

オペレータの検索行動に基づく誘導方式の基本検討

5S-7

永吉 剛 岩瀬 成人
NTT情報通信網研究所

1 はじめに

名義、住所といった複数種類の文字列を検索キーとする大規模データベースは、電話帳や顧客管理のデータベースに利用されている。

このようなデータベース検索システムにおいて、利用者が最初に入力する検索条件と、目標とする解を出力するための必要条件との間には不一致を生じるケースが多い。その場合は一回の検索で解を得ることができない。これに対してオペレータをはじめとする経験の豊富な利用者は、試行錯誤的に入力条件の修正を行い目標とする解に至ることができるが、システムやデータに関する知識の少ない利用者はその修正方法を知ることができない。

本稿では、オペレータの番号情報検索行動を実験により分析し、曖昧な情報にもとづく大規模データベース検索の望ましい誘導方式を提案する。

2 前提条件

以下の章では質問者がオペレータに伝える情報を問い合わせ条件、オペレータからシステムへの入力を検索条件と呼ぶ。また検索条件に対してシステムが出力する検索結果を検索解、質問者の問い合わせ要求を満足する検索解を目標解と呼ぶ。

2.1 実験システム

電話帳データベース検索システムを対象とした。オペレータは端末より複数種類の電話帳データファイル(個人用、企業用、職業用等)の中から一つを選択し、仮名文字列により住所、名義を入力する。名義入力是一定文字数以上は末尾切り捨てして前方一致検索を行う。該当する検索解が既定件数を越える場合、検索解は出力されず「解過多」となる。

2.2 あいまいさの特徴

一般に見られる番号案内の問い合わせは、電話帳に掲載されている住所と名義の正確な指定ではなく、目標とする掲載の住所や名義の一部分であることが多い。また通称名による問い合わせや、日本語表記の曖昧さ、地理的な認識の曖昧さを含むことも多く、

その場合には目標解に無関係な検索解しか出力されない。さらに住所、名義以外の問い合わせ情報からオペレータが検索条件を決定する場合もある。

2.3 実験

11名のオペレータが33種類の問い合わせを受けて検索を行う実験を行った。応対中のディスプレイ表示画面、キーボード操作、及び質問者とオペレータとの音声対話を記録し、これらの分析を行った。

実験では以下の分類による問い合わせを用いた。

a 目標解を得られるケース

a1. 問い合わせ条件の変更なし

検索条件は修正され得るが問い合わせ条件に変更はない場合。例えばデータファイルの選択といった検索条件は、問い合わせでは指定されずオペレータの判断で選択の幅が生じる。

a2. オペレータが問い合わせ条件を変更する

問い合わせ条件を満たす範囲に目標解がないため、オペレータが住所階層の拡大や単語分割できる名義の冠称名削除といった修正を自発的に行って目標解に至る場合。

a3. 質問者が問い合わせ条件を追加・変更する

問い合わせ先に関する質問者の固有の知識が検索条件修正の決め手となる場合。

b 目標解なしのケース

目標解がデータベースに存在しない問い合わせ。

3 実験結果の分析

3.1 応対時間の分布

応対を始めてから検索結果を特定するまでに要する時間を応対時間と呼ぶ。問い合わせ別の平均応対時間は全体の平均応対時間の0.4倍から2.1倍に及び、オペレータ別の平均時間は0.7倍から1.5倍までの分布を示した。応対時間の長い問い合わせでは検索条件の修正回数も多い。また応対で生じる聞き違いやミスタッチは、オペレータ間の応対時間の個人差の大きな要因となる。

3.2 条件修正行動の分析

応対時間の個人差が大きかった問い合わせの一例として、住所情報に変更を要する以下の問い合わせにおける各オペレータの条件修正行動を分析する。

問い合わせ「千代田区外神田あたりでドレミ出版社」
 →目標解：台東区上野3丁目、ドレミ出版
 (外神田と上野3丁目は隣接している)

図1.に各オペレータの条件修正の回数を示す。このうち2名は目標解に達する前に解無しと結論した。また別の2名は1回の条件修正で目標解に達した。

図2.に修正回数の最も多かったオペレータの検索推移を図示した。図の破線で囲まれた領域は検索解を1つ以上出力する条件を表す。この領域内の実線で囲まれた領域が目標解を出力する検索条件の範囲である。このオペレータは名義、住所、データファイルの各入力キーについて交互に、問い合わせ条件を少しずつ緩和した検索条件を入力していった。

