

3C-9

日本語交通規則文に関する Q-Aシステムの意味解析部

竹本 信治 · 中川 聖一
(豊橋技術科学大学工学部)

1. はじめに

本Q-Aシステムは運転免許の学科試験で出題される交通規則の問題文を自然言語で入力し、その文を論理表現に変換して問題の真偽を判定するものである。^[1] 本システムは、構文解析部・意味解析部・推論部・学習部の4つの処理系と、内部データとしての見出し辞書・日本語辞書・知識データベースから構成されている(図1)。

意味解析部は構文解析部と推論部の橋渡しの役目を担う部分であり、係り受け処理により解析された文のtree構造から、論理的枠組みの明解な一階述語論理の意味表現に様相演算子(must, can)を追加した様相論理形式に変換する。交通規則の問題文は自然言語の持つ構文的意味的曖昧さが比較的少なく、論理表現で記述することが一般的な文に比べて容易であるために問題の対象世界として選択した。この意味解析部は「日本語の曖昧な表現をどの程度正確に述語論理に変換できるか」という問題を抱えている部分である。

本稿では、このQ-Aシステムの意味解析部のアルゴリズムの改良を行ったので報告する。

2. 日本語辞書の形式^[1]

日本語辞書は見出し語辞書の内容を frame名とするフレーム表現で表される。フレーム表現はWinstonのLISP(2nd ed.)^[2] で使用されているものと同一のものとする。一般的な形式を以下に示す。(以下、*: 0回以上、+: 1回以上の繰り返し)

```
<frame> ::= (<frame名><slot>*)
<slot> ::= (<slot名><facet>*)
<facet> ::= (<facet名><値>*)
<slot名> ::= ako/cat/struc ...
<facet名> ::= value/if-needed/if-added
```

なお、スロット名の "struc" は意味解析のための構造化データであり、次の形式の値を持つ。
(<述語> ([<格><限量子> (<属性>*)]))
属性はその格の取り得る要素の性質や上層概念を表し、限量子は∇: 全称作用素、∃: 存在作用素である。

3. 意味解析部

3.1 意味解析部の構成

図2に、意味解析部の基本構成を示す。様相論理変換プログラムは大きく分けて、1) 自立語の探索、2) 述語の格フレームの同定、3) 変数処理の3部分より構成される。ここで変数処理とは一階述語論理式の束縛変数、全称作用素(全称記号∇にかかるとある変数)と存在作用素(存在記号∃にかかるとある変数)の処理に関するところであり、これらの情報はすべて日本語辞書の動詞の struc中に蓄えられている限量子の場所に記述されている情報を用いる(2. 日本語辞書の項参照)。意味解析部は構文解析結果を入力とする。本Q-Aシステムの構文解析部は本学で製作されたものであり^[1]、その結果は文節の切り出し・係り受け処理を施され格構造に変換される。以下に、その一般形式を示す。

A Semantic Analysis Part of Q-A System
on Japanese Traffic Regulations
Shinji Takemoto and Seiichi Nakagawa
Factory of Engineering, Toyohashi University of Technology

