

E-017

連結定数で結合された素式群による複文の述語知識表現法とそれへの変換法

Transformation of Complex Sentence to Predicate Logic Formula
Using Multiple Literals with Same Constants佐々木 智彦[∧]

Tomohiko Sasaki

石川 勉[∧]

Tsutomu Ishikawa

1 はじめに

我々は、“言葉で考えるコンピュータ”の実現を目指して研究している。このコンピュータ [1] では、不完全な知識を類似知識で補い近似解を導く概略推論法 [2] と、常識知識ベースが核となり、知識表現には拡張した述語論理形式を採用している [3]。本稿では、この知識表現法を前提とした連結定数を用いて素式間を結合した複文の表現法と、自然言語文からのこの形式への変換法について提案する。

2 知識の表現形式

すべての自然言語文を単純な構造の素式 (以下、述語式) および事象規定式や特性規定式、あるいはそれらを連言で結合した組み合わせにより表現する。詳細は文献 [3] を参照されたい。

(1) 述語式 述語式は以下のように表現する。知識の意味を一意に特定可能とするため、引数はラベル付きとする。

$$sP(r_1:t_1, r_2:t_2, \dots, r_n:t_n)$$

ここで、 s :様相子、 P :述語、 r_i :ラベル、 t_i :引数である。様相子は述部が否定や時制を伴う場合の意味的な差異を扱えるようにするため付加する。この知識表現は言葉をベースとしているため、述語と引数には基本的に単語または名詞句を利用する

(2) 事象規定式 述語式は一般に述部が引数部の性質を表すための形式であるため、述語式やその連言のみでは表現できない文がある。例えば、「～であることを知っている」等の文である。このような文に対しては以下のように表現する。これを事象規定式と呼ぶ。

[連結定数=述語式あるいはそれらの連言]

ここで、“=”の右辺の述語式はその真偽ではなく、述語式が表現している内容を表す。連結定数については後述する。

(3) 特性規定式 二者が等値のような関係である場合、述語式の述部にその関係を示す記号“EQ”と属性のラベル“atr”と値のラベル“val”を用いて、

$$EQ(atr:t_1, val:t_2)$$

と表現する。例えば、「太郎の年齢は20歳である」のような文は“EQ(atr:太郎の年齢, val:20歳)”と表現する。

3 知識変換法

3.1 単文の変換法

自然言語文を形態素・構文解析 (それぞれ茶筌 [4], 南瓜 [5] を利用) し、その結果より述部の品詞を調べ、様相子を獲得し引数部にラベル付けをする。単文における基本的な述語式の表現形式と変換法を述部の品詞毎に以下に示す。ここで、 P は述語 (動詞、形容詞または名詞)、 A, B は名詞または名詞句である。

(1) 述部が動詞の場合 「AはBをPする」などの文型は、“ $P(agt:A, obj:B)$ ”のように表現する。すなわち、述部には主動詞を、引数にはその深層格を用いる。この深層格に対するラベル付けについては、必須格に関してはEDR電子化辞書の日本語動詞共起パターン副辞書 [6] (以下、共起辞書) を利用する。取り出した文節を共起辞書中の記述と照合する。この照合では、対象の文節と共起辞書に書かれている格助詞が一致し、かつ対象文節の名詞が共起辞書中の概念識別子群に属していればラベルを付与する。一方、任意格のラベル付けについては、格助詞と意味カテゴリーを利用してラベルを付与する。

(2) 述部が形容詞の場合 「AはP」の場合、基本的にPがAを形容するとし、一項述語の場合、ラベル“sbj”を用いて、

$$P(sbj:A)$$

と変換する。また、二項述語の場合、 $s-obj, s-gol$ などのラベル (表1) を格助詞より決定し付与する。ただし、例外として「AはBがP」の文型の場合は、格助詞のみでラベルを付与せず、Pが感情・感覚形容詞または形容動詞 (好き、苦しいなど) であれば、

$$P(sbj:A, s-obj:B)$$

とする。また、Pが属性形容詞 (美しいなど) であれば、

$$P(sbj:A \text{ の } B)$$

と表現する。

表1 述語が形容詞の場合に用いるラベル

ラベル	深層格	概要	主な格助詞
sbj	主体 (subject)	形容の主体	は
s-obj	対象 (object)	形容の対象	を
s-gol	目標 (goal)	形容の終点	に, まで
s-sou	源泉 (source)	形容の始点	から
s-cmp	比較 (compare)	主体との比較	より

(3) 述部が名詞の場合 「AはPである」の文型は、基本的にAがPの性質を表すとし、一例を示すラベル“inst”を用いて、

$$P(inst:A)$$

と変換する。ただし、AやPの意味カテゴリーをシソーラスで調べ、一定の上下関係がないまたはPが数値や単

[∧]拓殖大学 工学部 情報工学科

位に関する語である場合は、特性規定式を適用する。例えば、「犬は動物である」の場合、“動物 (inst:犬)”と表現する。また、「太郎の給料は20万円である」などAがPの性質を表さない場合、特性規定式を用いて、“EQ(atr:太郎の給料,val:20万円)”と表現する。

なお、ここでは基本的には二項述語は利用しない。例えば、「花子のおじは太郎である」の場合、“おじ(太郎,花子)”という表現も考えられる。しかし、どの語彙が二項述語となるかを正確に判定するのが困難なため、ここでは“花子のおじ (inst:太郎)”のように述部を名詞句とし、一方をその修飾部に置くこととする。

