

## 2F-3

## 自然言語インタフェース構成モジュールにおける領域知識共用

山口 智治

市山 俊治

NEC 関西 C&amp;C 研究所

## 1 はじめに

自然言語インタフェースの対象領域に対する可搬性(対象領域適応性)は、その処理機構の対象領域に独立な部分と依存する部分をいかに切りわけかが課題となる。

筆者らが開発をすすめている自然言語インタフェース構築キット“IF-Kit”[市山91, 谷91]では、構文解析部、文脈処理部[宮部92]、タスク解析部[山口91]、対象言語生成部[飯野91]の各モジュールの結合による構成をとり、システムの動作を定義する処理知識について対象領域に依存する部分をタスク解析部に限定し、他のモジュールを対象領域に独立な部品とすることによって、インタフェースとしての可搬性を高めた。

一方、処理するデータに関する対象知識についても、その利用形態に合わせてモジュール毎に記述されることもあるが、各モジュールで対象領域に依存する知識を持ったままにするならば、対象知識の大規模化とともにその内容に重複が生じ、整合性の管理が困難になる。また、このことは対象知識ベースを構築する際に、重複する情報を各モジュール毎に登録する必要があることを示唆し、余計な労力を要することになり、可搬性を損ねてしまう。

本報告では、対象領域に対して、より可搬な自然言語インタフェースの構築を目的として、領域に依存する対象知識へのアクセスをひとつのモジュールに限定し、かつ他の各モジュールからもその対象知識を参照可能にする、領域知識共用の枠組について述べる。また、関係データベースを対象システムとする日本語インタフェースの開発例について、その各モジュールで参照する領域依存な対象知識についてふれる。

## 2 処理モジュール毎の対象知識

IF-Kitでは、自然言語インタフェース本体が大きく次の4つのモジュールに分けられている。

構文解析部: 領域独立なレベルでの意味処理を含む文構造の解析  
文脈処理部: 文内および前文を参照した省略補完、照応解消  
タスク解析部: 対象領域依存な意味処理、アプリケーションでの実行タスクの決定

対象言語生成部: データベース検索言語(SQL)生成

Sharing Domain Knowledge among Modules of Portable Natural Language Interface

Tomoharu YAMAGUCHI and Shunji ICHIYAMA

Kansai C&C Research Lab., NEC Corp.

各モジュールは、モジュールの動作を決定する処理知識と、入出力データに関する対象知識をもつ。領域に依存する処理知識については、タスク解析部のモジュールに集中させ、他のモジュールの処理知識から領域依存の情報は排除されて、処理機構のモジュール化がなされている。

しかし、各モジュールの入出力データに領域依存なものが含まれることは回避できない。対象知識に関しては、タスク解析部以外のモジュールでも対象領域依存情報の利用が必要である。例えば、構文解析部で参照するシソーラスや対象言語生成部で参照するデータベーススキーマなどである。

処理のモジュール性から見て、それぞれのモジュールで利用する形態の対象知識をモジュール毎に記述することもできるが、対象知識には対象領域世界の記述については複数のモジュールで類似した情報が利用されるので、全体を見通した時に知識の重複が生じてしまう(図1)。

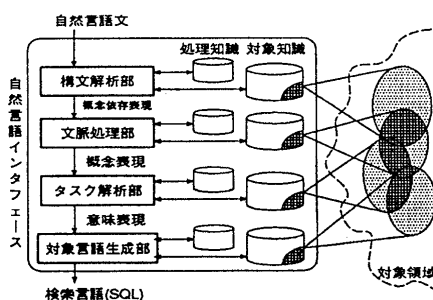


図1. モジュール構成と対象知識の重複

## 3 対象知識の共用

## 3.1 対象領域依存知識管理モジュールの限定

モジュール毎にもつ対象知識に重複があると、対象知識構築時の労力の増大や、変更が発生した時にそれらの整合性の管理の問題が生じる。そこで、処理モジュール構成にとられることなく、かつ各処理モジュールで利用可能であるように対象知識を記述・利用する枠組が望まれる。

領域依存な対象知識もタスク解析部に集中すれば、対象領域に依存する情報は処理知識、対象知識ともタスク解析部に限定して管理される。こうすれば領域知識構築および更新は、タスク解析部のもつ知識のみを操作すればよいことになる。

また、タスク解析部のもつ領域知識は意味ネットワークに基づく表現を用いて記述されており[山口91]、タスク解析部以外のモジュールで利用する知識の追加も容易である。

3.2 領域知識の共用

タスク解析部は領域知識の参照に関してネットワークの検索機能を持つ。他のモジュールからの領域知識の参照がタスク解析部の機能を部分的に呼び出すことで実現される。

各モジュールのデータ構造に依存することなくタスク解析部の機能を利用するために、他のモジュールとタスク解析部の間に、「領域知識要求発生部」と「領域知識要求処理部」を設ける(図2)。領域知識要求発生部は各モジュールのデータ構造と領域知識要求表現の変換をおこなう。領域知識要求処理部には、領域知識要求表現から要求に応じた手続きを起動してタスク解析部の機能を利用するために、タスク解析部の機能を利用する手続きの集合(表1)をもたせる。

