

# 音声によるデータベースアクセスの同音語問題

5K-9

荒川ゆう子 太田由香 増永良文

図書館情報大学

## 1. はじめに

本研究では音声によるデータベースアクセスのプロトタイプシステム[1,2]を評価した結果明らかになった同音語の問題についてインスタンスキーという新しい概念を定義・導入し、問題の解決にあたった。本稿では我々が構築しているプロトタイプシステムの概要と、特に同音語の問題解決に向けた研究・開発の結果を報告する。

## 2. 音声によるデータベースアクセスシステムのプロトタイプ

本研究ではまず、現在利用可能な音声認識装置や音声合成装置を使用して日本語音声によるリレーショナルデータベースアクセスシステムを試作し、問題点を明らかにした。このシステムでの処理の流れを図1に示す。本研究で開発したソフトウェアモジュールは大別すると次の4つになる(図1)。

- (a) センテンス確定モジュール：音声認識装置によって認識された文字列のセンテンスパターンを同定し、SQL文への変換が可能な形に変形する。
- (b) SQL生成モジュール：認識されたセンテンスをSQL文に変換する。
- (c) SQL処理モジュール：上記モジュールで生成されたSQL文でリレーショナルデータベースへ質問を発行し、結果として結果リレーションを受け取る。
- (d) 質問結果発声モジュール：結果リレーションを音声合成装置の入力系列に変換する。

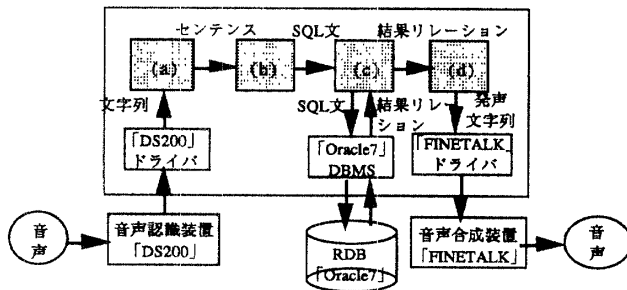


図1 音声によるデータベースアクセスシステムの処理の流れ

## 3. プロトタイプシステムにおける同音語問題

このプロトタイプシステムにより、不特定話者が音声で入力した質問文でRDBにアクセスし、検索結果を音声で聴くことが

Homonym Problem of a Query-by-Speech Interface to Relational Databases

Yuko Arakawa, Yuka Ohta and Yoshifumi Masunaga

University of Library and Information Science

1-2 Kasuga, Tsukuba, Ibaraki 305, Japan

可能となった。しかし音声アクセスの際、同音語の識別に問題があることが明らかになった。以下で、この問題を入力時と出力時に分けて述べる。

### (a) 入力時の問題

本研究で用いた音声認識装置は、センテンスパターンの登録単語として「児島[こじま]」、「小島[こじま]」、「古島[こじま]」をこの順に設定した場合、「こじま」と音声で入力すると、最初に登録されている「児島」を返してしまう。したがって、機械的制約により小島と古島はそもそも認識できない。そこで、人名など同音語問題を引き起こす可能性のある属性については、(i) センテンスパターンの単語登録を「読み」で行い、リレーションに「読み」という新たな属性を加えることで読みで検索を行い検索漏れをなくし、さらに、(ii) その中から所望のタプルを一意に識別するための特別の手立てを導入する。

### (b) 出力時の問題

音声出力では漢字の情報までは伝えられないため、出力された結果に対して的確な判断ができない。このような場合は、名前だけでなく他の情報も付け加えて出力するなどの機能が必要となる。

## 4. インスタンスキーの導入による同音語問題の解決

### 4.1 インスタンスキー導入の目的

音声を用いたシステムでは、同音語の問題はさまざまな局面で問題を引き起こすことが考えられる。本稿では、特に音声による質問の入力時に発生する曖昧性を解消する目的で、インスタンスキーなる概念を新規に定義・導入しこの問題に実用的な解決を図る。

3章であげた入力時の問題において、結果リレーションが複数のタプルからなる場合に、ユーザとのインタラクションをもち、所望のタプルを同定することを考える。そのためにはタプルを唯一に識別するキーが必要であるが、主キーのような属性値は一般にユーザにとっては記憶しにくいものである。そこで、探索条件を満たす中間結果リレーションを対象としてインスタンスキーの概念により、ユーザとインタラクティブに結果を絞り込むスキームを考える。

### 4.2 インスタンスキーの求め方

#### 4.2.1 インスタンスキーの定義

【定義】 (インスタンスキー)

$R(A_1, A_2, \dots, A_n)$  をリレーションスキーマとし、時刻  $\tau$  における  $R$  のインスタンスを  $R_\tau$  とする。  $K = \{Ak_1, Ak_2, \dots, Ak_p\} (1 \leq k_1 < k_2 < \dots < k_p \leq n)$  を属性の集合とする時、  $K$  が  $R_\tau$  のインスタンスキーと呼ばれるのは、次の条件を満たすときである：

$$(\forall t, t' \in R_\tau)(t[K_\tau] = t'[K_\tau] \Rightarrow t = t')$$

つまり、インスタンスキーは指定された時刻でのみタプルの一意識別能力をもつ属性の組をいう。一般に、インスタンス

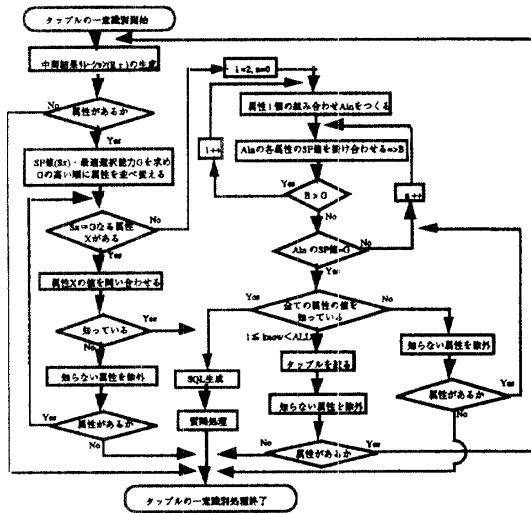
キーは複数あるかもしれない。また、ビューにおいてもインスタンスキーは表と同様に定義できる。

