論理・形式表現の対象
 - 条項の第1文
- 条項と論理構造の関係
 - 幾つかの型に分類
- 変換法
 - ルールベース (JUMAN, KNP, 文パターン, 語彙体系, 格解析, Jurisin, 2009)
- 実験対象(1文が対象)
 - 富山県条例, 広島県条例
 - 国民年金法

条項の言語表現

- 条文は、複数の項からなり、項は1～2文からなる(国民年金法)。
- 文には括弧による挿入文があり、語の定義、付加条件などが書かれる。

年金給付(この法律の他の規定又は他の法令の規定によりその全額につき支給を停止されている年金給付を除く。)は、その受給権者の申出により、その全額の支給を停止する。ただし、この法律の他の規定又は他の法令の規定によりその額の一部につき支給を停止されているときは、停止されていない部分の額の支給を停止する。

(国民年金法第20条の2)

- 様々な条件などを号で列挙する場合もある。

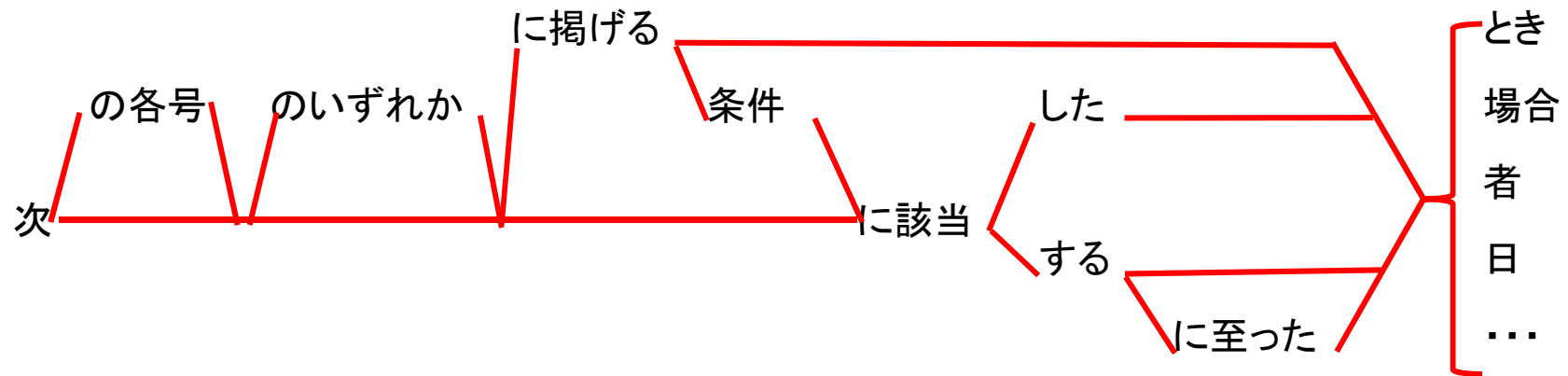
遺族基礎年金の受給権は、受給権者が次の各号のいずれかに該当するに至ったときは、消滅する。

- 一 死亡したとき。
- 二 婚姻をしたとき。
- 三 養子となつたとき(直系血族又は直系姻族の養子となつたときを除く。)

(国民年金法第40条)

号参照表現

- 条項の号参照の表現形式は一定のパターンがあり、参照文と参照される号とを合わせて一つの文にすることができる。(論理・形式表現に変換する観点からは、論理・形式表現を組み立てるときに組み込んでよい。)



次に該当する者が ... とき, ... する.
一 □□する者. (条件1)
二 ○○する者. (条件2)



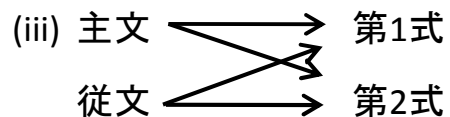
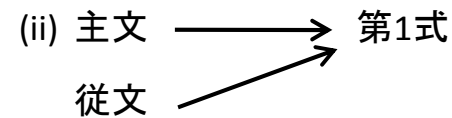
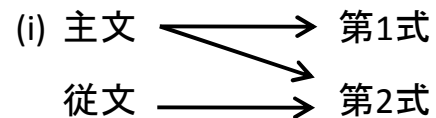
□□する者が ... とき, ... する.
○○する者が ... とき, ... する.

