

2G-1

意味構造記述データを備えた会話文データベースと
英日会話文の意味処理の試み

河 口 英 二 鎌 田 清 一 郎 脇 山 正 博

九州工業大学 工学部

1. はじめに

従来、自然言語処理としては、学術論文や技術文献のように“整然とした”自然言語文を対象とするものが中心で、日常生活に最も馴染み深い“会話文”処理の研究はまれであった。その理由は、

- (1) 会話文は文法規則による解析が難しい、
- (2) 会話でやりとりされる情報は、発話者の意志、意図、感情などの言外の状況に左右され易い、
- (3) 呼掛け・返答・感嘆・独言のように、通常の文書には現われない文もある、
- (4) “標準的会話文のデータベース”がない、

等が考えられる。本稿は、著者らが提案している自然言語文の意味構造記述式 (Semantic-structure Description formula, SD式)⁽¹⁾⁽²⁾を用いて会話文の意味構造を記述する具体例と、英日会話文データベースの試作、および、SD式を基にした会話文の意味処理のための“言及関数”について述べる。

2. SD式の概要

SD式は“概念ラベル”を原始記号として、“機能項目”、“修飾子”、“結合子”等の記号を組み合わせ、1. 個々の文の意味構造、2. 知識や推論規則、3. 発話者の心理言語行動規則等を一つの文脈自由言語の形式で表記するものである。このSD式は英語や日本語に対する一つの具体的な中間言語である。

まず、機能項目としては

S : 主語	A : 注意の喚起/呼掛け
V : 述語	R : 返答
C : 補語	E : 感嘆
O : 目的語	
OD : 直接目的語	
OC : 間接目的語	
B : 受身表現に於ける行為者	

を用いる。修飾子としては、

NEGA : 否定概念を表わす
PASS : 受身的な述語表現
ASSU : 仮定

MOOD : 法、様相等を表わす

FCUS : 焦点(すなわち、強調)を当てる

を設定している。結合子としては以下の例を含め20種を定めている。

/ : 右辺で左辺を修飾

INCL : 左辺が右辺を包含する

EXCL : 左辺は右辺を含まない

EQUA : 左辺と右辺は等価である

DEFI : 左辺は右辺のように定義する

ANDX : 左辺と右辺の論理積

ORXX : 左辺と右辺の論理和

EORX : 左辺と右辺の排他的論理和

PARA : 左辺と右辺の並列的記述

CAUS : 左辺から右辺が導かれる

3. 種々の自然語文や知識データに関するSD式

会話文では、呼掛け、返答、感嘆等の文型がしばしば現われる、上述のようにSD式ではこれらに対して特別な機能項目を設けている。更にSD式では様々な知識や推論規則も表現できる。以下に例を示す。

(1) 陳述に対応するものとして、[s, v], [s, v, o], [s, v, o, c], ……のような文型を9種定めている。

例01 [s(太郎), v(知る/過去), c([s(林檎), v(赤い)])]

(2) 陳述的なSD式であり会話中の発話意図を表現する

例02 [s(自分), v(質問), o(相手),

c([s(氏名/相手), v(である), c(何)])]

(3) 会話文に特有なものとして、a, r, eを用いて呼掛け、返答、感嘆表現を示す。

例03 a(挨拶/(朝)para(相手/目上))

例04 r(肯定/相手/子供)

例05 e([s(自分), v(気づく), o(失敗/自分)])

(4) 変数や関係詞ラベルを用いて知識の記述や関係節に関する記述力を高めている。

例06 (assu([s(x), v(である), c(親/y)]))

caus([s(y), v(である), c(子/x)])

例07 [s(太郎(\$1)), v(手渡す/過去), od(花子),

oi(本(\$2)) /

[s(\$1), v(発見/昨日), o(\$2)])]

(5) 話者の心理言語行動データ等も表現可能である。

例08 (assu([s(自分), v(知る), c([s(相手), v(咳/継続的)])])) indu

([s(自分), v(予想), c([s(相手), v(風邪/状態)])])]

(6) その他、概念の階層構造の表現

例09 (人間)incl(日本人)

4. 会話文データベースとSD式データの作成

本研究では、会話文(日/英)の例として「NHKラジオ/テレビ会話テキスト」を取り上げ、その中の会話例文を収集し、SD式データを作成し全体を「会話文データベース」として完成されることを狙っている。これは末尾に示すように、1. 話題名、2. 会話

A Database of Conversational Language Texts and a Semantic Processing Approach to Conversational English and Japanese

Eiji KAWAGUCHI, Seiichiro KAMATA and Masahiro WAKIYAMA.

