

(論文内容の要旨)

本論文は、異なる文化に属する人たちが与えた絵文字解釈のデータから文化差を解析する方法について研究した成果をまとめたものであり、6章から構成されている。

第1章は序論であり、本研究の背景と目的について概観している。

第2章では、本研究で用いた日米絵文字解釈データの特徴を述べている。

本研究ではインターネットを用いたウェブサーベイを日本と米国で実施し、各絵文字の日米での解釈を自由回答形式で収集している。そして収集した絵文字解釈を分析し、一つの絵文字に複数の解釈が与えられたり、複数の絵文字に同じ解釈が与えられたりする現象から、絵文字に多義性や同義語があり、意味的あいまい性が存在することを報告している。また、一部の絵文字において日米の解釈が大きく異なることを指摘し、文化的あいまい性が存在することを報告している。

第3章では、絵文字解釈の意味的あいまい性を扱う方法として、ある解釈語と、一つの絵文字に付与された絵文字解釈語集合の間の意味的関連度 (Semantic Relevance Measure) を計算する方法を提案している。

意味的関連度の計算には、絵文字に付与された各解釈語の確率に基づき、絵文字集合全体における二つの解釈語の類似度を計算している。また、意味的関連度の応用として、与えられた検索語にふさわしい絵文字を順位付けて提示する方法を提案している。しかし、意味的関連度の計算では絵文字を用いる際の様々な視点を区別なく扱うため、意味的関連度の数値が平均化され、検索における意味的関連度の閾値を適切に設定することが困難であった。そこでEDR電子化辞書の上位概念体系を借用し、絵文字解釈語を事前にAGENT, MATTER, EVENT, LOCATION, TIMEの五つのカテゴリに分類し、カテゴリごとに意味的関連度を計算し、特定のカテゴリで関連度が高ければ検索結果として提示する方法を提案している。

事前にカテゴリ分類された解釈語データと、そうでない解釈語データを基にした絵文字検索性能を比較した結果、事前分類された方が、そうでないものより、幅広い閾値に対し、安定した高い検索性能 (F値は0.7) を示すことを確認している。

第4章では、人間が絵文字解釈のどの部分に注目し文化差を読み取るかを、被験者を用いた文化差判定実験で明らかにしている。その方法として、日本人と米国人からなる被験者に、日本と米国から集めた絵文字解釈を対比させて提示し、日米の絵文字解釈の文化差を判定させ、事後インタビューを通して文化差判定基準を調べている。インタビュー結果から、被験者が次の基準を用いて、二文化の絵文字解釈の文化差を判定していると結論している。

- (1) 二文化に類似する解釈が少ない、もしくは異なる解釈が多い、
- (2) 二文化で類似する解釈が多いが、解釈の割合が異なる、
- (3) 一方の文化では解釈が意味的にばらついているが、もう一方では意味的にまとまっている、

- (4) 国名や市名などの固有名詞が含まれている,
- (5) 一方の文化では肯定的な解釈が多いが, もう一方では否定的な解釈が多い.

第5章では, 人間の文化差判定基準を定式化し, 二文化の絵文字解釈の文化差を自動的に検出する方法を提案している. そのために, 被験者実験で明らかになった人間の文化差判定基準を,

- (1) 二文化の絵文字解釈の共通性,
- (2) 二文化の共通解釈の比率の違い,
- (3) 二文化の絵文字解釈のばらつき度合いの違い

を判定する式として定義し, 三つの判定式を論理積で結合することで文化差判定式を定義している. この文化差判定式は, 三つの判定基準を全て満たした絵文字解釈を文化差のあるものと判定する. 各判定式は, 対訳辞書やシソーラスを用いて二言語の絵文字解釈を対応付けた後, 第3章で定義した意味的関連度の計算式を適用して得られたものである.

この文化差判定式の性能を, 日米絵文字解釈データを用いて調べた結果, 高い文化差検出性能 (F値は0.7) が示されている.

第6章は結論で, 本論文で得られた成果を要約している.

(論文審査の結果の要旨)

本論文は、異なる文化に属する人たちが与えた絵文字解釈データから文化差を解析する方法について研究した結果をまとめたものであり、得られた主な成果は次のとおりである。

1. 絵文字解釈の意味的あいまい性を扱う方法を提案した。一つの絵文字が多様な解釈を持ち、複数の絵文字が同じ解釈を共有するとき、絵文字と絵文字解釈は多対多の関係にあり、絵文字の解釈に意味的あいまい性が生じる。

そこで与えられた解釈語と絵文字に付与されている解釈語集合との意味的な関連度を測る尺度 (Semantic Relevance Measure) を定義し、この尺度を用いて、与えられた解釈語に適した絵文字を順位付ける方法を提案し、検索語による絵文字検索システムを提案している。さらに、絵文字解釈語を概念辞書の上位概念体系を用いて予めカテゴリ分類し、カテゴリごとに意味的関連度を計算し、特定のカテゴリで関連度が高ければ検索結果として提示することで、絵文字検索の性能を向上できることを示している。

2. 人間が二文化の絵文字解釈のどの部分に注目し、文化差を読み取るかを明らかにした。これを実現するために、日本人と米国人からなる被験者に、日本と米国で収集した絵文字解釈の比較表を提示し、日米絵文字解釈の文化差を判定するようにしている。そしてインタビューを実施して被験者の文化差判定基準を調べている。インタビュー結果から、被験者が二国の絵文字解釈の意味的な違いだけでなく、解釈語の割合や解釈語集合のばらつき具合、固有名詞の有無や肯定的、否定的含意などの文化差判定基準を用いていることを明らかにした。

3. 人間の文化差判定基準を定式化し、二文化の絵文字解釈の文化差を機械的に読み取る判定式を定義した。これを実現するために、人間の文化差判定基準を、

- (1) 二文化の絵文字解釈の共通性、
- (2) 二文化の共通解釈の比率の違い、
- (3) 二文化の絵文字解釈のばらつき度合いの違い

を判定する式として定義し、三つの判定式を論理積で結合することで文化差判定式を定義している。この文化差判定式は、三つの判定基準を全て満たした絵文字解釈を文化差のあるものと判定する。日米絵文字解釈データを用いて文化差判定式の性能を調べた結果、高い文化差検出性能 (F値は0.7) を示している。

以上、本論文は異なる文化に属する人たちが与えた絵文字解釈のデータから文化差を解析する方法について研究した成果をまとめたものであり、学術上、実際に寄与するところが少なくない。よって、本論文は博士 (情報学) の学位論文として価値あるものと認める。

また、平成21年2月9日に実施した論文内容とそれに関連した試問の結果、合格と認めた。