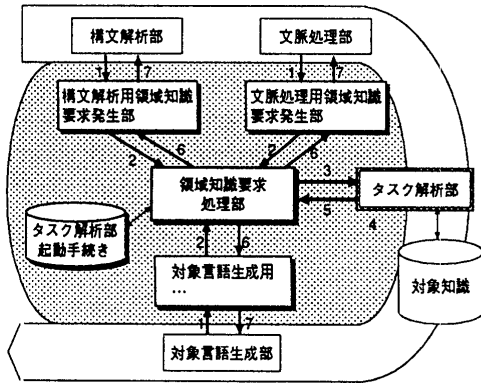


図2. モジュール間での領域知識共用の流れ

- 1: 領域知識利用状況の検出
- 2: 領域知識要求の形式化
- 3: 要求の解析およびタスク解析部機能利用手続き起動
- 4: タスク解析部機能利用手続きの実行
- 5: 実行結果の返信形式化
- 6: 返信結果の利用形式化
- 7: 領域知識の利用

表1. タスク解析部機能利用手続き群(部分)

機能分類	基本機能利用手続き内容
概念検索	異なる日本語表記の単一概念指示の判定 上位概念関係の判定(視点指定あり) 下位概念関係の判定(視点指定あり) 特定の概念に最も近い概念を候補概念群から選択
属性参照	属性関係の判定 属性間問の属性軸の一致の判定
日本語表記取得	対象システム記号からの日本語表記の取得 比較対象に依存する日本語比較表現表記の取得 数量概念に対する単位の取得
システム記号取得	日本語表記からの対象システム記号の取得 日本語表記から取得されるシステム記号の分類
データベース構成	データベーステーブルを代表するフィールドの取得 データベーステーブルのキーフィールドの取得 対象データベースでのデータ型の取得 データベーステーブルの結合条件の取得 特定フィールドから結合可能なフィールドの取得

4 各モジュールで利用する領域知識

関係データベースを対象アプリケーションとして構築された自然言語インタフェースの各モジュールでの領域知識の利用例を以下に示す。

構文解析部

並列構造推定 語概念の組合せで並列になりやすいものの推定(ex. 「ホテルとレストランの駐車場…」 “ホテル” に対して “レストラン” と “駐車場” のどちらが並列になるか)

比較対象判定 比較対象の属性判定(ex. 「100室以上のホテル」 “ホテル” は “100” の比較対象になるか)

ヘッド概念判定 修飾関係のある2概念についてどちらがヘッド概念であるかの判定(ex. 「最大の利益」「利益の最大」についてヘッド概念は “最大”)

文脈処理部

相対的兄弟関係判定 ある概念を視点としたときの2つの概念の兄弟関係の判定

属性関係 2つの概念について一方が他方の属性であるかの判定、また、どちらが属性であるか

範囲判定 大小関係、包含関係、重複関係の判定(ex. “四国” は “徳島” を包含する, “会社員” と “研究者” には重複がある)

間接的関連判定 ある概念を通じた他の2つの概念の関連性の判定(ex. “パソコン” と “clock 数” は “CPU” を通じて関連する)

検索言語生成(および確認文生成)部

データベース構成 フィールドの属するテーブル、キーフィールド、テーブル結合条件など

日本語表記 データベース構成要素(テーブル、フィールド)、検索言語構成要素(コマンド、関数)、フィールド間の関係などに相当する概念の日本語表記

5 おわりに

自然言語インタフェースの可搬性の向上のためのモジュール間での領域知識共用の枠組について述べた。

処理知識のみでなく対象知識についても、対象領域に依存する知識をひとつのモジュールに集中した。また、領域知識をもつモジュールへの情報要求を処理する機構を導入し、他のモジュールでの領域知識の利用を可能にすることによって、本来の言語処理能力を損なうことなく可搬性を向上できる。領域知識要求処理部を充実させれば、言語処理能力の向上の可能性も期待される。

さらに、各モジュールで利用する領域知識について、関係データベースを対象システムとする日本語インタフェースの例を通じて述べた。

[参考文献]

[市山91] 市山俊治, 村木一至: 自然言語インタフェースの構築キットの提案, 43 回情処全大 1H-2, 1991.  
 [谷91] 谷幹也, 飯野香, 山口智治, 市山俊治: 自然言語インタフェース構築キット:IF-Kit, 信学技法 NLC91-62, 1991.  
 [飯野91] 飯野香, 市山俊治: 概念依存構造からのデータベース検索式の生成, 43 回情処全大 1H-4, 1991.  
 [山口91] 山口智治, 市山俊治: 拡張意味ネットワークの連想検索に基づく言語理解, 43 回情処全大 5H-10, 1991.  
 [宮部92] 宮部隆夫: ヒューマンマシン対話における照応解析, 44 回情処全大 5P-5, 1992.