

状況表現に基づく文生成方式

5J-4

瀧塚孝志

(財) 新世代コンピュータ技術開発機構

1. はじめに

談話理解システムDUALSの第2版の作成に当たり、状況理論により意味表現された内部データから、日本語の文章を生成する機能を追加した。本稿では、状況表現から文生成を行う手法と、質問に適した答えを合成する手法を紹介し、意味表現からの文生成のモデルを提案する。

2. DUALSの概要

DUALSシステムは、次の5つのモジュールからなる。

①解析部：日本語文を状況理論に基づく意味表現に変換する。②オブジェクト同定：文中の指示代名詞や固有名詞などの同定を行う。③理解表示：入力された内部表現を日本語文に戻す。④問題解決：入力文が質問文であるとき、その解答となる状況を検索・推論する。⑤応答生成：質問文と問題解決の結果を融合して、応答文を作成する。

これらのモジュール間で受け渡される内部データの表現形式の枠組みを例を用いて示す。

“太郎は急いで隣の花子を呼んだ。”

という文は、およそ次の様に表わされる。

事態：

関係：条件付引数（“呼ぶ”，条件：“急いで”）  
 引数：[ 主格，対格 ]  
 主格：{ 太郎 }  
 対格：条件付引数（{ 花子 }，条件：“隣”）  
 場所：過去<発話時点より前>  
 極性：肯定

注。{}内はオブジェクト，“”内は関係名

入力された1文は、1つ以上の事態(SOA)からなる発話状況(DS)として内部される。事態は、論理的な意味を表わす「関係」と、関係が成立する時間と空間を表わす「場所」と、肯定が成立するか否定が成立するかを表わす「極性」より構成される。

「関係」には、文から時制や時空に関する修飾および否定の情報を取り除いた用言的關係と、限量子の付いた名詞を数量と要素の關係で表わした様な抽象的關係がある。動詞に対応する關係は、深層格で名札付けされた引数を持っている。引数が修飾節/句を持っている場合には、修飾部を条件として持つ「条件付引数」により表わされる。

3. 文生成の手法

文生成モジュールは、前記の内部表現を次の手順により日本語化する。

①出力すべき情報の構築

事態や事態に対する制約は、独立して存在することが出来る。そのため出力する状況が複数の事態を含むときは、他の事態の引数の修飾として表現出来るかを検査する。

また、連体修飾中の表層上は冗長となる引数、例えば“煙を発見した機長”における(機長)\*を削除する。

条件付引数（(機長)，  
 関係：“発見(する)”  
 引数：[ 主格，対格 ]  
 主格：{ 機長 } \*  
 対格：{ 煙 }  
 …………… )

条件の意味が簡単には表現出来ない修飾、例えば“魚を焼く煙”は、何等かの条件があるとして表現されるため、削除する引数は存在しない。

条件付引数（(煙)，  
 関係：“焼く”  
 引数：[ 主格，対格 ]  
 主格：{ }  
 対格：{ 魚 }  
 …………… )

その他意味処理だけに必要な情報も削除する。

②文形/文体の決定

原因結果とか時間的順序関係など2つの事柄の關係、および質問文などの文形を決定する情報は、特別な關係名で示されている。特別な關係名をもとに文形と構文の枠組みを決め、「場所」のスロットに格納されている時制及び時間や位置に関する情報を助動詞や連用修飾として取り出す。こうして、日本語として出力するのに適した意味表現形式にする。

次に、生成される文が“です・ます”調になるように、枠組みとなっている關係に文体の情報を付与する。生成モジュールは名詞だけでも文として生成出来るが、取り扱う通常の文の形態は、一般動詞が關係となっている動詞文と、

「です」が関係となる名詞文，“方がいい”など形容詞が関係となる形容詞文に分けられる。

### ③格助詞／語順の決定

“主格，対格，場所起点”など深層格の名札に対し，“が，を，に，から，より”といった表層格を対応付ける。“は”の生成は，格情報に与えられた定題化の印で生成すると共に，主語の属性や構文の構造，他の語句との関連をもとに“が”から“は”への変換も行っている。

表層的必須格の標準的な語順を「が→を→に」の順に定めた。意味的に強い自由格は前方に，弱い自由格は後方に置いて生成することにより，“エンジンから煙が吹き出した”というような，物語文らしい雰囲気のある文を生成している。質問に答える際など，この語順を変えて特定の格を強調したほうがよい場合，強調されている格に印を付けることによりその格情報を先頭に出力するようにしている。