条項と論理構造の関係

- 次の条項は、2文からなり、2つの要件効果構造を表すが、それらは第1文、第2文と1対1対応ではない。実際、以下の項に示すように様々な対応がある。

遺族基礎年金、寡婦年金又は死亡一時金は、被保険者又は被保険者であつた者を故意に死亡させた者には、支給しない。被保険者又は被保険者であつた者の死亡前に、その者の死亡によつて遺族基礎年金又は死亡一時金の受給権者となるべき者を故意に死亡させた者についても、同様とする。

- 第1文、第2文、あるいは被挿入文、挿入文をそれぞれ主文、従文とし、第1の要件効果構造を第1式、第2の要件効果構造を第2式として、一般に、次の対応が見られる。



- 条項と論理構造の関係は、大きく7種類、細かくは20種類程度に分類できる。(Jurix, 2010)
- 国民年金法2007版, 条項数, 556, 複数文条項, 200, 分析対象, 189

複数文条項と論理構造(要件効果構造)

元文(国民年金法第20条の2)

年金給付(この法律の他の規定又は他の法令の規定によりその全額につき支給を停止されている年金給付を除く。)は、その受給権者の申出により、その全額の支給を停止する。ただし、この法律の他の規定又は他の法令の規定によりその額の一部につき支給を停止されているときは、停止されていない部分の額の支給を停止する。

各文

| | |
|-----|---|
| 第1文 | [年金給付は、] T_1 <その受給権者の申出により、 A_1 {その全額の支給を停止する。} C_1 |
| 第2文 | ただし、<この法律の他の規定又は他の法令の規定によりその額の一部につき支給を停止されているときは、 A_2 {停止されていない部分の額の支給を停止する。} C_2 |
| 挿入文 | <この法律の他の規定又は他の法令の規定によりその全額につき支給を停止されている年金給付 A_3 を除く。 |

