JAIST Repository
https://dspace.jaist.ac.jp/

Title	法令文書の自然言語処理
Author(s)	島津，明
Citation	
Issue Date	2007-03-07
Type	Presentation
Text version	publisher
URL	http://hdl.handle.net/10119/8296
Rights	
Description	4th VERITE : JAIST/TRUST-AIST/CVS joint workshop on VERIfication TEchnologyでの発表資料，開催：2007年3月6日～3月7日，開催場所：北陸先端科学技術大学院大学・知識講義棟２階中講義室

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

法令文書の自然言語処理

2007. 3. 7

島津　明

目次

- 法令工学と自然言語処理
- 法令文の分析
- 法令文の解析

法令工学

- 法令が，各法令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れていることを検査・検証し，法律の改定に対しては，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（片山）
- 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（片山）
- 法令が対象とする世界のモデル化，対象とする概念や概念関係，それらを表現する語彙の選択，それらに基づく文章化等の支援

法令工学の視点からの自然言語処理

- 法律条文を論理式で形式的に表現して，論理的矛盾や整合性の検査・検証を行うために，法律条文を自然言語処理により形式表現に変換する．
- 自然言語処理により論理的矛盾や整合性の検査・検証を行う．
- オントロジー，辞書，文章生成，支援インタフェース等を作る．

これまでの取り組み

- 論理表現の検討
- 法令文の分析
- 法令文の論理表現への変換処理システムの開発
- 法令文の見易い表示の検討
- 言語処理による法令文書の整合性検査の検討

論理表現の検討

- 法令文を要件・効果構造（田中93）として解釈
  - 「～の場合は（要件），．．．である（効果）．」
- 1階述語論理表現と様相記号で記述
  - Davidsonian style
  - イベント変数（事象の参照に必要）
  - 格要素を述語で表現（arityを考慮しないですむ）

論理表現の例

公共の場所等において，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．（千代田区条例53号）

∀x1 x2 x3 x4 ∃e1 e2 e3 e4 e5
人(x1) ∧ チラシ等(x2) ∧ 公共の場所等(x3) ∧人(x4) ∧
((配布する(e1) ∧ agt(e1, x1) ∧ obj(e1, x2) ∧ loc(e1, x3)) ∨
(配布させる(e2) ∧ agt(e2, x1) ∧ obj(e2, x2) ∧ rcp(e2, x4) ∧ loc(e2, x3)))
∧ 散乱する(e3) ∧ obj(e3, x2) ⇒
O(回収する(e4) ∧ agt(e4, x1) ∧ obj(e4, x2) ∧
速やかに(e4) ∧ 清掃する(e5) ∧ agt(e5, x1) ∧ obj(e4, x3))

法令文の分析－論理表現による記述（1）

- 目的
  - 処理内容の理解
  - 正解データの作成
  - 機械学習のための訓練データ

法令文の分析－論理式による記述（2）

- 2005（昨年度）
  - 富山県条例第54号
    - 富山県行政手続等における情報通信の技術の利用に関する条例
    - 全10条，34文
  - 千代田区条例第53号
    - 安全で快適な千代田区の生活環境の整備に関する条例
    - 全28条，86文
  - 所得税法
    - 全244条，その内の86文

法令文の分析－論理式による記述（3）

- 2006（今年度）
  - 所得税法
    - 前年の86文の内，43文を修正
  - 国民年金法
    - 全148条，その内の45文＋α

法令文の分析－論理表現の生成規則

- 目的
  - 言語表現の形と論理表現の対応の明確化
- これまで
  - 要件効果構造に対応した文レベルのパタン規則の抽出
  - 言語表現の形と論理構造とを対応付ける規則の検討
  - どんな言語表現に対して，どんな述語や項を設定するか
  - 述語の変数のタイプや個数の検討
  - 様相を述語論理で置き換える検討

法例文の分析－言語表現と論理構造

第7条 次の各号のいずれかに該当する者は、国民年金の被保険者とする。
1. 日本国内に住所を有する20歳以上60歳未満の者であつて次号及び第3号のいずれにも該当しないもの . . . . .

第7条は，
　1号に該当する者 ⇒ 国民年金の被保険者
　2号に該当する者 ⇒ 国民年金の被保険者
　3号に該当する者 ⇒ 国民年金の被保険者
と解釈されるが，
　国民年金の被保険者(x) ⇒ 1号(x) ∨ 2号(x) ∨ 3号(x)
でもあろう．

第7条 次の各号のいずれかに該当する者は、国民年金の被保険者とする。
1. 日本国内に住所を有する20歳以上60歳未満の者であつて次号及び第3号のいずれにも該当しないもの（被用者年金各法に基づく老齢又は退職を支給事由とする年金たる給付その他の老齢又は退職を支給事由とする給付であつて政令で定めるもの（以下「被用者年金各法に基づく老齢給付等」という。）を受けることができる者を除く。以下「第1号被保険者」という。）
2. 被用者年金各法の被保険者、組合員又は加入者（以下「第2号被保険者」という。）
3. 第2号被保険者の配偶者であつて主として第2号被保険者の収入により生計を維持するもの（第2号被保険者である者を除く。以下「被扶養配偶者」という。）のうち20歳以上60歳未満のもの（以下「第3号被保険者」という。）

