

自然言語入力に基づく常識的心情判断方式

齋藤 安彰† 渡部 広一† 河岡 司†

†同志社大学大学院 工学研究科 知識工学専攻
知識情報処理研究室

概要

人間にやさしい、使いやすいコンピュータが求められる。そのために、人間と円滑なコミュニケーションを行えるコンピュータが必要となる。そのようなコンピュータを実現するためには、人のもつ「常識」を判断する能力をコンピュータにもたせる必要があると考えている。本稿では、常識的に話者の抱いている感情を判断する「心情判断メカニズム」について報告する。具体的には、話者の入力文解析から常識判断や概念連想を利用した感情導出までの手法である。入力文をフレームと呼ばれる成分に分解し、それぞれにカテゴリ化を行う。そのカテゴリの組み合わせ処理によって話者の感情を導出する手法を提案する。

Commonsense emotion judgment method based on natural language input

Yasuaki SAITO † Hirokazu WATABE † Tsukasa KAWAOKA †

† Intelligent Information Processing Laboratory,
Department of Knowledge Engineering and Computer Sciences,
Graduate School of Engineering, Doshisha University,

Abstract

We need useful and tenderness computer for human. Therefore we need computer, which can communicate with man naturally. It is necessary to give the capability to judge the "common sense" which people have to the computer. In this paper, the compound noun processing method in emotion judgment is proposed. It is a technique from speaker's input sentence analysis to the emotion judgment using the common sense judgment and the concept association. The input sentence resolves to the element which is called a frame, and it makes it to the category. The technique for deriving emotions by the combination processing of the category is proposed.

1. はじめに

近年、コンピュータは不可欠なものになり、社会におけるコンピュータの担う役割は多種多様なものとなっている。そのために、より使い易いインターフェースが望まれている。コンピュータを少しでも人間に近づけ、人間との双方向の会話を実現するためには、コンピュータに人間の感情を推測させることが重要な要素となる。例えば話し手が「プレゼントを貰った」と言えば、聞き手は話し手が喜んでいることを理解し、それに対して「よかったね」や「うらやましい」といった返答を返す。

このように、普段我々人間が会話を行う際には、その根底に流れる感情をお互いに読みとることで、会話はスムーズに進むと考えられる。

コンピュータと人間がこのような知的コミュニケーションを行うためには、コンピュータに人間が持つ常識的判断のための知識や考え方を教える必要がある。そこで、コンピュータに感情知識ベース、および人間の考え方を論理的にモデル化したメカニズムを与え、入力された文章に対して利用者の感情を推測させる。この「感情判断メカニズム」については、すでに報告されている^{[1][2][3]}。既に報告されている感情判断メカニズムにおいては単語区切り入力を対象としていた。本稿では、その感情判断メカニズムを心情判断と呼び、処理できなかった文型の入力に対する拡張を行った心情判断手法を提案する。

2. 概念ベースと関連度計算

ある語 A をその語と関連の強いと考えられる語 a_i と重み w_i の対の集合として定義する。

$$A = \{(a_1, w_1), (a_2, w_2), \dots, (a_m, w_m)\} \quad (1)$$

ここで、 a_i を 1 次属性と呼ぶ。また便宜上、A を概念表記と呼ぶ。このような属性の定義された語（概念）を大量に集めたものを概念ベースと呼ぶ。ただし、任意の 1 次属性 a_i は、その概念ベース中の概念表記の集合に含まれて

いるものとする。すなわち、属性を表す語もまた概念として定義されている。したがって、1 次属性は必ずある概念表記に一致するので、さらにその 1 次属性を抽出することができる。これを 2 次属性と呼ぶ。概念ベースにおいて、「概念」は n 次までの属性の連鎖集合により定義されている。本稿では、複数の国語辞書から、各見出し語を概念表記、その見出し語の説明文中の自立語を 1 次属性として抽出し、出現頻度に基づく重みを付加した約 4 万の概念からなる概念ベース^[4]を基に、新聞などから抽出した概念表記や属性を加え、質の向上を目的にした精練操作（属性の追加・修正など）を施し、更に概念間に成り立つ一般的なルールに基づく適切な重みを付加した約 9 万の概念からなる概念ベース^[5]を構築し利用している。

