

# HICATS / JE の文法作成支援環境

6J-7

坂本浩一、白井孝雄（日立ソフトウェアエンジニアリング（株））  
吉村紀久雄、梶博行（（株）日立製作所）

## 1. はじめに

日立日英翻訳支援システムHICATS / JE (Hitachi Computer Aided Translation System / Japanese to English)は、プログラム部・文法部・辞書部の3つの部分より構成される。プログラム部は、文法規則のインタプリタであり、機能は固定化されている。文法部は、翻訳用文法規則の集りであり、文法記述言語GDL (Grammar Description Language)によって記述されている。辞書部は、単語の集りである。

このうち、文法部については、翻訳精度向上のための改良・変更が頻繁に行われるため、HICATS / JEの開発にあたっては、文法作成支援環境の整備が重要な課題の一つとなった。本稿では、HICATS / JEにおける文法の作成に用いた文法作成支援ツールについて述べる。

## 2. 文法記述言語GDLの概要

GDLはHICATS / JEにおける文法記述言語である。その概要を以下に述べる。

### (1) 概念依存図式

HICATS / JEでは、入力日本文の中間表現として概念依存図式を用いている[1][2]。これは、一種の木構造であり、ノードが概念を、アークがノード間の関係を表している。さらに、各ノードには、種々の属性（「時制」など）が、各アークには格コード（主格を表す「A」など）が付与されている（図1）。

日本文

「計算機によって効率が向上した」

概念依存図式

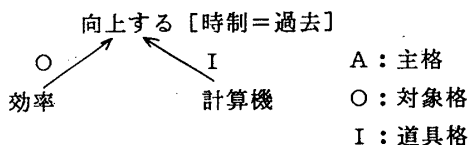


図1. 概念依存図式の例

### (2) 文法規則

GDLで記述した各文法規則は、概念依存図式に対する

- (a) 木構造の変形操作
  - (b) ノード内属性値の参照/更新
- を指定するものである。例を図2に示す。

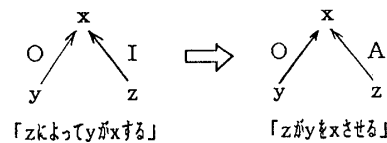


図2. 文法規則の例

### (3) 文法部の構成

文法部は、部分文法という単位に分割される（図3）。部分文法は、通常のプログラミング言語におけるサブルーチンと同様に、1つの機能（例えば係り受けの決定など）を実現するための文法規則の集りである。

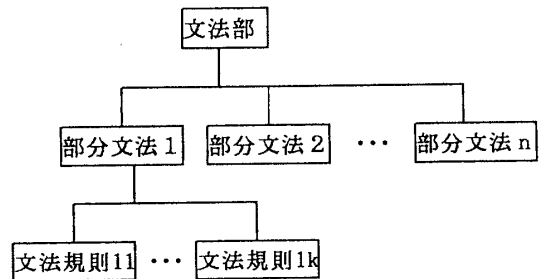


図3. 文法部の構成

## 3. 文法作成支援ツールの開発

GDLは、HICATS / JE用の独自の言語であるため、処理系をはじめ各種の開発ツールは一切存在しなかった。そこで、GDLによる文法作成作業に対して通常のプログラミング言語と同様の作成支援環境を提供するため、インタプリタと併せ以下のツールを用意した。

- (1) GDLデバッガ
  - (2) GDLクロスリファレンス出力ツール
  - (3) GDLドキュメンテーションツール
- 以下、これらのツールの概要を述べる。

4. GDLデバッガ

本ツールは、GDLで記述した文法規則のデバッグツールであり、以下の機能を持つ。

- (1) シングルステップ実行機能  
ブレークポイントとして指定された部分文法の実行中は、1つの文法規則の適用ごとに実行を一時中断し、ユーザの指定を待つ。ここでユーザが指定可能な機能として、(2),(3)がある。
- (2) 適用規則の表示機能  
直前に適用した規則の内容を表示する
- (3) 概念依存図式の表示機能  
概念依存図式の現在の状態を表示する。この際、木構造の表示とともに、各ノードの持つ属性値の表示も可能である。
- (4) 実行時間表示機能  
部分文法ごとに、実行に要したCPU時間を表示する。
- (5) 統計情報表示機能  
適用した規則の個数を表示する。  
(4),(5)の機能は、作成した文法の処理効率を評価する際に有効である。

5. GDLクロスリファレンス出力ツール

本ツールは、GDLで記述された文法に対してクロスリファレンスリストを出力するツールであり、以下の機能を持つ。

- (1) 属性クロスリファレンス出力機能  
概念依存図式の各ノードが持つ属性に関するクロスリファレンスリストを出力する機能である。出力例を図4に示す。本図中、例えば属性「時制」の属性値「過去」を、部分文法「ABC」の規則「C555」が参照していることを表す。ただし、「\*」は属性値の更新を意味する。

<属性名>	<属性値>	<部分文法名(規則名)>
時制	現在	ABC(A001) *DEF(B002)
	過去	*ABC(C555)
		:
		:

図4. 属性クロスリファレンスの出力例

- (2) 格コードクロスリファレンス出力機能

概念依存図式の持つ格コードに関するクロスリファレンスリストを出力する機能である。出力例を図5に示す(見方は図4と同様)。

<格コード>	<部分文法名(規則名)>
A	XYZ(X621)
I	*PQR(C114) *PQR(K809)
	:
	:

図5. 格コードクロスリファレンスの出力例

6. GDLドキュメンテーションツール

本ツールは、GDLで記述された各文法規則のドキュメント作成を支援するツールであり、以下の機能をもつ。

- (1) 木構造変形操作の図示機能  
GDLで記述された各文法規則(1次元的な文字列)を解析し、その文法規則が意味する、概念依存図式に対する木構造の変形操作を(2次元的に)図示する機能である。例として、図2に相当する文法規則の図示例を、図6に示す。

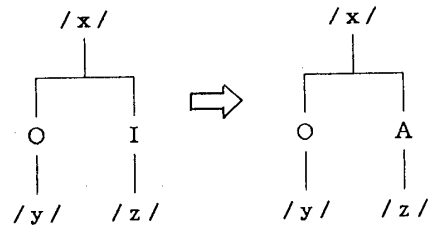


図6. GDLドキュメンテーションツールの出力例

7. おわりに

HICATS/JEで用いている文法作成支援ツールについて述べた。さらに各種のツールを開発することにより、より良い文法作成支援環境を提供してゆくことが今後の課題であると考えます。

参考文献

[1] 梶、新田：概念依存図式からの英文生成；情報処理第28回全国大会 5L-3 (昭59-3)  
 [2] 新田、梶：概念依存図式による日本語文の意味のモデル化；情報処理第28回全国大会 5L-4 (昭59-3)