

(様式2)

学位論文の概要及び要旨

氏 名 徳久雅人 印

題 目 パターン言語処理に基づく情緒推定に関する研究

学位論文の概要及び要旨

概要

本論文は、計算機による言語意味理解のための基礎技術の一つとして、言語表現から話者や登場人物の情緒を推定する方式を提案する。そして、情緒推定に必要な言語知識ベースを構築することによって、提案した方式の精度を確認する。従来の意味処理では、表現に結び付けられた話者の認識に迫るような意味理解の研究はまだ少ない。その理由として、意味理解で必要と考えられる膨大な知識ベースについて、その具体的内容と構成法が明確になっておらず、構築が進んでいないことが挙げられる。そこで、以下の手順により、言語意味理解の情緒的側面に対する問題の解決を目指す。

- (1) 「情緒主」、「情緒の生起/反応の過程」、「言語の解釈の種類」に着目して、従来の情緒推定の方式を分類して問題点を明らかにし、本研究の狙いと方式を示す。
 - (2) 情緒処理の精度上の目標と評価基準を明らかにするために、言語表現に対する人間の様々な情緒推定能力を実験的に評価し、本研究の目標と評価方法を定める。
 - (3) 情緒処理の新しい方式として、「非線形言語モデル」に基づく「パターン言語処理」を応用した方式を提案する。この方式は、あらかじめパターンと情緒の関係を対応付けた知識ベースを構築しておき、実行時に、入力文と適合したパターンから、入力文の情緒を推定する方式である。そこで、新方式の実現に必要な知識ベースを設計し、パターン記述言語とパターン照合アルゴリズムを開発する。
 - (4) 前項の設計に従って、情緒推定のための知識ベースを構築し、それを用いた情緒推定方式の精度を実験的に評価する。そこで、まず、パターンと情緒の関係を判断する基準を明確にするために、情緒が生起する「原因」を分析する。次に、「日本語語彙大系」の文型パターン(1.4万件)に対してその情緒の対応付けを行う。その後、得られた知識ベースを用いて情緒推定の実験を行う。
- 以上により、言語表現から話者や登場人物の情緒を推定するための新しい方式とその実現に必要な言語知識ベースの構成法を提案するとともに、その実現性と効果を確認する。

要旨

本論文は、計算機による言語意味理解のための基礎技術の一つとして、言語表現から話者や登場人物の情緒を推定する方式を提案するものである。情緒推定に必要な言語知識ベースを構築することによって、提案した方式の精度を確認している。

最近、意味処理では、単語や表現の持つ意味を対象とした意味解析の研究が盛んであるが、表現に結び付けられた話者の認識に迫るような意味理解の研究はまだ少ない。その理由として、意味理解で必要と考えられる膨大な知識ベースについて、まだ、その具体的内容と構成法が明確になっておらず、構築が進んでいないことが挙げられる。

本論文では、言語表現から話者や登場人物の情緒を推定するための新しい方式とその実現に必要な言語知識ベースの構成法を提案すると共に、その実現性と効果を確認した。本研究の成果は、以下の4点にまとめることができる。

- (1) 「情緒主」、「情緒の生起/反応の過程」、「言語の解釈の種類」に着目して、従来の情緒推定方式を5つのタイプに分類した。「情緒状態」を表す言語表現からの情緒推定は、方式と網羅的な言語知識ベースが存在するのだが、「情緒生起の原因」に基づく情緒推定には、日本語の基本的な用言を網羅し、かつ、情緒主や情緒名などの情報を備えるような言語知識ベースは開発されておらず、表層的な言語情報に基づく方式が整備できていないことが分かった。
- (2) 情緒処理の精度上の目標と評価基準を明らかにするため、言語表現に対する人間の様々な情緒推定能力を実験的に評価した。人々に共通的に読み取られる情緒を、情緒推定の正解とした。日記文やテキスト対話において、2者間の情緒推定の一致率が52%~65%であることが分かった。
- (3) 情緒処理の新しい方式として、「非線形言語モデル」に基づく「パターン言語処理」を応用した方式を提案した。この方式は、あらかじめ表現パターンと情緒の関係を対応付けた知識ベースを構築しておき、実行時に入力文と適合したパターンから入力文の情緒を推定する方式である。そこで、新方式の実現に必要な知識ベースを設計し、パターン記述言語とパターン照合アルゴリズムを開発した。ATNを高速化したアルゴリズムとすることで、実効的には $O(n) \sim O(n^2)$ の計算量で、約20万件のパターンとの照合が可能となった。
- (4) 前項の設計に従って、情緒推定のための知識ベースを構築し、それを用いた情緒推定方式の精度を実験的に評価した。新方式において、パターンは言語表現で表された概念を掬い取る網である。そこで、まず、概念と情緒の関係を判断するための基準を明確にするために、情緒が生じる「原因」を分析したところ、「喜び」、「怒り」などの8種類の基本的な情緒に対して、約120種類の原因を見いだした。その後、「日本語語彙大系」の文型パターン(1.4万件)に対して情緒の原因を対応付けてみると、約52%のパターンが情緒の原因と対応することが分かった。そして、得られた知識ベースを用いて新方式の情緒推定能力を評価した。情緒推定の初心者には正解率が49%であったところ、本知識ベースを手動で運用した場合は48%、自動で運用した場合は45%であった。また、誤り分析によると、本知識ベースの不備による誤りは3%であった。

以上の結果、新方式では、通常の人に匹敵する情緒推定能力が実現できることが分かった。言語表現から情緒を読み取る能力は、人間でも個人の感性に依存する部分が存在する。また、通常の対話では、情緒は、表情、身振りなど言語外の表現で表される部分も大きく、必ずしも言語情報だけで推定できるとは限らない。本研究の成果は、Web上の記事やブログなどに現れた話者の感情や意志の自動分析など、多くの応用が期待される。