4 望ましい検索誘導方式の提案

図1.、2.の分析からわかるように、曖昧な要素を含む問い合わせに対するオペレータの検索行動は、一般に図3.に表すような、問い合わせ条件を原点にとり各検索キーを軸として展開する空間内の探索と見ることができる。

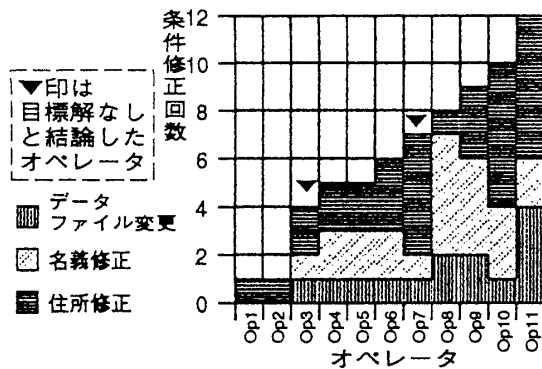


図1. 各オペレータの条件修正回数

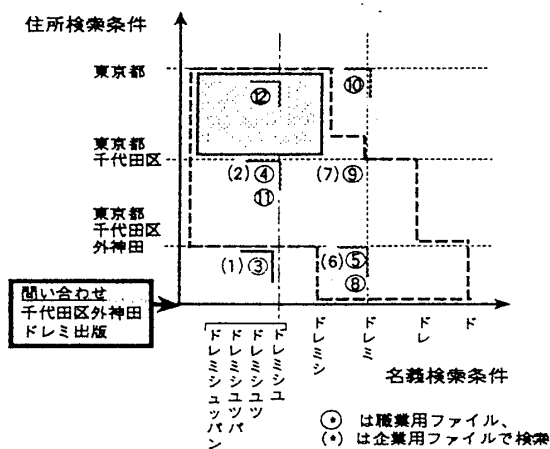


図2. 修正回数の多いオペレータの検索推移

検索推移情報の管理支援

情報検索に”成功する”ためには、目標解の有無を正しく、かつ手早く結論することが要求される。すなわち効率的でもれのない探索を行わなければならない。前記事例で2名のオペレータは、目標解を出力する検索条件の領域を調べていない。また図2.でとりあげたオペレータは11回目の検索で4回目の検索条件と同じ条件を入力している。こうした探索のもれや冗長性を生じる原因の一つとして、検索過程のある時点で未だ確かめていない検索条件と、既に調べ終わった条件との正確な把握が利用者にとって容易でないことがあげられる。このように時間軸上で生起する検索推移情報の管理をシステム側が支援する事によって、正確で効率的な検索が可能となると考えられる。

データ属性に関する知識の提供支援

さらに、より効率的な探索を行う為には、実世界及びデータベース内における問い合わせ対象の統計的分布や意味的関連等といった知識を活用する必要がある。前記事例で、一回の条件修正で目標解に達したオペレータの一人は「外神田が千代田区と台東区の境界に位置する」という地理的知識をあらかじめ持っていた。こうした知識は専門的かつ網羅性を要するものであり、計算機の支援が有効である。

5 おわりに

曖昧な問い合わせに対するオペレータの大規模データベース検索行動について実験による分析を行い、望ましい検索誘導方式について提案した。今後電話帳データベースを対象にこれらの機能を実現していく予定である。

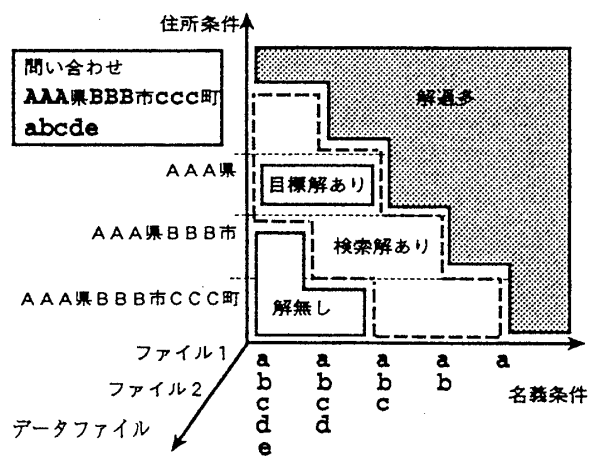


図3. 曖昧な問い合わせに対する修正条件の探索空間