法例文の分析－言語表現と論理構造

第5条 この法律において、「被用者年金各法」とは、次の各号に掲げる法律をいう。
1. 厚生年金保険法（昭和29年法律第115号）
2. 国家公務員共済組合法（昭和33年法律第128号）
3. 地方公務員等共済組合法（昭和37年法律第152号）（第11章を除く。）
4. 私立学校教職員共済法

x＝被用者年金各法 ≡
x＝厚生年金保険法 ∨ x＝国家公務員共済組合法 ∨ . . .
厚生年金保険法 ⇒ 被用者年金各法

法例文の分析－言語表現と述語

「被保険者」
被保険者(x)
「被保険者の資格」 ？
資格(被保険者(x)) ？
「被保険者の資格を取得する」
取得する(e), 者(x), agt(e, x), obj(e, 被保険者の資格)
含意:
取得する(e), 者(x), agt(e, x), obj(e, 被保険者の資格)
⇒ 被保険者(x)

法例文の分析－論理

第16条 給付を受ける権利は、その権利を有する者（以下「受給権者」という。）の請求に基いて、社会保険庁長官が裁定する。

循環？

法令文の分析－言い替え

- 処理内容の理解
- 論理表現の分析のための理解
- 法令文書の見易い表示のための基礎データ

法令文の分析－言い替え（1）

- 2006（今年度）
  - 国民年金法
    - 被保険者，受給権者に関連する条文，約70
    - 社会保険労務士によるチェック（予定）

法令文の分析－言い替え（2）

第8条 前条の規定による被保険者は、同条第１項第2号及び第3号のいずれにも該当しない者については第1号から第3号までのいずれかに該当するに至つた日に、20歳未満の者又は60歳以上の者については第4号に該当するに至つた日に、その他の者については同号又は第5号のいずれかに該当するに至つた日に、それぞれ被保 険者の資格を取得する。

1. 20歳に達したとき。
2. 日本国内に住所を有するに至つたとき。
3. 被用者年金各法に基づく老齢給付等を受けることができる者でなくなつ たとき。
4. 被用者年金各法の被保険者、組合員又は加入者の資格を取得したとき。
5. 被扶養配偶者となつたとき。

法令文の分析－言い替え（3）

1. 第一号被保険者は次のA、B、Cのいずれかの時期に至ったときに被保険者の資格を取得する。
2. 二十歳未満の者又は六十歳以上の者はDの時期に至ったときに被保険者の資格を取得する。
3. 第二号被保険者はDの時期に至っときに被保険者の資格を取得する。
4. 第三号被保険者はAあるいはEの時期に至ったときに被保険者の資格を取得する。

A：二十歳に達したとき
B：日本国内に住所を得たとき
C：被用者年金各法に基づく老齢給付等の受給権者でなくなったとき<受給権者の定義は第16条>
D：被用者年金各法の被保険者、又は組合員、又は加入者の資格を取得したとき
E：被扶養配偶者となったとき

条文の論理式への変換

- 手続き方法による変換（信岡，中村）
- 機械学習による変換（Nguyen Le Minh)

条文の論理式への変換（手続き的方法）

1. 形態素解析・構文解析
2. 要件効果構造の解析
3. 言い替え処理
4. 格構造と名詞句の解析
5. 論理式の合成

要件効果構造の解析

公共の場所等において，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．（千代田区条例53号）

要件：

公共の場所等において，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，

効果：

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．

言い替え処理（多様性の吸収）

．．．事業所得の金額又は山林所得の金額の計算上，必要経費に参入する．（所得税法，第六十三条1項）

．．．事業所得の金額又は山林所得の金額を計算するとき，必要経費に参入する．

これまでに，530文を分析し，9パターンを求めている．

格構造と名詞句の解析（再帰処理）

第十一条　．．．．被保険者の資格を取得した日の属する月からその資格を喪失した日の属する月の前月までをこれに参入する.

（VP（NP （NP （VP （NP （VP （NP 被保険者の資格を）
取得した）
日の）
属する）
月から）
（NP （NP （VP （NP （VP （NP その資格を）
喪失した）
日の）
属する）
月の）
前月までを））
（NP これに）
参入する）

手続き方法による変換システム

- 作り直した（PerlからACL lisp）.
- 要件効果構造のパターン規則を増やした.
- 言い替え規則を加えた.
- 格解析と名詞句意味解析を再帰的にした.
- 格解析を一般化した（可能な解釈すべて行う）.
- 格フレーム辞書を増やした.
- 名詞句の意味解析の範囲を広げた.

条文の論理式への変換（機械学習の適用）

- 方法の検討
- 実験
- 日本語法令文と論理式対のデータ作成

今後

- 変換システムの精度とカバリッジの向上
- 論理表現の調整
  - 推論部と表現を合わせる
- 言語レベルでの法令文書の検査法の検討
  - 文書における定義，概念分類，語句の関係等の解析