

Q&A の自動回答支援システム

大元 聡、林 暁紅、永井 秀利、中村 貞吾、野村 浩郷

九州工業大学情報工学部
〒820-8502 飯塚市川津 680-4
nomura@ai.kyutech.ac.jp
http://www.dumbo.ai.kyutech.ac.jp/nomura/

あらまし パソコン技術サポートを行うメールコールセンターの自動化に関して、本研究では、既存の質問
応答メールデータをカテゴリ化することによる質問応答データベースを構築し、さらにカテゴリごとにカテゴリ
回答文を作成し、未知の問い合わせメールが文書ベクトルを用いて類似度を計算する質問カテゴリ判定シ
ステムによって属するカテゴリに判定され、そのカテゴリのカテゴリ回答文を用いて回答を作成し、作成した
回答を最終査読者に提示する手法とシステムを提案する。提案手法の有効性を確認するため、実際のシ
ステムを作成して実験を行った結果、145 個の質問カテゴリを用いて、86%の割合で正解の質問カテゴリを
上位 3 位以内に推定した。このシステムは、メールコールセンターで実務に供用できる状態にある。
キーワード メールコールセンター 質問応答データベース 文書ベクトル 類似度判定

E-mail Question Automatic Answering Support System

Satoshi Ohmoto, Xiaohong Lin, Hidetoshi Nagai, Teigo Nakamura and Hirosato Nomura
Kyushu Institute of Technology
Iizuka 820-8502 Japan
nomura@ai.kyutech.ac.jp

Abstract The present paper discusses a support system at a mail call center. The mail call center receives inquiries from computer users by e-mails. Then the mail call center replies to them within 24 hours. In order to cope with such a hard task, we propose a system that can perform almost all automatic processing except the final check by a human before sending off the reply to the user. First, we develop a Q&A database by using the reviewed Q&A mails. We categorize these Q&A data depending on their contents. Every category has its own answer respectively. Second, we apply the document vector method so that the system calculates the similarity between the new inquiry and the question categories in the Q&A database. From the answers attached to the found question categories, it generates the candidate new answers that might be adequate to the question. Finally, they are transferred to the reviewer in order to select the most adequate one. According to our examination, by using 145 question categories, we obtained 86% precision about determining the right question category. The system is ready to use at a mail call center.

Keyword mail-call center, Q&A database, document vector, similarity evaluation

1. はじめに

近年インターネットやパソコンの普及により、アフターサービスの一環としてパソコン技術サポートの必要性が高まっている。多くのパソコン技術サポートセンターでは、主に電話で技術サポートを行う従来型のコールセンターに加えて、インターネット経由でE-mailでの問い合わせを受け付けるメールコールセンターがたくさん設置されてきている。

メールコールセンターで行われている技術サポートは、すべて無料サポートである。質問メールは、夕刻から深夜にかけて多く送付されてくる。問い合わせメールの受信から回答の発信までは24時間以内に完了することが求められている。このような制約があるため、企業にとって、正確かつ迅速なサポートを行うには人件費などのコストが膨大なものになりつつある。そこで、メールコールセンターの自動化が強く求められている。

メールコールセンターにおけるサポートの自動化にあたっては、100%に近い精度が要求されるため、現在の技術では完全な自動化は実現が困難である。また、質問者はパソコンの顧客であるため、それへの対応は丁寧、親切かつ正確なものでなければならない。そのため、質問者への回答は、人間による最終査読が不可欠である。

そこで、パソコンユーザーからの質問に対して、システムが少数の回答候補を提示して、それらから人間が最適なものを選択し、最終査読して回答を仕上げるような支援システムを作成し、使用するのが現実的であると思われる。このようなシステムでは、最終査読以外は自動的に処理されることになり、コストの大幅な削減が実現できる。我々は既にそのようなシステム ACCESS (Automated Call-Center Service System) を作成し、実際に使用できる状況にある。

本論文では、実際のメールコールセンターでの約三年間の実務により収集した三万件以上の最終査読済みの「質問応答」データの中から、約一万件を使って「質問応答データベース」を構築した。

実際の質問には、同一内容のものや似た内容のものが多い。したがって、ユーザからの質問が来る度に一々回答を作成するのは無駄である。そのため、既に回答した質問応答データから「質問応答データベース」を構築し、それを再利用できるようにすると、大幅なコストダウンができる。

