

A Study on Information Retrieval with Vagueness Requirement

(あいまい要求を考慮した情報検索に関する研究)

学位論文内容の要旨

本論文は、情報システム上に存在する“知識資源”から必要な情報を抽出するために、ファジィ関係不等式を用いることで自然言語によるユーザの入力を適切に処理し、かつ、ユーザとのインタラクションを繰り返すことによって適合性を高める情報検索システムを提案するものである。

近年の情報技術の発展と普及により膨大かつ多種多様な情報にアクセスすることが可能となっている。従来、情報検索の領域では、ディレクトリ、全文データベース、ヒューリスティックを利用したパターンマッチング等により情報を抽出する研究がなされてきた。しかし、現状のシステムは十分にユーザの意図を反映した検索結果を示すことができているとはいえない。ユーザの意図をくみとるためにはあいまいな自然言語の処理、およびシステムとユーザのインタラクションを必要とする。そこで本論文では、ファジィ関係とユーザとのインタラクションを用いた情報検索システムを提案する。

情報検索は、多くの場合自然言語によって行われ、その要求は常にあいまいさを含むため、提案システムではファジィ関係を用いることでこの問題を処理する。自然言語は文脈に依存して様々な意味を持つと共に、その意味する内容は常に揺れ動いているため、語釈の明確な境界を予め定義することが困難である。また、あいまいな情報間の入出力関係、因果関係を代数的に表現するファジィ関係式は、ファジィシステムの一つの数学的モデルとして重要であり、同定問題、診断問題、推論問題などを中心に幅広く応用されている。ファジィ関係は通常の数学で扱う関係の拡張であり、自然言語に含まれる二つの集合間のあいまいな関係を定義することができる。一方、「結果は少なくともある最低限の制約を満たさなければならない」などの制約付関係式は関係不等式として表現され、この関係不等式も関係式同様重要である。本論文では、max-min 合成演算を対象としたファジィ関係不等式の解法について論じている。

さらに、ユーザからの要求が様々に解釈されるのであれば、システムはその解釈の正当性をユーザに確認する必要がある。ユーザは最初の要求においては自由に検索語を記述できるため、システムからは意図を推定するために必要な情報の不足から適合する文書集合が大きすぎる場合や、過度に詳細な指定によって適合する文書集合が空になる場合が存在

すると考えられる。これを回避するためには、システム側からユーザへ要求の意図を推定するのにふさわしい入力を求める必要がある。本論文では、このようにシステムとユーザとのインタラクションを繰り返すことでユーザの意図に適合した検索結果を提示する手法について述べている。

最後に実験およびその結果から、Web 上の情報を対象とした情報検索に対して、提案する情報検索システムが有効な方法論であるということを示している。

本論文は、6章から構成されている。以下にその概要を示す。

第1章は序論であり、本研究の背景、現状の問題点、研究目的について述べている。はじめに、情報検索におけるアプローチ手法について分類しその特徴をまとめ、比較することで従来の手法の問題点を明らかにしている。次に、ファジィ理論に基づく情報検索手法の特徴について述べ、さらに、本論文で提案する情報検索手法について概念を述べ、その特徴をまとめている。

第2章では、情報検索において用いられる自然言語による情報アクセス技術についてとりあげている。ここでは、情報を整理して提示する情報抽出技術、データベースを対象と検索技術および Web 上の情報を対象とした検索技術、テキストデータを要約して提示するテキスト要約技術、そして質問応答技術について述べている。ここでは、各技術に関してその背景や性質などについて述べ、さらに問題点についてまとめている。

第3章では、5章で実験を行う情報検索システムについてその詳細な特性について論じている。情報検索システムは、ユーザからのあいまい要求をファジィ関係を用い、クリस्प値で表現可能な評価項目に置換することによって検索対象の評価を行っている。また、ユーザとのインタラクションを行うことによりユーザの意図をより明確にしている。

第4章では、提案システムで用いるファジィ関係不等式と、その一解法について述べている。この解法は、演算子を用いることによって転置行列を利用した通常の代数方程式に近い解法で解を求めることができるうえに、アルゴリズムも簡単で分かりやすいものとなっているという特徴を持っている。

第5章では、提案システムの有効性を Web 上に存在する情報を対象とした実験を行い、検証している。この実験において、ユーザの意図をくみとり、検索結果に反映させることが困難であるという問題に対してファジィ関係不等式の導入することによりユーザのあいまい要求を考慮することが可能であることが明らかになった。また、システムとユーザのインタラクションを行い、それを再検索する際の情報として用いることでよりユーザの意図に近い検索結果を提示することができた。

第6章において、論文全体のまとめと総括を行っている。

学位論文審査の要旨

主査	教授	嘉数	侑昇
副査	教授	大内	東
副査	教授	和田	充雄
副査	教授	大森	隆司
副査	助教授	吉井	伸一郎

学位論文題名

A Study on Information Retrieval with Vagueness Requirement

(あいまい要求を考慮した情報検索に関する研究)

本論文は、情報システム上に存在する“知識資源”から必要な情報を抽出するために、ファジィ関係不等式を用いることで自然言語によるユーザの入力を適切に処理し、かつ、ユーザとのインタラクションを繰り返すことによって適合性を高める情報検索システムに関する研究成果をまとめたものである。