3.2 複文の変換法

複文は原則として、同一の連結定数(S)により意味的に結合された複数の述語式の連言で表現する。ただし、述語式の連言のみで表現できない場合は、事象規定式や含意を用いて表現する。

3.2.1 変換の考え方

(1) 関係代名詞や関係副詞で結合している複文 英語における関係代名詞や関係副詞で結合している複文は、特定の先行詞がある場合にはその先行詞を示す述語式を含ませる。また、文中に“～とき”、“～ところ”などのように先行詞がない場合には“とき (inst:S)”、“ところ (inst:S)”といった意味の無い述語式は含ませない。これらを省略してもラベルのみで解釈可能である。以下に表現例を示す。

- 2 父は母が帰宅する夕方に会社へ行く
⇒ 帰宅する (agt:母,tme:S1)
 ∧ 行く (agt:父,gol:会社,tme:S1) ∧ 夕方 (inst:S1)
- 2 父は母が帰宅する時に会社へ行く
⇒ 帰宅する (agt:母,tme:S1)
 ∧ 行く (agt:父,gol:会社,tme:S1)

(2) thatで結ばれた複文 英語のthat(同格を表す接続詞的用法)で結ばれた複文の場合、事象規定式を用いる。“～こと”など先行詞がない場合は、(1)と同様にこの述語式は含ませない。以下に変換例を示す。

- 2 私は太郎が犯人であることを知っている
⇒ 知る (agt:私,obj:S) ∧ [S=犯人 (inst:太郎)]
- 2 私は太郎が犯人であるという事実を知っている
⇒ 知る (agt:私,obj:S) ∧ 事実 (inst:S)
 ∧ [S=犯人 (inst:太郎)]

(3) “～ため”など単純な接続詞で結ばれた複文 “～ため”、“～ので”等の接続詞で結合された文は、これら接続詞を伴う従属節の内容を連結定数で表し、主節を表す述語式の適切な個所(ラベルに対応させる)にこの連結定数を含ませ表現する。以下に表現例を示す。なお、接続詞“ので”の場合はラベル“cnd”を用いて表現する。

- 2 太郎は、将来、会社を経営するため、経済学を専攻する。
⇒ 専攻する (agt:太郎,obj:経済学,gol:S)
 ∧ [S=経営する (agt:太郎,obj:会社,tme:将来)]

(4) “～ならば”などで結ばれた複文 接続詞“～ならば”、“～れば”で結ばれた文の場合、以下のように含意を含む式で表現する。

- 2 私は天気が良ければゴルフに行く
⇒ 良い (sbj:天気) → 行く (agt:私,gol:ゴルフ)

3.2.2 基本的な変換手順

基本的な変換手順を複文を例に図1に示す。まず、自然言語文を形態素・構文解析し、その結果より主節と従属節を取り出し、各述部に係り受け関係のある文節を引数部とする。次に、必須格についてEDR電子化辞書に記載されている情報を利用してラベルを付与する。任意格については、その語に付加しているその語の意味カテゴリーを用いてラベル付けを行う。最後に、主節と従属節から重複している引数を連結定数に置換し、述語式を生成し連言で結合する。

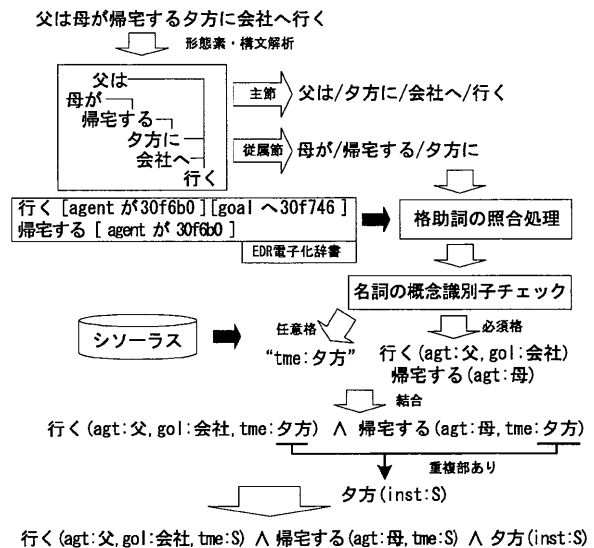


図1 基本的な変換手順

4 おわりに

自然言語文をすでに提案している拡張した述語知識表現[3]へ変換する方法について述べた。複文については、連結定数で結合した素式群で表現する方法について述べた。今後は、より複雑な文を対象に知識表現と変換法の研究開発を行う予定である。

参考文献

- [1] 石川勉:言葉で考えるコンピュータの実現にむけて, 人工知能学会論文誌 (第17回) 論文集, 3C1-04(2004).
- [2] Nguyen Viet Ha, 石川勉, 阿部明典:知識の類似性を利用した概略推論法, 電子情報通信学会論文誌 D 1, Vol.J84-D-1, No.4, pp.389-400(2001).
- [3] 石川勉, 佐々木智彦, 佐藤雅彦:言葉をベースとする拡張型述語論理形式の知識表現法, 人工知能学会研究会資料, SIG-LSE-A303, pp. 25-32(2004).
- [4] 茶筌:http://chasen.aist-nara.ac.jp/hiki/ChaSen/
- [5] 工藤拓, 松本裕治:チャンキングの段階適用による日本語係り受け解析, 情報処理学会論文誌, Vol. 43, No. 6, pp. 1834-1842(2002).
- [6] 日本電子化辞書研究所:EDR 電子化辞書, 日本語動詞共起パターン副辞書,(1996).