

日本語文解析器および生成器のサーバ化に関する研究

5 B-2

原田 雅之, 唐澤 博

山梨大学 工学部 電子情報工学科

E-Mail: (harada, karasawa)@opal.esi.yamanashi.ac.jp

1 はじめに

自然言語を入出力インタフェースとするシステムを多数のユーザが使用する環境では、ユーザ数分の日本語文解析器および生成器を逐一起動して用いるのはメモリ効率及び実行効率が悪い。そこで、マルチユーザ対応の漢字変換システム Wnn の場合と同様な考え方で、解析器・生成器をサーバ化して、各ユーザ・アプリケーションをクライアントとする構成のシステムを開発した。クライアントは文解析サーバから、ユーザの入力した日本語文が解析されて意味構造になったものを受け取る。一方、F構造を文生成サーバに送ることによって日本語出力文を受け取る。双方の過程で、文脈情報も受け渡しが行われて更新される。この方式を自然言語対話 Agent システムに適用した結果を報告する。

2 日本語文解析器および生成器

2.1 日本語文解析部

ユーザからの自然言語での入力をシステムに理解できる構造に変換するモジュールである。まず、入力日本語文は LAX 流の形態素解析部において形態素列に変換される。次に、構文・意味解析、文脈解析部 [1][2] において LFG(語彙機能文法)に変換される。さらに、意味構造生成部で USF(発話意味構造)に変換され、意図解析部において UGF(発話意図構造)に変換される。最後に、社会関係解析部 [3] においてシステムの行動が決定されて適切な実行スキーマが呼び出される。

2.2 日本語文生成部

システムの発話内容を自然言語に変換するモジュールである。実行スキーマから UGF を受け取り、意味生成部において USF に変換される。次に、構文生成部において F 構造(語彙機能構造)に変換され、形態素生成、代用・省略処理などの文生成部を経て日本語文が出力される。

3 サーバ・クライアントモデルの構成

自然言語対話 Agent システムは図 1 に示す構成となる。まず、トップ・スキーマが起動され、ユーザに対して開始・終了の挨拶をするものとなっている。

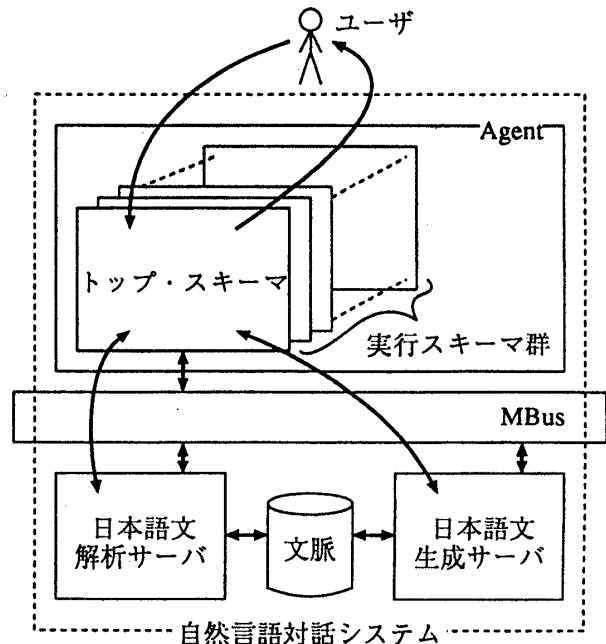


図 1: 自然言語対話 Agent システム

3.1 サーバ化の目的

逐一起動型の構成の場合、ユーザ数分の非常に巨大な日本語文解析器および生成器を毎回立ち上げて用いることから、メモリ効率及び実行効率が悪いことが多い。そこで、解析器・生成器をサーバ化して、各ユーザ・アプリケーションをクライアントとする構成のシステムの場合では、メモリ効率及び実行効率が良くなると考えられる。さらに、1つのサイトにおいて数多くの言語が扱われていることは稀であることから、そのサイトで扱われている言語分の文解析・生成サーバ対のみが常駐していれば良いことになる。

3.2 日本語文解析部の実現方法

3.2.1 クライアント

ユーザの入力する文章を標準入力より読み込み、文脈情報と共に解析器専用入力スプールに置いて、ユーザが扱う言語の文解析サーバに対して要求を出す。そのサーバから終了信号が送られて来たら、解析器専用出力スプールに置かれた更新後文脈情報と発話意図構造を受け取る。