近年の情報技術の発展と普及により膨大かつ多種多様な情報にアクセスすることが可能となっている。従来、情報検索の領域では、ディレクトリ、全文データベース、ヒューリスティックを利用したパターンマッチング等により情報を抽出する研究がなされてきた。しかし、現状のシステムは十分にユーザの意図を反映した検索結果を示すことができているとはいえない。特にユーザにあらかじめ明確に決定された検索項目が無いような多くの場合、ユーザの意図をくみとるためにはあいまいな自然言語の処理、およびシステムとユーザのインタラクションを必要とする場面が多々ある。これらを特に意識して本論文は、ファジィ関係とユーザとのインタラクションを用いた情報検索システムを提案している。

情報検索は、多くの場合自然言語によって行われ、その要求は常にあいまいさを含むため、提案システムではファジィ関係を用いることでこの問題を処理する。自然言語は文脈に依存して様々な意味を持つと共に、その意味する内容は常に揺れ動いているため、語釈の明確な境界を予め定義することが困難である。また、あいまいな情報間の入出力関係、因果関係を代数的に表現するファジィ関係式は、ファジィシステムの一つの数学的モデルとして重要であり、同定問題、診断問題、推論問題などを中心に幅広く応用されている。ファジィ関係は通常の数学で扱う関係の拡張であり、自然言語に含まれる二つの集合間のあいまいな関係を定義することができる。一方、「結果は少なくともある最低限の制約を満たさなければならない」などの制約付関係式は関係不等式として表現され、この関係不等式も関係式同様重要である。本論文では、max-min 合成演算を対象としたファジィ関係不等式の解法について論じている。

さらに、ユーザからの要求が様々に解釈されるのであれば、システムはその解釈の正当性をユーザに確認する必要がある。ユーザは最初の要求においては自由に検索語を記述できるため、システムからは意図を推定するために必要な情報の不足から適合する文書集合が大きすぎる場合や、過度に詳細な指定によって適合する文書集合が空になる場合が存在すると考えられる。これを回避するためには、システム側からユーザへ要求の意図を推定するのにふさわしい入力を求める必要がある。本論文では、このようにシステムとユーザとのインタラクションを繰り返すことでユーザの意図に適合した検索結果を提示する手法について述べている。

最後に実験およびその結果から、Web上の情報を対象とした情報検索に対して、提案する情報検索システムが有効な方法論であるということを示している。

本論文は、7章から構成されている。以下にその概要を示す。

第1章は序論であり、本研究の背景、現状の問題点、研究目的について述べている。

第2章では、はじめに、情報検索におけるアプローチ手法について分類しその特徴をまとめ、比較することで従来の手法の問題点を明らかにしている。次に、ファジィ理論に基づく情報検索手法の特徴について述べ、さらに、本論文で提案する情報検索手法について概念を述べ、その特徴をまとめている。

第3章では、一般的な情報検索の定式化を行っている。この定式化により提案手法の関連研究との位置づけを明確にしている。多くの場合、一度の要求ではユーザの求める情報が獲得できるとは限らない。そのため、システムとユーザによる複数のインタラクションを考慮する必要がある。そこで、本論文では情報検索をユーザとシステムが複数回 Query/Reply を繰り返す対話として定式化している。

第4章では、提案する情報検索システムに用いるファジィ理論についてとりあげている。このファジィ理論は、知識工学において幅広く研究されたアプローチであり、人間の用いる自然言語が含有するあいまいさを扱うのに適している。本論文では、このファジィ理論を用いることによって、情報検索におけるユーザの要求のあいまいさという問題を処理している。

第5章では、提案する情報検索システムについてその詳細な特性について論じている。情報検索システムは、ユーザからのあいまい要求をファジィ関係を用い、クリस्प値で表現可能な評価項目に置換することによって検索対象の評価を行っている。また、提案システムで用いるファジィ関係不等式と、その一解法について述べている。この解法は、演算子を用いることによって転置行列を利用した通常の代数方程式に近い解法で解を求めることができる。また、アルゴリズムも簡単で分かりやすいものとなっているという特徴を持っている。また、提案システムの有効性をWeb上に存在する情報を対象とした実験を行い、検証している。この実験において、ユーザの意図をくみとり、検索結果に反映させることが困難であるという問題に対してファジィ関係不等式の導入することによりユーザのあいまい要求を考慮することが可能であることが明らかにしている。

第6章では、Web上に存在する情報は、近年のネットワークインフラの普及によって様々な種類かつ膨大な量のデータベースとみなすことが可能である。その中でも特に、画像データに注目し、ユーザの与える質問キーワードを基にそのキーワードに関連するイメージを出力するシステムについて述べている。

第7章において、論文全体のまとめと総括を行っている。

これを要するに、本論文は、複数回の対話をも含む情報検索問題の定式化を行い、それに基づいて情報検索を行う際のユーザからの要求が含有するあいまいさに対して、ファジィ関係を用い、またシステムとユーザがインタラクションを行うことによってユーザの意図に適合した検索結果を提示可能とする手法を与えたものであり、Web検索、特に画像データの検索に関する実験を通して提案手法の有効性を評価している。これらの結果得られた成果は複

雑系工学及び情報工学の分野に寄与するところ大なるものがある。よって著者は、北海道大学博士(工学)の学位を授与される資格があるものと認める。