従来のキー（主キー、候補キー）はリレーションスキーマを設計した時点で決定されるが、インスタンスキーはリレーションが変更されるごとに新たなインスタンスキーを導出する必要がある。これがインスタンスキーが従来のキーと決定的に異なる点である。

#### 4. 2. 2 インスタンスキーの算出アルゴリズム

図2にタプルの一意識別処理のフローチャートを示す。本研究でインスタンスキーとは中間結果リレーションというビューにおいてselectivity factor[3]が最適選択能力を持つ属性の組をいう。中間結果リレーションはもとのリレーションから最初の質問文で使用した属性を除外したものである。またselectivity factorとは属性の選択識別能力の高さを示す値であり、最適選択能力Gは「1/リレーションのタプル数」で計算される。

図2 タプルの一意識別処理



#### 4. 2. 3 タプルの一意識別の流れ

実際の処理の様子を以下で説明する。図3にサンプルリレーションを示す（ユーザはWとYの属性について知識があると仮定）。質問文として「名前がこじまの社員の給与は」を与えると、結果が複数となるため中間結果リレーションが作成される（図4）。次にそれぞれの属性についてのselectivity factorとそのリレーションの最適選択能力Gを計算する。例では属性Vが最初のインスタンスキーとなるので、システムはユーザにその値を問い合わせる。仮定よりユーザはその属性値を知らないため新たなインスタンスキーが探索される。その際ユーザの知らない属性Vは探索から除外される（図5）。例では属性1つの組でインスタンスキーはこれ以上存在しないので、属性2つの組で探索が行われる。属性の組(W, X)がインスタンスキーとなるためユーザにその値を問い合わせるが、仮定よりユーザはWの値しか答えられない。よって、知っていたWの値をもとにタプルを絞り込み、知らなかった属性Xを探索から除外する（図6）。こうして得た中間結果リレーションにおいて再びselectivity factorを計算し、インスタンスキーを求める。例では属性Yがインスタンスキーとなり、ユーザはその値を知ってい

ため、先程のWの値とともにSQL文を作成し質問処理を行うと、ユーザの意図するタプルが一意に識別され、「こじまさんの給与は80万円です」という結果が返される。

サンプルリレーション

V	親み姓	W	X	Y	Z	給与
1	こじま	♥	松	●	上	80
2	こじま	♠	松	●	下	30
3	こじま	♥	竹	×	上	70
4	こじま	♠	梅	×	上	30
5	こじま	♠	梅	●	下	50
6	こじま	♠	竹	×	下	60

図3 サンプルリレーション

中間結果リレーション

V	W	X	Y	Z
1	♥	松	●	上
2	♠	松	●	下
3	♥	竹	×	上
4	♠	梅	×	上
5	♠	梅	●	下
6	♠	竹	×	下

最適選択能力 G = 1/6  
各属性の selectivity factor  
S<sub>w</sub> = 1/6  
S<sub>x</sub> = 1/4  
S<sub>y</sub> = 1/3  
S<sub>z</sub> = 1/2

図4 中間結果リレーション

W	X	Y	Z
♥	松	●	上
♠	松	●	下
♥	竹	×	上
♠	梅	×	上
♠	梅	●	下
♠	竹	×	下

図5 次のインスタンスキー

Y	Z
●	上
×	上

最適選択能力 G = 1/2  
各属性の selectivity factor  
S<sub>y</sub> = 1/2  
S<sub>z</sub> = 1

図6 次の中間結果リレーション

#### 5. おわりに

本研究では、音声によるデータベースアクセスのプロトタイプシステムを評価した結果明らかになった、音声入力時における同音語による曖昧性を解消するため、インスタンスキーという新しい概念を定義・導入した。これにより、同音語を含んだ質問文からでもユーザの意図する結果が一意に決定できるようになった。今後、より自然な日本語での入力や電話によるアクセスが行えるよう実装をはかり、システムの実用性を高めていく予定である。

#### 【謝辞】

本研究を遂行するうえで貴重な御意見を下さった様々な方々に深く感謝申し上げます。また、本研究を行うにあたり、ハードウェアやソフトウェアをご提供いただいた（株）オージス総研、日本オラクル（株）、（株）明電舎に心より感謝申し上げます。

なお、本研究は一部平成8年度文部省科学研究費補助金基盤C一般「音声によるデータベースアクセスの研究」（課題番号07680333）の補助を得て行われているものであることを記す。

#### 【文献】

- [1] 荒川ゆう子, 増永良文: "音声によるデータベースアクセスの研究," 情報処理学会研究報告, 96-DBS-109, 1996
- [2] 荒川ゆう子, 増永良文: "リレーショナルデータベースへの音声アクセスインタフェースの構築実験," 情報処理学会第53回全国大会講演論文集(4), 3Q-6, 1996
- [3] P. Griffiths Selinger: "Access Path Selection in a Relational Database Management System," Proceedings of ACM-SIGMOD 1979 International Conference on Management of Data, pp. 23-34, 1979.