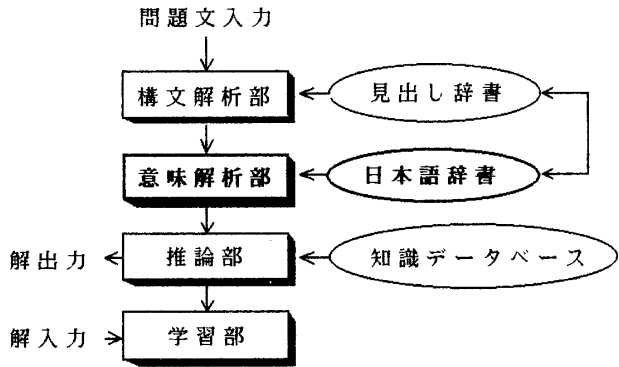


図1 システムの構成

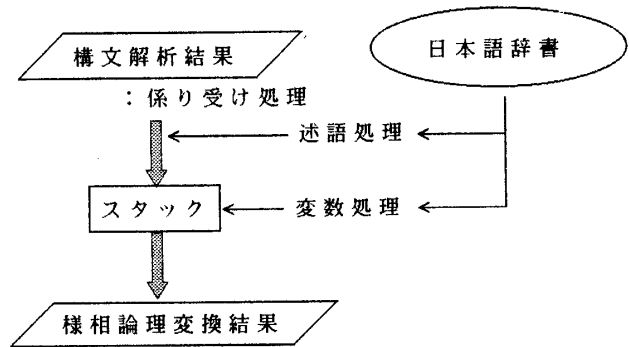


図2 意味解析部の構成

```
<句> ::= (<種別><単語><活用><句>*)
<種別> ::= <潜在格> | <接続詞>
           | rentai | renyou | bun | fukusi
```

意味解析部は、上記の構造を持つリストを以下に示す様相論理の形の意味表現に変換する。

```
<述語式> ::= (<prefix><述語式>*)
           | (<述語名><引数>*)
<prefix> ::= (∇<変数>*) | (∃<変数>*)
           | must | can | impl | and | or | not
```

ここで対象世界を交通規則文に限定した事により、全称作用素∇を存在作用素∃の前とした。交通規則の世界において、全ての対象物がある特定の1つのものに対して成り立つ事があるてはならないからである。

3.2. アルゴリズム

従来のアルゴリズムは「~のとき」のフレーズの中を特別に処理していたが、本アルゴリズムでは「~のとき」も<種別>に "bun" を持つ<単語>に係るものとみなし、単一の処理によって論理式を求める。また、変数の同定の処理を各<単語>の処理と同時に進めていたのに対し、すべての述語に変換した後に変数の同定に関する処理を行った。この2点を改良すること

によって、アルゴリズムを統一化し、変数の同定はより広範囲の述語を参照することができた。従って、変換された論理式はより正確なものとなった。
以下に、構文解析結果から意味表現の内部形式への変換手順を示す。

- (1) その時点で最初にある<句>を取り出す。
- (2) その<句>の<活用>の箇所をみて、それがmust, canを含むか事実を表す文かを調べ、記憶する。
- (3) その<句>の先頭の<単語>の意味データを辞書から取り出し、スタック中の変換の終了した述語の変数との同定を行って、スタックに積む。
- (4) その<単語>に係る格や修飾するものを構文解析結果中から探し述語の変数同定を行ってスタックに積む。(1)~(4)で処理された<句>はその時点で文より取り除かれる。
- (5) 文に<句>が残っていれば、(1)~(4)を繰り返す。(複文の処理)
- (6) <単語>の取る潜在格に相当する<句>がなければ、その<単語>の辞書中の "if-added" 項を参照してdefaultの述語を挿入する。
- (7) すべての変数に対し、∀:全称作用素、∃:存在作用素のどちらになるかを辞書の知識を用いて決定する。
- (8) 最後に、(3)のところで求めておいた様相論理をスタックに積む。
- (9) スタックから意味表現の内部形式を取り出す。

上記のアルゴリズムでの実行結果とトレース結果を図3に示す。

4. むすび

今回、述語論理表現への変換法を改良した。このアルゴリズムでは、「~のとき」の処理を他の<句>と同一のものとしたことにより、述語の相互関係を構造的に統一することができた。また、変数同定により広い範囲から情報を得て、正確な論理表現を得ることができた。意味解析部は現在、自立語の辞書登録数が約50単語で、主に交差点付近での通行ルールについて処理可能である。この意味解析部はシステム各部の中で最も重要な部分であり、さらに一般的な文まで扱えるように改良を検討している。

謝辞

議論して頂いた大学の山本幹雄氏に感謝いたします。

【参考文献】

[1] 中川聖一, 千田滋也: 交通規則文章題の理解と質問応答
電子情報通信学会論文誌
Vol. J70-D NO. 11 pp. 2280~2286 (1987年11月)
[2] Winston P. H., Horn B. K. P.: "LISP" (2nd ed.)
Addison-Wesley (1982)
[3] 自動車免許試験問題専門研究会編:
普通免許試験によくでる学科テスト、日本文芸社