「ユーザからの質問」には、「質問」とは思われな

いものも含まれている。このような「質問」に対しては、最終査読者が「査読」するのではなく、異なる観点からの対応が必要になるものもある。

ユーザからの質問は自由記述であるので、質問文にはミスタイプ、かな漢字変換の誤り、文法的不完全さなどが多く見られる。このような現象を前提として処理しなければならない。しかし、最終査読済みの質問応答データではそれらは修正・訂正されているので、質問応答データベースを検索して、うまく活用することには大きな利点がある。

質問応答データベースは、二層にカテゴリ化した構造としている。それぞれのカテゴリにはそれぞれを特徴づけるタグを付与している。最下層のカテゴリには、実際の質問応答データが格納されている。二層にカテゴリ化した理由は、最下層でうまくマッチするものがないとき、いわゆるシソーラスにおける上位概念を利用するという考えに似ている。

質問応答データベースの再利用に関しては、パソコンユーザーから送られてきた問い合わせメールから、その質問がどの質問カテゴリに属するのかを統計的処理などにより推定することにより行う。推定された質問カテゴリごとのカテゴリ回答文を用いて作成した回答候補を最終査読者に提示することで、回答作成を効率化することになる。

質問カテゴリ推定の精度を検証するために、システムを作成し、評価実験を行った。145個の質問カテゴリを用いて実験したところ、86%の割合で正解の質問カテゴリを上位3位以内に推定した。蓄積されている三万件以上の「質問応答」データの中から、約一万件を使って「質問応答データベース」を構築してみたため、データ数は十分ではないことが確認され、残りの約二万件の「質問応答」データを「質問応答データベース」に加えることにより、正解の質問カテゴリを発見する精度は大幅に改善できると思われる。

なお、本稿では、著作権や知的所有権等の制約の可能性などにより、データや処理の具体的な内容は概略のみを示すに留めることがある。

2. 質問応答データベース

2.1 質問応答データベースの構築手法

質問応答メールデータとは、パソコンユーザーから送られてきた問い合わせメールとそれに対する

査読済み回答文のペアのことである。

質問応答メールデータの中には、同一データないしは類似データが多数存在する。したがって、問い合わせメールの内容または意味が同一または類似で、それらの回答文の文章表現も同じまたは類似である場合、それらを「類似データ」とみなす。

メールコールセンターの質問応答データベースを構築するにあたって、実際には、10,135 件の質問応答メールデータを使用した。これらに対して、「質問カテゴリ」を作成し、類似データの「質問カテゴリ分類」を行った。質問カテゴリは、上に述べたように二層構造にし、上位層をブランチカテゴリ、下位層をリーフカテゴリと呼ぶ。類似データは同一リーフカテゴリに分類し、さらに相関関係があるリーフカテゴリは同一ブランチカテゴリに分類する。

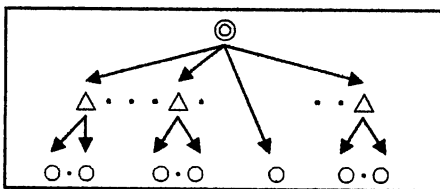
2.2 質問応答データベースの構成

構築したメールコールセンターの質問応答データベースは二段階（二層）のツリー構造である。リーフカテゴリはブランチカテゴリに属する場合もあるし、直接ルートカテゴリに属する場合もある。ルートカテゴリはブランチカテゴリの上位カテゴリであるが、ツリー構造のルートノードであるので、「層」とはみなさない。すなわち、全体を三層構造とは呼ばないことにしている。

ブランチカテゴリはデータを持たず、リーフカテゴリは同一データないしは類似データを持つ。ルートと各カテゴリの相関関係は下記の通りである：

- ・ ルート → ブランチカテゴリ → リーフカテゴリ
- ・ ルート → リーフカテゴリ

これを図示すると概略次のようになる。



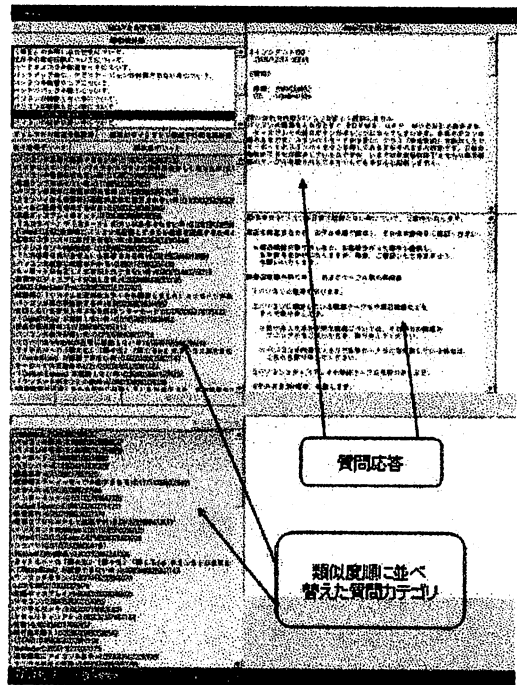
構築したメールコールセンターの質問応答データベースでは、10,135 件の質問応答メールデータの内、利用対象外データ3,598 件を除き、計6,537 件に対して、83 個のブランチカテゴリおよび634 個のリーフカテゴリが設定された。利用対象外データとは、いわゆるすなおな形・内容のものではなかったものな

