

Yes/No 選択による音声検索インタフェース

1N-06

宮本 勝 植原佳織 稲垣博人

日本電信電話株式会社 NTT サイバーソリューション研究所

1. 背景

音声認識、合成技術の向上により、音声インタフェースを利用した検索サービスが普及してきた。音声による入出力は、ハンズフリー操作を可能にする。このため、車内や歩行中など操作姿勢が不安定な場合のインタフェースとして、音声は有効である。

操作の自由度が高い反面、音声インタフェースには、入出力手段の制約がある。入力に関しては、音声認識率の問題がある。電話を利用する場合には、周囲の騒音、通信路のノイズ、発話自由度を高めたための認識辞書数の増加などが、認識率低下の要因となる。

一方、出力に関しては、音声出力の一覧性がない。ディスプレイにおける 2 次元的な表現とは異なり、音声情報が時間軸に沿って 1 次元的に再生される。情報を選択するためには、ユーザの短期記憶が、一時的に情報を記憶する必要がある。

2. 目的

本研究では、音声インタフェースの入出力の制約を考慮した検索インタフェースの構築を目的とする。そのための入力手段として、予め登録してある検索条件を採用するか否かを、Yes/No で選択していくことで、検索条件を決定する方法をとる。これにより、音声認識は 1bit の情報を認識すればよい。また、検索条件を示す単語だけを電話に発話する恥ずかしさから解放される。

また、出力は、Yes/No 型の質問木と、検索結果から構成される(図 1)。質問木は、質問回数の上限を設けることで、検索条件入力負荷の上限を保証する。また、ユーザの短期記憶が一時的に記憶可能な件数を設定し、ここまでの検索効果を最適化するように、2 分木構造を持つ質問木を生成する。

このように、数ステップの操作による検索効果を最適化することで、付加価値を高めた簡易操作による双方向サービスを実現し、ラジオ等の放送型メディアと差別化したサービスを目指す。

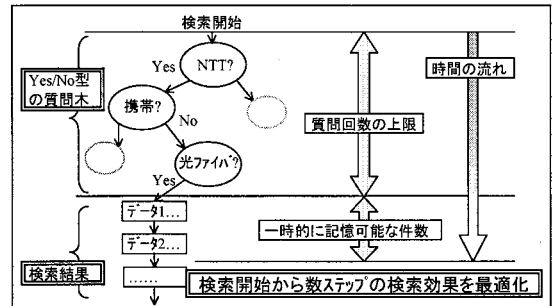


図 1: 対話の流れ

3. 本研究の位置付け

我々は、登録検索条件の組合せの中で、携帯端末でも一覧可能な検索結果数に絞り込める組合せを抽出し、メニューとして提示する方法を提案してきた[1]。本研究では、従来の複数項目型のメニューではなく、Yes/No 型メニューを採用するため、メニュー選択のための短期記憶負荷を軽減できる。また、選択される検索条件とデータとの適合度を、質問木全体として最適化することで、適合度が高いデータを効率良く探せるような質問木を生成する。

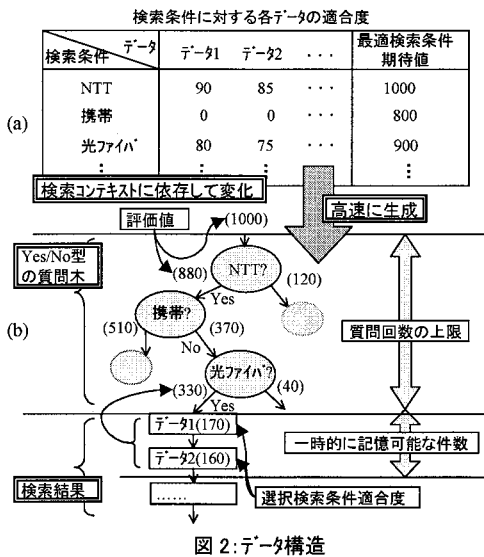
一方、2 分木を平衡化し、データの検索効率高める方法がある[2]。しかし、本研究は、ユーザインタフェースとしての 2 分木に主眼を置いており、必ずしも平衡な 2 分木が最適ではない。