Kyushu Institute of Technology, Faculty of Engineering

状況, 3. タイトル, 4. キーワード, 5. 本文 (日/英) に関する情報を整理したものである。そして, キーワードを除く全ての項目についてSD式データを付加するようにしている。現在, 日本語/英語原文の台詞をそれぞれ約 23,000 文程度収集しており, さらにキーワードやSD式データの作成作業を続けている。

5. SD式に関する言及関数 refer(x,y)

具体的な意味データによる意味処理では各文 (意味表現式) 相互のマッチングが必要である。この場合, 見かけ上では異なる意味表現式であっても知識や推論操作を通じて正しいマッチングが可能となる仕組みが必要である。SD式モデルではSD式同士の「言及」を明確に定義している。言及関数 refer(a,b)は, "aはbを言及している" との意味を持つ。aがbを言及しているとき refer(a,b)=yes で, そうでないとき refer(a,b)=noとなる。以下 refer(a,b) の例を示す。但し a, b, ... 等はいずれもSD式である。: の右側に refer(a,b)=yes となる条件を示す。但し, : が無い場合は無条件に yes となる。

refer(a,a).

refer(a/b,a).

refer(a,b) : (b)incl(a) または (b)equa(a).

refer([s(a1),v(b1),o(c1)], [s(a),v(b),o(c)]).

: refer(a1,a), refer(b1,b), refer(c1,c).

refer(pass(a),a).

refer(a,x). 但し x はSD式変数。

例えば, このような関数を用いて末尾の会話文データベースに対して

[s(自分),v(伝達),o(x),

c([s(y),v(である),c(準備完了)])]

(誰かに何か準備済みであることを伝達する)

なる英文を検索すると以下の答えが得られる。

<ER8104103> Breakfast is ready.

6. むすび

本稿では, 自然語文の意味表現式であるSD式を用いて会話文の意味データを作成し, そのデータを基にして意味処理を行う試みについて述べた。今後の問題点としてはSD式と自然語との相互変換システムや心理言語行動データの整備等がある。

参考文献

- (1) 河口, 他, "会話文の文意構造記述に関する試み", 昭63年人工知能学会全国大会論文集, No.9-1, 1988.
- (2) 鎌田, 他, "会話文と図形情報の意味記述に関する試み", 電子情報通信学会技術研究報告, NLC88-30, 1989.

会話文データとSD式データの例 (R8104)

(話題名) <STORY81041> こちらは僕のお父さんです/併述/[s(これ),v(である),c(父/自分)]

503オーチャード・レイン (こちらは僕のお父さんです)

LIFE AT 503 ORCHARD LANE (This is My Father)*

(会話状況) <SITUR81041> [s(キャサリン(\$a))/[s(\$a),v(である),c(娘/デビットソン)]],

v(準備/(進行)para(様態/繁忙)),

c([s(\$a),v(させる),o([s(家族/当該),v(準備),o(活動/今日)])])]

(ある朝, デビットソン家の娘キャサリンが忙しく, その日のために家族を送り出す準備をしています。)

(One morning, the Davidson's daughter, Catherine, is busy getting the family ready for the day.)

(タイトル) [SDTTL#01:[s(自分),v(告知),o(相手),c([s(朝食),v(である),c(準備完了)])]]]

[JPTTL#01:朝食の用意ができたわよ]

[EGTTL#01:Breakfast is ready]

(キーワード) [[時間:[朝]], [場所:[自宅]], [登場人物:[姉妹]], [話題:[朝食, 用意, 気分, 上々, 父母, 仕事, 出勤]]]

(本文:SD式)

キャサリン:<SR8104101> [a(HAROLD)]+[a(SYLVIA)]

<SR8104102> ++[s(自分),v(促す),o(相手),c([s(相手),v(来る/終点/階下)])].

<SR8104103> ++[s(自分),v(告知),o(相手),c([s(朝食),v(である),c(準備完了)])].

シルビア:<SR8104104> [s(自分),v(告知),o(相手),c([s(自分),v(行く/(mood(意志))para(即刻)])].

キャサリン:<SR8104105> [r(挨拶/相手/親密)]+[s(自分),v(質問),o(相手),c([s(体調/(相手)para(今朝)),v(である),c(如何)])].

シルビア:<SR8104106> [r(良好)]

<SR8104107> ++[s(自分),v(質問),o(相手),c([s((父)para(母)),v(所在/何処)])]

(本文:日本語文)

キャサリン:<JR8104101> ハロルド, シルビア。

<JR8104102> 下に降りていらっしゃい。

<JR8104103> 朝食の用意ができたわよ。

シルビア:<JR8104104> 今いきます。

キャサリン:<JR8104105> 今朝は調子どう?

シルビア:<JR8104106> いいわよ。

<JR8104107> お父さんとお母さんはどこ。

(本文:英語文)

<ER8104101> Harold, Sylvia.

<ER8104102> Come down stairs.

<ER8104103> Breakfast is ready.

<ER8104104> I'm coming.

<ER8104105> Hi! How are you this morning?

<ER8104106> O.K.

<ER8104107> Where are Mom and Dad?