

## BTH パーザを使用した音声入力質問-応答システムの試作

屋野 武秀, 笹島 宗彦, 河野 恭之

株式会社東芝 関西研究所

神戸市東灘区本山南町 8-6-26 / TEL:078-435-3104 / FAX:078-435-3162

{yano,sasa,kono}@krl.toshiba.co.jp

## 1. はじめに

本稿では、大規模なラティスに対しても、高速に解析する文テンプレートハッシュ方式 (BTH) [1] によるパーザを使用した質問-応答システムについて報告する。このシステムは、カーナビゲーション (以下カーナビ) タスクをモチーフとして、場所/区間に関する質問に対して回答し、ノート PC 上でほぼ実時間で動作している。

## 2. 文テンプレートハッシュ方式 (BTH)

口語音声の不要語などの取り扱いが容易な音声認識方式として、指定したキーワードを抽出するワードスポッティング方式がある。その認識結果は抽出された単語 (不正解のものを含む) をノードとするグラフ形式 (以後ラティス) で与えられる (図 1)。そして、音声区間に発声されたキーワード列を推定し、発話内容を推定する。しかし、ラティスから生成可能な単語系列数が多くなると、受理可能な単語系列を選択する計算量が増加する問題があった。

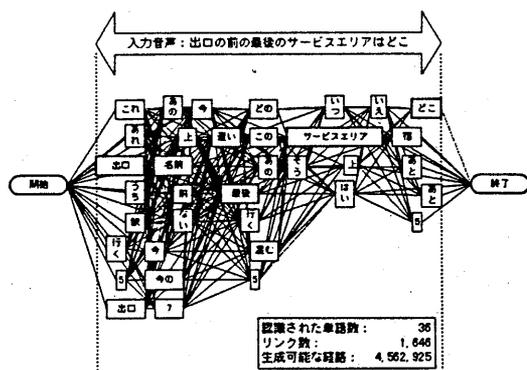


図 1 単語ラティス

BTH[1]は、高速なラティス解析方式であり、その特徴的な構成として、文型辞書、文テンプレートハッシュ辞書が挙げられる。

文型辞書: 受理可能な品詞系列と文型 ID とを記述

文テンプレートハッシュ辞書: 文型辞書をコンパ

ルし、辞書に登録されている各品詞について文中の出現箇所と、所属する文型の集合との対応関係を要素とする表。つまり、「ある品詞に属す単語が文中のある位置 (順序) にあるならば、その単語を採用した文はどの文型に所属しうるか」という記述を要素に持つ。

BTH パーザは、ラティス上にある前方 (図 1 では左側) に接続されたノードから渡される受理可能な文型集合と、文テンプレートハッシュ辞書で得られる対象ノードを通過しうる文型集合の積集合を計算することにより受理可能な文型の絞り込みを行う。また、絞り込み結果を後方のノードに伝播し、後方のノードで同様の処理をすることによって、更に絞り込みを進める。以上の絞り込み手続きを開始ノードから終了ノードに向けて進め<sup>1</sup>、終了ノードに伝播された文型が受理可能な文型の集合となる。受理可能な単語系列は得られた受理可能な文型をもとにラティスをスキャンすることによって得られる。

## 3. 質問応答システム

次に、BTH パーザを利用した質問-応答システムについて述べる。この質問応答システムは、カーナビタスクをモチーフとして、場所/区間に関する質問を受け付けるものである。

本システムの構成は、地図を表示し、経路を移動しながら現在地を更新するアプリケーションモジュール (図 2) と、BTH パーザと問題解決器を含む対話モジュール、音声認識エンジンの大きく 3 種類に分けられる。これらはすべて一台のノート PC 上で動作している (図 3)。