どであり、再利用にはむかない、あるいは現在の研究の守備範囲を逸脱しているものなどである。

2.3 質問応答データベース構築支援システム

質問応答データベースの構築には、多くの工数を要する。したがって、当初は、1,705 の質問応答データについて人手で質問応答データベースを構築した。これは、どのようなカテゴリを設定すべきかの事前検討のためでもある。

質問応答データベース構築の効率をあげるため、その後、質問応答データベース構築支援システムを作成して活用した。上に述べた 6,537 件の質問応答データは、この質問応答データベース構築支援システムを使用して構築したものである。そのスクリーンショットを下に示す。



2.4 カテゴリ回答文作成への応用

リーフカテゴリに分類された同一データないしは類似データの回答文は同じまたは類似であるため、リーフカテゴリに属する質問メールに対して、共通回答文であるリーフカテゴリ回答文を作成する。未知の問い合わせメールの回答文を作成する際、その質問メールの属するリーフカテゴリのカテゴリ回答文を用いて、回答文の作成を行う。

リーフカテゴリ回答文には、定型回答文とテン

プレート回答文の2種類を用意する。定型回答文は機種関連情報などを含まない場合のためのものであり、回答文を作成する際にそのまま出力する。一方、テンプレート回答文は、ハードウェアやソフトウェアなどの多種類の機種関連情報に関するスロットが用意されており、それらの機種関連情報を機種関連情報データベースから抽出し、スロットに入れ、回答文を作成し出力する。

3. 特徴ベクトル

パソコンユーザーから送られてきた問い合わせメールの特徴を正確に捉えるために、以下の特徴ベクトルを複合的に用いる。

- ・TF / IDF による特徴ベクトル
- ・体言と用言の共起を考慮した特徴ベクトル
- ・文タイプを考慮した特徴ベクトル

3.1 TF / IDF による特徴ベクトル

システムでは、問い合わせメール中に出現する語のTF / IDFによる重みを要素とした文書ベクトルを拡張したもので質問文を表現する。TF / IDFはよく使われるものであるため、ここでは説明を省略する。

3.2 体言と用言の共起を考慮した特徴ベクトル

TF / IDFによる重み付けは、通常、ある語が特定の文書の特徴付ける尺度を表現するものであり、文の構造を反映しない。したがって、

- ・「電源を切る。」
- ・「電源を入れる。」

という二つの文に対して、「電源」という語は同じ重みが与えられる。だが実際には、目的とする質問文のカテゴリ判定においては、この二つは違う特徴を持つものとして認識すべきである。これは、語の出現頻度だけを考えていると、とらえにくい特徴である。そこで、TF / IDFによる重み付けに加えて、体言に対する用言の一文での共起の割合を重みとして用いることを考える。それぞれの体言について、一文中で共起した用言の頻度を要素とする特徴ベクトルを用いる。

3.3 文タイプを考慮した特徴ベクトル

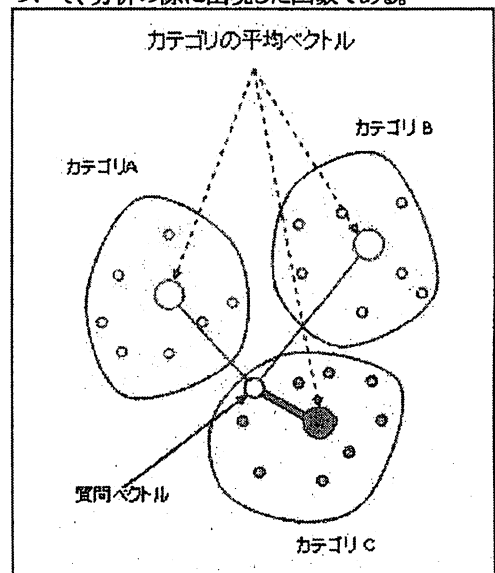
パソコンユーザーから送られてきた問い合わせメールの内容をより正確に反映した特徴ベクトルを作成するため、文中の語がどのような意味の文に出現