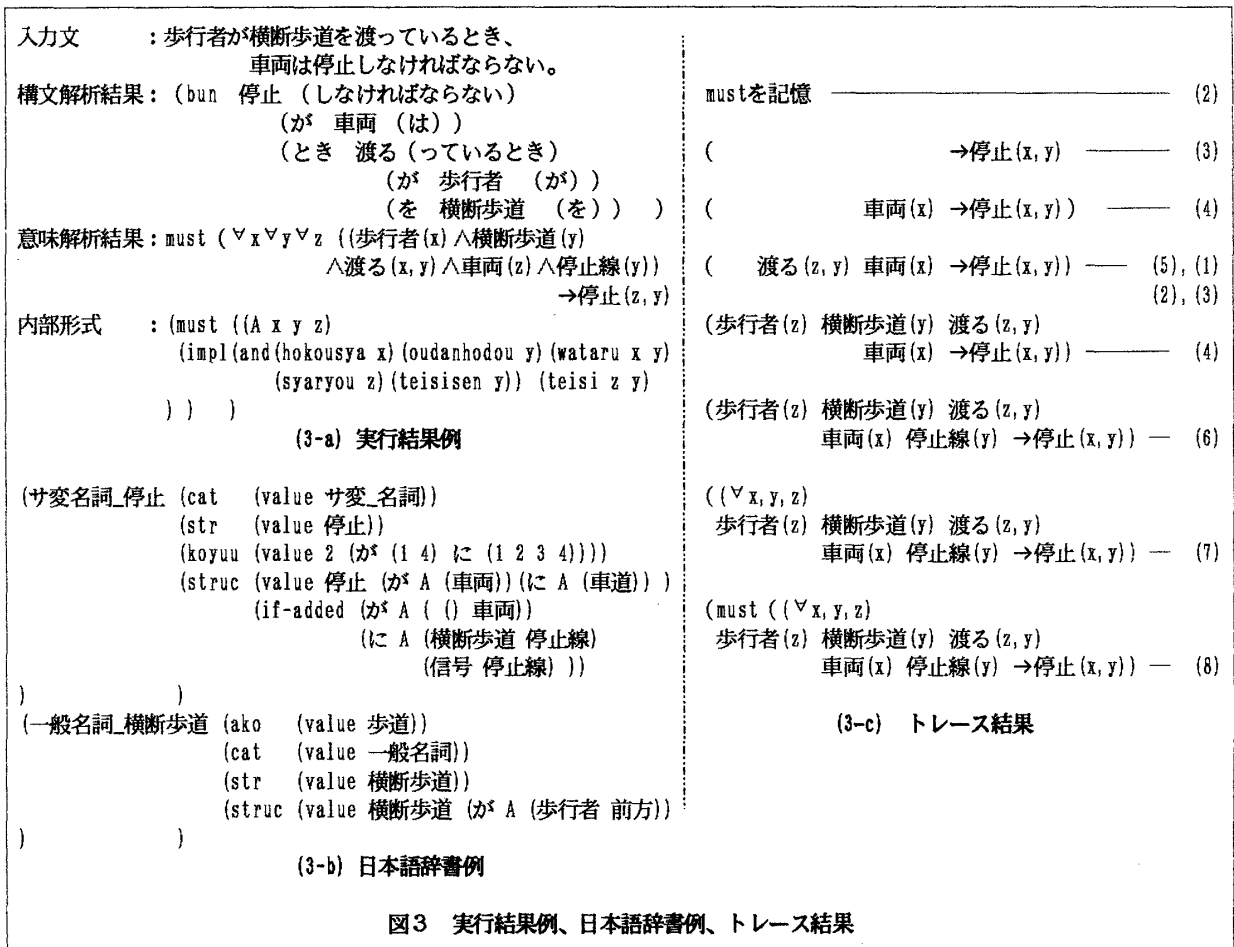


図3 実行結果例、日本語辞書例、トレース結果