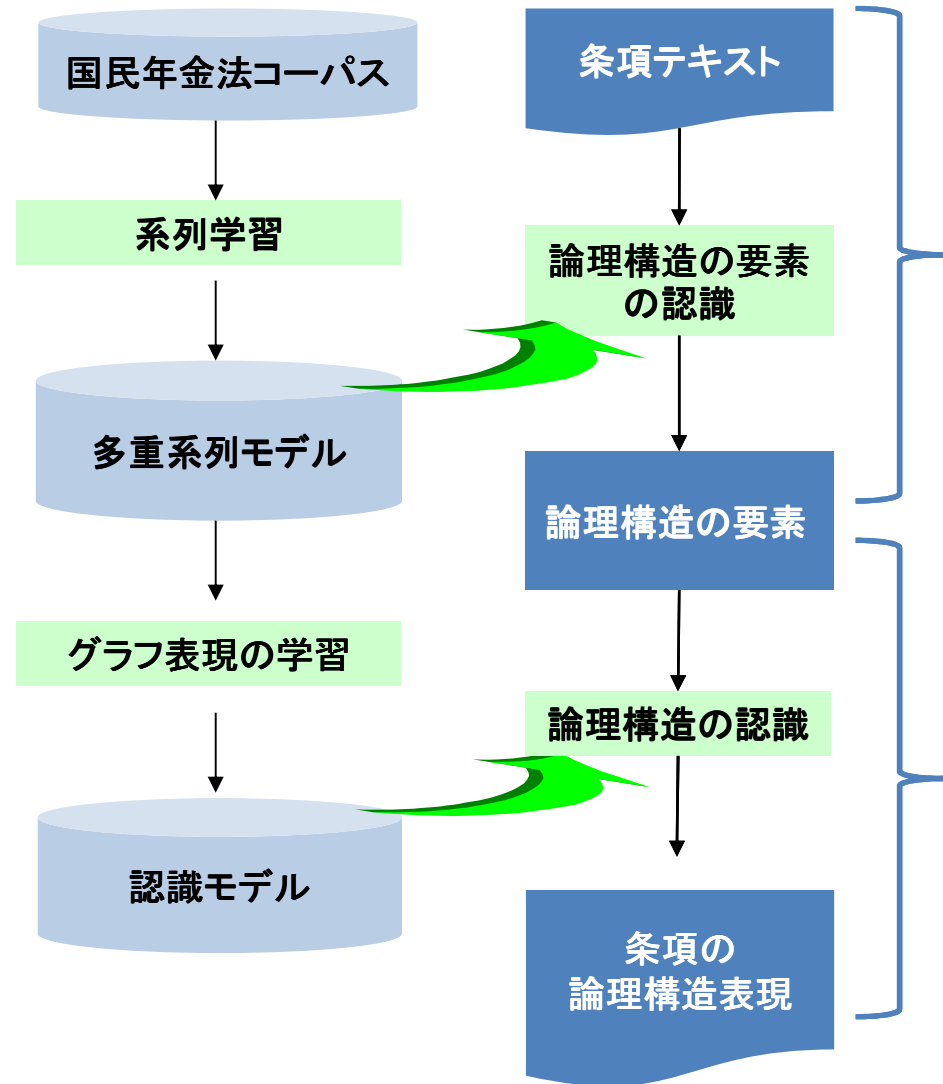
論理構造: 第1文と第2文:(iii)型, 第1文と挿入文:(ii)型, 第2文と挿入文:(ii)型

| | |
|-----|--|
| 第1式 | $\neg(A_3) \wedge (T_1, A_1) \wedge \neg(T_1, A_2) \rightarrow (T_1, C_1)$ |
| 第2式 | $\neg(A_3) \wedge (T_1, A_1) \wedge (T_1, A_2) \rightarrow (T_1, C_2)$ |

注. (A)はAを内容とする命題を表す.

複数文条項の要件効果構造の解析

- 2段階の解析法を提案(ACM TALIP, 2013)
 - 1) 論理構造の要素の学習と認識
 - ✓ 各文の文字列表現
 - ✓ 多重系列学習と要素の認識(CRFs)
 - 2) 論理構造の学習と認識
 - ✓ 論理構造要素のグラフ表現
 - ✓ 重み付きグラフの学習(MEM, SVM)
 - ✓ ILPによる適切な部分グラフの探索
- 実験
 - ✓ 論理構造要素の認識, F: 74.37%
 - ✓ 論理構造の認識, F: 83.38%
 - ✓ 全体の認識, F: 62.00%



共参照関係

- 共参照関係には、法令、条、項、号など、全テキストを指す場合、テキストの特定部分を参照する場合がある。従来研究の対象は前者であった。本研究は、後者も対象。

6 前項の届出は、厚生労働省令で定める場合を除き、厚生年金保険法の被保険者である第二号被保険者の被扶養配偶者である第三号被保険者にあつては、その配偶者である第二号被保険者を使用する事業主を経由して行うものとし、国家公務員共済組合法若しくは地方公務員等共済組合法の組合員又は私立学校教職員共済法の規定による私立学校教職員共済制度の加入者（以下「私学教職員共済制度の加入者」という。）である第二号被保険者の被扶養配偶者である第三号被保険者にあつては、その配偶者である第二号被保険者を組合員又は加入者とする国家公務員共済組合、地方公務員共済組合又は日本私立学校振興・共済事業団を経由して行うものとする。

7 前項に規定する第二号被保険者を使用する事業主とは、厚生年金保険法の被保険者である第二号被保険者を使用する事業所（同法第六条第一項に規定する事業所をいう。）の事業主（同法第二十七条に規定する事業主をいう。第百八条第三項において同じ。）をいう。