### ④単語の生成

生成すべき単語が修飾を受けているときは，修飾部を生成し，次にその単語を生成する。短い連用修飾（なぜ，急になど）は，主語の直後や文の先頭に出すことによりリズムを付けるようにしている。1つの単語が複数の修飾を受けているときは，接頭語など強い関係にあるものを最も近くに，短い修飾を次に近く配置するようにしている。

格情報が名詞でなく関係となっているときには，“こと，もの，の”などの形式名詞を挿入し，文法上の調整を行う。数量を表わす修飾部を持っている名詞に関しては，1人，1機，1台など名詞の属性に合った数詞を選択する。接尾語や助動詞は後方から修飾する条件として取り扱っている。

配置の決まった単語（語義を表わすアトム）に対し，生成用の辞書を引き表層上の文字列（語幹＋活用形）を得る。単語の一次的順序を決めて行くときに，活用語尾や音便などの形態素処理を行う。句点は，構文構造の区切りや定題の“は”，長い連用修飾，文間接続詞の後などに挿入している。

## 4. 質問応答

DUALSで可能な質問の形式は，①「なに」型質問（いつ，何処，誰）と②「なぜ」型質問と③「はい／いいえ」型質問である。

### ①「なに」型質問

問題解決より渡される状況は「なに」に相当する情報である。この情報を質問文の「なに」の位置に埋め込み，質問文の文形を平叙文のみに直して生成すべき状況を作り上げる。この状況を3章で述べた手法を適用して解答文を作る。

語順を変更することにより，解答部分を強調出来るならば，データ構造中に強調を示す印を付ける。単純に「なに」の位置に埋め込めない場合は，単語の概念に関する簡単な知識を用いて問題解決より渡された状況を変形し埋め込む。

解答すべき状況が長い修飾条件を持っている場合には，その修飾を取り去り簡潔な解答文になるようにしている。

### ②「なぜ」型質問

問題解決が推論した状況に対し，“～だからです”という関係を被せて生成すべき状況にしている。“なぜだと思いますか”のように「なぜ」が引数中（埋め込み文中）にある場合，“～だからだと思います”のように外側の言い回しも残した構造を作成している。

### ③「はい／いいえ」型質問

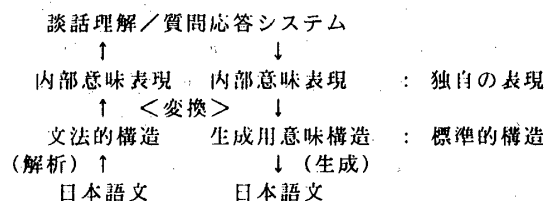
現在のところ問題解決から渡される結果は‘yes’と‘no’だけであり，生成される答えは“はい，～は～です”と“分りません”の2通りしかない。

“いいえ”と答えるためには，質問された状況に対し極性が逆の状況を見付けて否定が成立していることを確認し“いいえ～は～ではありません”と答えるか。常識的な制約や発話された関係から質問状況と両立しない状況を見付け出して“いいえ，～は～です”と答えなければならない。

## 5. 文生成モデル

自然言語処理において文生成という用語は，話題の流れを監視し聴き手の理解に合わせて文を生成するという意味を持つことがある。DUALS-IIシステムでは，問題解決モジュールと文生成モジュールの境界を明確にするため，文生成モジュールは与えられた意味表現を日本語文に直すだけとし，それ以上の機能は問題解決モジュールが行うものとした。

ICOTではDUALS-IIの作成経験をもとに，日本語文標準解析ツールと生成ツールの仕様固めを行った。仕様検討に際し，各種の意味表現形式（DUALSでは状況理論に基づく意味表現形式）と独立であり，汎用性があるインタフェースとなるよう注意が払われた。このとき採用した自然言語処理の標準モデルは，



である。生成用意味構造は，3章の②文形／文体の決定までの情報を意味表現に反映させたレベルに設定した。DUALSではほぼ作り付けになっている生成用の文法も，ユーザが追加出来るよう設計している。

## 6. おわりに

本稿では，状況理論に基づき意味表現された内部構造から日本語文を生成する手法と，いくつかの質問パターンに対し適した解答を生成する手法について述べた。そして，内部の意味表現に独立した標準日本語文生成のモデルを提案した。