するのか、という傾向について考える。そのために、まず問い合わせメールを分析してそれぞれの文タイプごとの特徴を調べ、分析結果をもとに文タイプ同定のルールを作成する。

文タイプは、我々の他の研究でも従来から使ってきた。情報抽出システム NIX、文書要約システム TESS、情報検索システム CODIFY、情報集約システム DOCS、などである。ここでは、質問メールを分析することにより、次のように少数の文タイプを設定した。

・Question	324
・Problem	648
・Intention	87
・Situation	398
・Think	37
・Other case	80
・About	368
・Message	96
・etc	34

質問について述べてある文や、パソコンの不具合・問題について述べてある文など、上記9種類の文タイプを設定し、約一週間分の問い合わせメール、323件2,072文を分析して、文末表現や機能語から、文タイプを同定するルールを作成し、各文タイプ中での各単語の頻度を要素とする特徴ベクトルを用いる。上の表の右端の数字は、参考までに記したものであり、それぞれの文タイプについて、分析の際に出現した回数である。



3.4 カテゴリの平均ベクトルを用いた類似度計算

未知の問い合わせメールがどの質問カテゴリに属するのかを計算するのに、各質問カテゴリ内の質問文の特徴ベクトルを平均化したものを便宜的に質問カテゴリのベクトルとする。その概念図を上に表示す。

これらに対して、未知の問い合わせメールのベクトルとの類似度を計算する。そして、質問ベクトルと各質問カテゴリとの距離を計算し、最も近い質問カテゴリをもつカテゴリに質問が属すると判断する。現在は、計算量を少なくするため、このような簡易な類似度計算方式を暫定的に使っている。

3.5 文タイプ同定ルール

文タイプを同定するルールによって、文中の語が出現する文タイプの傾向を求めて文タイプベクトルとし、文タイプベクトルが類似している場合には文章間の類似度が高くなるように、文書ベクトルの類似度を定義する。

上述した TF / IDF による重みを w 、共起ベクトルを \vec{c} 、文タイプベクトルを \vec{i} として、以下の式で文書ベクトル \vec{T} と \vec{T}' の類似度を定義する。

$$\vec{T} \cdot \vec{T}' = \sum_{i=1}^n \left\{ \omega_i \cdot \omega'_i (\vec{c}_i \cdot \vec{c}'_i) (\vec{t}_i \cdot \vec{t}'_i) \right\}$$

共起ベクトルについては、質問カテゴリ内の質問文の特徴ベクトルを平均化することで、少数のノイズを取り除き、質問カテゴリ内で真に特徴的な語の情報のみを残すことができると考える。

4. システム

システムは、ユーザから質問メールを受け取ると、自動的に処理を始める。処理の結果、すなわち、回答候補は質問者への返答メールの形に整形されて最終査読者の査読を待つ。

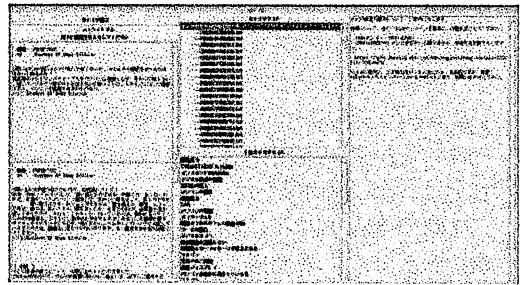
質問文の解析は、形態素解析のみを行い、その結果を言語データベースおよび知識データベースに照らして、質問応答データベース検索の準備をし、自動的に検索を実施する。言語データベースは、言語的素性を持つ辞書のようなものであり、知識データベースはパソコンに関する事典のようなものである。

質問文の解析として、形態素解析のみを活用する

理由は、依存構造解析などの信頼性に起因するものである。

最終査読者の査読済み質問応答データは、質問応答データベースの更新に供される。質問応答データベース自動構築システムは、現在、作成中である。

システム画面の例を下に示す。いくつかの操作機能が用意されており、最終査読者が効率的な査読を行えるよう配慮されている。最終査読者が三つの回答候補のいずれかが妥当であると判断すると、それが質問者に自動的にメール返信される。微細な加工が必要な場合には、この画面の上で加工し、それが質問者に回答メールとして自動的に返信される。システム画面の説明は省略する。



5. 実験

質問カテゴリの判定の精度を評価するために、問い合わせメールに対して、正解質問カテゴリをどれだけ上位に判定するかを調べた。

リーフカテゴリについての結果を下に示す。

一位に判定	三位以内に判定
2,505 件 (62.3%)	3,459 件 (86.0%)