音声認識方式は、ワードスポッティング方式を採用している。基本的に自立語のキーワードを指定

<sup>1</sup> 開始ノードからは全ての文型が伝播される

し、認識結果のワードラティスを BTH パーザで解析する。そして、受理可能な単語系列と対応する文型をもとに各候補に対応するクエリを生成する。

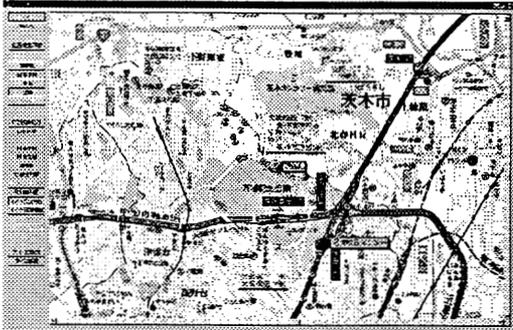


図 2 アプリケーション画面例

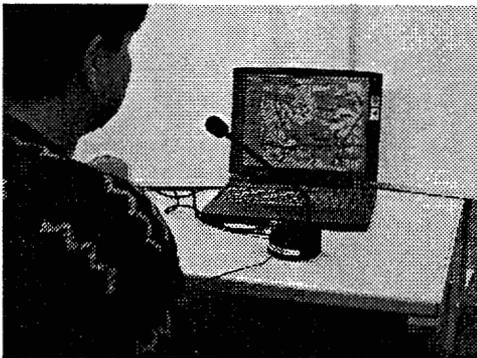


図 3 カーナビ質問-応答システム

音声認識結果は曖昧性を持つため、受理可能な単語系列も複数存在する。このため、複数の候補単語列を考慮して問題解決し、回答内容を決定する。具体的には、認識スコアに加えて問題解決時にヒューリスティックによるスコアリングを加味し、意味的スコア調整の処理を行う。また、回答時に解析結果のスコアが低い場合には確認の問い返しを行う。

本システムは、ドメイン/タスク汎用性を保つために、ドメイン/タスクに固有であると考えられる発話意図理解や、問題解決、応答生成は、意図変換テンプレート、問題解決スクリプト、応答生成テンプレートを準備し、それぞれインタプリタを介して問題解決手続き、応答文生成を実行する(図4)。

本システムが受理する文法は、場所を現す場所句文型と、場所/区間について問い合わせる文型からなる<sup>2</sup>。文法は、倒置文型などを考慮して作られ

<sup>2</sup> 問い合わせ文型の中に場所句文型を挿入し、組み合わせ

ており、より口語的な発話も受理可能なものとなっている。対話例は次のようになる。

U1: 出口はどこ。

S1: 出口は西宮インターです。

U2: 目的地までどれくらい。

S2: 現在地から目的地までについてお尋ねですか。

U3: はい。

S3: 現在地からポートタワーまでの距離はおよそ30km、45分程度かかります。

システムは、PC (CPU: PentiumII 266 MHz) 上で、概ね約 2 秒以内 (音声認識約 1 秒、認識結果解析・問題解決約 1 秒) で返答する。

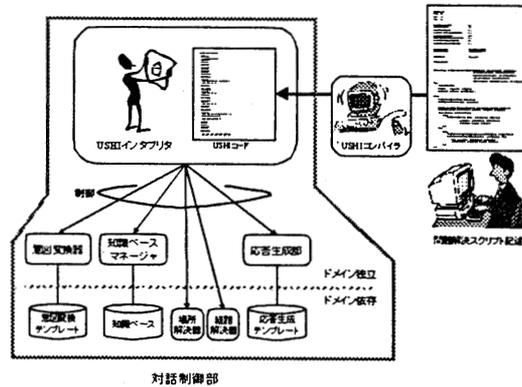


図 4 対話部のモジュール構成

#### 4. まとめ

BTH パーザを利用した質問-応答システムを試作した。本システムは、カーナビをタスクとし場所、区間に関する質問を受理するものである。本システムをノート PC 上で動作させた結果、概ね約 2 秒以内で返答することを確認した。

今後の課題として、対話履歴を用いた省略補完、参照同定処理の導入が挙げられる。

#### 参考文献

[1] 屋野, 笹島, 河野: "文テンプレートハッシュ方式によるラティス解析方式と質問-応答システム", 情処研報, SLP23-13 (1998-10)

た文法となる。