4. Yes/No 質問木の生成方法

4.1. データ構造

本システムは、予め登録された検索条件に対する各データの適合度のマトリックスを入力とする(図 2(a))。このマトリックスに、質問回数の上限と、検索結果を一時的に記憶可能な上限数の制約を加え、最適な質問木を生成する(図 2(b))。この適合度は、検索条件に対する各データの適合する度合いであれば構わず、適合度 0 は、検索条件の検索結果に含まれないことを示す。

また、検索条件とデータの組合せは、ユーザが検索するコンテキストによって、動的に変化する。端末の位置、個人の嗜好、時間帯等に依存して変化するマトリックスに対して、高速に質問木を生成し、双方向サービス可能な方法を構築する。



4.2. アルゴリズム

各データを検索条件の検索結果に含まれるグループと、含まれないグループに分割することを、質問木の根から、再帰的に質問回数の上限の深さまで行う。このとき、質問木を比較評価し、最適解を探索する。しかし、検索条件数と質問回数の上限が増加すると、質問木のパターン数が指数級数的に増大する。全パターンを評価しては、高速に計算できない。

そこで、最適値を導く期待値（以下、最適検索条件期待値）を求め、この値が閾値より高い検索条件だけを評価する。

(1) 評価関数

親の質問のうち、Yes と回答し選択した検索条件の適合度の合計値を、選択した検索条件に対する適合度（以下、選択検索条件適合度）として算出する（例、図 2(b)の「データ 1」→ $90(\text{NTT}) + 80(\text{光ファイバ}) = 170$ ）。

選択検索条件適合度が高いデータを、一時的に記憶可能な検索結果内に多く集めたい。そのために、選択検索条件適合度が上位のデータを、一時的に記憶可能な件数分だけ合計した値を、検索結果の評価値とする（例、図 2(b)の「光ファイバ」の Yes 側検索結果 → $170(\text{データ 1}) + 160(\text{データ 2}) = 330$ ）。質問木全体の評価値合計が、その質問木の評価値となる。

(2) 最適検索条件期待値

分割対象データに対する適合度の各検索条件ごとの合計値を、最適検索条件期待値とする（図 2(a)）。この値が大きくても、選択検索条件適合度が高いデータが、質問木の一部の枝に偏ると、評価値は高くないが、その確率は低いと考えた。

5. 評価

- 検索対象: ニュースのテキストファイル 160 件
- 検索条件: 関心のあるキーワード 20 件
- 適合度: TF/IDF 値を加工した 0 から 99 の適合度
- Yes/No 選択上限数: 3 回、■ 一覧可能上限数: 3 件

上記の条件で、最適値と計算量に対する閾値の影響を評価した（図 3）。閾値がある場合とない場合を評価し、閾値がある場合は、閾値を段階的に厳しくし、計算量を減少させた。その結果、最適検索条件期待値による閾値を用いることで、真の最適値と同等の最適値を維持しつつ、計算量を急激に減少させることを確認した。

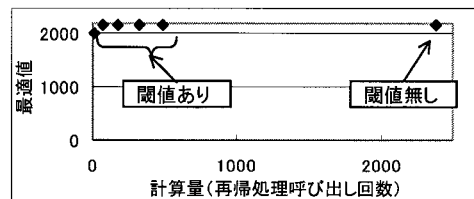


図 3: 最適値と計算量に対する閾値の影響

6. 考察

評価条件の網羅性は、要検討であるが、今回の評価は、本方法が計算量を削減しつつ、最適値の精度を確保可能である可能性を示している。このことは、選択検索条件適合度が高いデータが、一部の枝に偏る確率が低いことを示唆している。

また、本方法は、適合度が高いデータが得られる質問木を生成するため、Web ページのように、適合度の低いデータが多く存在する場合に効果が期待できる。

7. 結論と今後の課題

音声インタフェースの入出力の制約を指摘し、Yes/No の選択による検索インタフェースを提案した。評価の結果、質問木生成処理コストを削減しつつ、最適値を確保可能なことを確認した。

また、評価値の定義により、生成される質問木が変わってくる。今後は、他の評価値の定義を比較検討していく。

謝辞

日頃ご指導戴く NTT サイバーソリューション研究所マルチメディア端末プロジェクト小川克彦プロジェクトマネージャーに感謝します。

参考文献

- [1] 宮本勝ほか (1999) 携帯端末のための操作回数保証型メニューの生成方法, DICOMO'99, pp195-200.
- [2] 岩下博一ほか (1993) 平衡二分木を構成する並列アルゴリズム, 情報処理学会研究報告, AL-32, pp1-8.