この結果により、最終査読者が最良と判定する回答返信メールが類似度順位三位以内にある確率は、この実験においては、86.0%であったことがわかる。

ブランチカテゴリについての結果を下に示す。

上位カテゴリを一位に判定	上位カテゴリを三位以内に判定
3,202 件 (79.6%)	3,805 件 (94.6%)

この結果により、最終査読者が最良と判定する上位カテゴリは、94.6%の精度で類似度上位三位以内に入っていることがわかる。

システムの検索結果は、上記の三つのみではなく、実際には、類似度の高い順に多くの候補が求められている。最終査読者は、必要と思えば、四位以下の候補もスクリーン上で見れるようになっている。実際には、特別な場合でも、五位くらいまでチェックすれば十分である。

なお、この実験は、質問応答データベース構築には使わなかった質問メールを使って行ったものであり、オープンテストである。

6. おわりに

本論文は、メールコールセンターでのパソコン技術サポートを取り上げ、パソコンユーザーから送られてきた問い合わせメールに対し、作成した質問応答データベースにあるどの質問カテゴリに属するのかを統計的な手法を用いて判定し、その結果を用いて回答文を作成し、回答者に提示する回答作成を支援するシステムを作成した。

実験の結果、86%の精度で類似度三位以内に正解の質問カテゴリを特定でき、システムの有効性を確認した。

したがって、本システムは、そのままメールコールセンターで実際に使用可能である。

今回は、手持ちの約三万件強の質問応答データに対してその約1/3を使って質問応答データベースを作成したが、実験によりデータ不足が確認できた。残りの2/3、すなわち、約二万件強の質問応答データをも使い質問応答データベースを拡充すると、精度はかなり改善されると思われる。

類似度計算は、本稿で述べたものの他に、k nearest neighbors (KNN) 法などについても実験してみたが、精度の改善はみられなかった。しかし、本稿の類似度計算法は簡易すぎると思われるので、今後は、精度の高い結果が得られる効率的な類似度計算法を見つけ出す必要があるかとも思われる。しかし、このとき、計算量と精度のトレードオフを考える必要がある。

本システムの名前は、ACCESS-mail である。この他に、現在、ACCESS-voice も開発中である。これについては、別の機会に述べる予定である。

参考文献

- 1] H. Nomura, Mail/Voice Call Center Answering and Information Obtaining/Tidying, Proc. Of IPSI-2006 FRANCE, CD-ROM 10 pages, 2006
- 2] H. Nomura, M. Itoh, H. Nagai, & T. Nakamura: Information Extraction on Non-presupposed but Interesting Items in Addition to Presupposed Items, Proc. of IPSI-2005 SLOVENIA, CD-ROM, 6 pages, 2005
- 3] H. Nomura: Information Retrieval and Integration of Relevant News Articles, Proc. of International Symposium on Machine Translation, NLP and Translation Support Systems, pp4-9, 2004
- 4] H. Nomura, H. Koga, H. Nagai, & T. Nakamura: Text Summarization Based on Linguistic Function, Conceptual Relationships and Partial Contextual Constraints, Proc. of IPSI VENICE 2004, CD-ROM, 6 pages, 2004
- 5] H. Nomura, S. Ohmoto, H. Nagai, & T. Nakamura: Information Extraction on Presupposed Items and/or Presupposed Topics from Text Documents, Proc. of IPSI-2004 STOCKHOLM, CD-ROM, 6 pages, 2004
- 6] H. Nomura, J. Akamatsu, H. Nagai, & T. Nakamura: Information Extraction on Presupposed Items from News Articles, Proc. of The Fifth International Conference on Electronic Measurement & Instruments, pp626-630, 2001
- 7] T. Nakamura, T. Matsubara, & H. Nagai, H. Nomura: A Method for Evaluating Uncertain Requirements in Travel Consultation Dialogues, Proc. of Third Natural Language Processing Pacific Rim Symposium, pp290-295, 1995
- 8] H. Nagai, T. Nakamura, & H. Nomura: Skeleton Structure Acquisition of Japanese Law Sentences based on Linguistic Characteristics, Proc. of Third Natural Language Processing Pacific Rim Symposium, pp143-148, 1995
- 9] H. Nomura: Information Extraction and Generation on Information Highway, Proc. of the 2nd Symposium on Natural Language Processing, pp7-14, 1995
- 10] S. Misumi, H. Nagai, T. Nakamura, & H. Nomura: Algorithmic Information Extraction from Natural Language Document, Proc. of Second Natural Language Processing Pacific Rim Symposium, pp417-420, 1993

(参考文献としては、我々の論文のみを掲げた)