- 4段階で指示箇所を同定する方法を提案 (ICAIL, 2013)
 - (1) 系列の機械学習(CFRs)により指示表現を同定。(2) 指示表現の解析(指示先, 内容語, タイプ), (3) 構文木により被指示対象の候補を抽出, (4) 機械学習により被指示対象を決定
 - 国民年金法を対象に実験. Listwise 法が pairwise 法より少しよい. (accuracy: 86.03)

法令文は長くて括弧が多い

配当所得とは、法人（法人税法第二条第六号（定義）に規定する公益法人等及び人格のない社団等を除く。）から受ける剰余金の配当（株式又は出資（公募公社債等運用投資信託以外の公社債等運用投資信託の受益権及び社債的受益権を含む。次条において同じ。）に係るものに限るものとし、資本剰余金の額の減少に伴うもの及び分割型分割（同法第二条第十二号の九に規定する分割型分割をいい、法人課税信託に係る信託の分割を含む。以下この項及び次条において同じ。）によるものを除く。）、利益の配当（資産の流動化に関する法律第百十五条第一項（中間配当）に規定する金銭の分配を含むものとし、分割型分割によるものを除く。）、剰余金の分配（出資に係るものに限る。）、基金利息（保険業法第五十五条第一項（基金利息の支払等の制限）に規定する基金利息をいう。）並びに投資信託（公社債投資信託及び公募公社債等運用投資信託を除く。）及び特定受益証券発行信託の収益の分配（法人税法第二条第十二号の十五に規定する適格現物分配に係るものを除く。以下この条において「配当等」という。）に係る所得をいう。

（所得税法第24条第1項）

国民年金法の構造的書き換え

- 可読性の高い法令記述言語を目指して
- 構造的書き換え
 - 条文の要件効果の論旨を明確にするという方針
 - 要件効果構造の表現に係る規則(10)と文表現に係る規則(7)
- 2008年の国民年金法を対象 (JAIST Press, 2009)
 - 被保険者に関する条文が主な対象
 - 書き換えた条文の項や号は328
(国民年金法: 条文は全部で216, 項は全部で548)
- 書き換え結果は社会保険労務士が確認
- 被験者による実験
 - 8条項について, 30人の被験者を元文対象と構造的書き換え対象の2群に分け, 各条項の内容について質問し, 解答時間と回数を計測
 - 時間も回数も構造的書き換えが一般に少なかった.

国民年金法の構造的書き換え

〈元文〉

年金給付の受給権者が死亡した場合において、その死亡した者に支給すべき年金給付でまだその者に支給しなかつたものがあるときは、その者の配偶者、子、父母、孫、祖父母又は兄弟姉妹であつて、その者の死亡の当時その者と生計を同じくしていたものは、自己の名で、その未支給の年金の支給を請求することができる。

(国民年金法第19条)

〈構造的書き換え〉

(AかつB)の場合、(CかつD)の者は、自己の名で、その未支給の年金を請求することができる。

- A: 年金給付の受給権者が死亡した。
 - B: 死亡した受給権者に支給すべき年金給付で、まだ支給しなかつたものがある
 - C: 受給権者の(配偶者又は子又は父又は母又は孫又は祖父又は祖母又は兄弟姉妹)
 - D: 受給権者の死亡の当時、受給権者と生計を同じくしていた者
-

おわりに

- 法令文書の論理構造に関する研究に加え、下記の研究を行ってきた。
 - ✓ 注釈コーパス作成
 - ✓ 構造的書き換えの自動化
 - ✓ 可読性向上のためのインタフェース
 - ✓ 法令QAシステム
 - ✓ 法令文書などの作成・変更を支援する情報更新法
 - ✓ 判例文書の構造解析(判例DB, 検索の支援)
 - ✓ 号の組込み生成法
 - ✓ 条項の論理構造に基づく翻訳法
 - ✓ 言語処理による矛盾検出法
 - ✓ 条項内容のUMLによる記述
- これらの研究を一般化・総合化するとともに、有効性を示すためにスケーラビリティを上げなければならないと考えている。

参考文献

- L.T.N. Tho, M.L. Nguyen and A. Shimazu. Unsupervised Keyword Extraction for Japanese Legal Documents, Legal Knowledge and Information Systems (JURIX 2013), Bologna, 2013.
- O. T. Tran, B. X. Ngo, M. L. Nguyen and A. Shimazu. Answering Legal Questions by Mining References Information. International Workshop on Juris-informatics (JURISIN 2013), Yokohama, 2013.
- O. T. Tran, M. L. Nguyen and A. Shimazu. Reference Resolution in Legal Texts. International Conference on Artificial Intelligence and Law (ICAIL), pp.101-110, Rome, 2013. (Award for Best Student Paper)
- B. X. Ngo, M. L. Nguyen, O. T. Tran, and A. Shimazu. A Two-Phase Framework for Learning Logical Structures of Paragraphs in Legal Articles. ACM TALIP, Vol.12, No.1, pp.3:1-3:32, doi>10.1145/2425327.2425330, 2013.
- H. Igari and A. Shimazu. Syntactic Document Structure of Legal Documents. International Workshop on Juris-informatics (JURISIN 2012), pp.49-62, Miyazaki, 2012.
- H. Bui, M. L. Nguyen and A. Shimazu. Divide and Translate Legal Text Sentence by Using its Logical Structure. International Conference on Knowledge, Information and Creativity Support Systems (KICSS2012), pp.18-23, Melbourne, 2012.
- M. L. Nguyen and A. Shimazu. A Semi Supervised Learning Model for Mapping Sentences to Logical form with Ambiguous Supervision. International conference on Applications of Natural Language Processing to Information Systems (NLDB 2012), pp. 116-127, Groningen, 2012.
- 島津 明. 法令工学: 安心な社会システム設計のための方法論—法令文書の解析を中心に— (解説論文). IEICE, Fundamental Review, Vol.5, No.4, pp.320-328, 2012.

参考文献

- M. Q. N. Pham, M. L. Nguyen, B. X. Ngo and A. Shimazu. A Learning to Rank Method for Information Updating Task. *Applied Intelligence*, DOI={10.1007/s10489-012-0343-2}, pp. 1-12, 2012.
- K. Takano, M. Nakamura, Y. Ohyama and A. Shimazu. Semantic Analysis of Paragraphs Consisting of Multiple Sentences - Towards Development of a Logical Formulation System -. *Legal Knowledge and Information Systems (JURIX2010)*, pp.117-126, 2010.
- 片山卓也. 法令工学: 安心な社会システム設計のための総合ソフトウェア科学 (招待論文). *情報処理*, Vol.51 No.5, pp.487-490, 2010.
- M. Nakamura and A. Shimazu. Semantic Analysis of National Pension Law, *International Workshop on Juris-informatics (JURISIN 2009)*, pp.83-94, 2009.
- 萩原信吾, 東条 敏. 対立を用いた法的知識の整合性検証. *人工知能学会誌*, 24, 5, p.405-416, 2009.
- 島津 明(編). 21世紀COEプログラム「検証進化可能電子社会」最終成果報告書, 北陸先端科学技術大学院大学, 2009.
- Y. Kimura, M. Nakamura and A. Shimazu. Treatment of Legal Sentences Including Itemized and Referential Expressions -Towards Translation into Logical Forms-, *New Frontiers in Artificial Intelligence: LNAI*, Vol.5447, pp.242-253, 2009.
- 島津 明. 国民年金法の構造的書き換え--法令工学の立場から--. *JAIST Press*, ISBN978-4-903092-19-5, 2009.
- 片山卓也(編). 法令工学の提案. *JAIST Press*, ISBN4- 903092-04-6, 2007.