関連度とは、概念と概念の関連の強さの度合いであり、関連度は関連度計算^[6]によって求めることができる。関連度計算とは、その度合いを定量的に評価するものである。具体例を表 1 に示した。“果物”と“夕日”より、“果物”と“林檎”の方が意味が近く、関連度が高いということになる。具体的には概念連鎖により概念を 2 次属性まで展開したところで、最も対応の良い 1 次属性同士を対応付け、それらの一致する属性の重みを評価することにより算出するものである。

表 1 関連度計算結果の例

概念 A	概念 B	概念 A と概念 B の関連度
果物	林檎	0.26
果物	夕日	0.01

3. 感覚判断と意味理解

3.1 感覚判断システム

感覚判断システム^{[7][8]}は、ある語（名詞）に対して人間が常識的に抱く印象を形容詞・形容動詞の形で出力するシステムである。『痛い』『臭い』といった人間が五感で感じる印象を“五感感覚語”、『めでたい』『不幸な』といった五感以外で感じる印象を“知覚語”と呼び、

この2つを総称して“感覚語”と呼ぶ。この感覚語は合わせて 203 語ある。感覚判断システムの処理は、単語とその特徴である感覚の関係に関する代表的な知識を持たせ（図 1）、その知識を概念ベースを用いて拡張し、あらゆる単語に関する感覚語を精度良く出力するよう工夫されている。

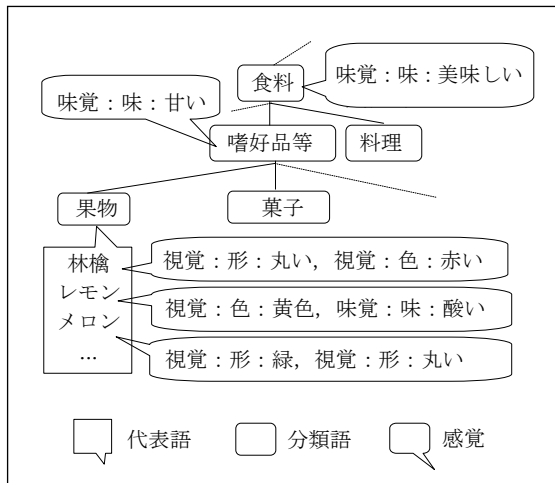


図 1 感覚知識ベース

3.2 意味理解システム

文型を限定しない形式での入力に対応するためには入力文を一定の形式に変換する必要がある。そこで意味理解システム⁹⁾を用いて単文をフレームに分解するという操作を行う。意味理解システムとは単文入力に対して 6W1H と用言という 8 個の成分に分けることのできるシステムである。例えば、「昨日店で妹が母にセーターを買ってもらった」という入力は表 2 のようなフレームに分けられる。

表 2 意味フレーム

フレーム	語
Who	妹
What	セーター
When	昨日
Where	店
Whom	母
用言	買ってもらう

このフレームを意味フレームと呼ぶ。この中

から心情導出に関係の深い 6 つのフレームを用いることとする。

4. 心情判断と感情判断

心情判断メカニズムとは入力された語から心情をコンピュータが判断するものである。心情判断システムでは、“感情語”と“心情語”を扱う。感情語とは発言の時点に瞬間的に発生すると考えられる『喜び』、『安心』、『悲しみ』、『恐れ』、『怒り』、『落胆』、『後悔』、『恥』、『罪悪感』に加え、感情『なし』の計 10 種である。結果はこの 10 種類のうちの 1 つ、または複数となる。心情語とは、『快』、『不快』をそれぞれ『強』、『中』、『弱』の三段階で表現したものである。例えば「私は宝石を貰う」といった場合、感情は“喜び”となり、心情は“快”となる。感情語を導出する操作を感情処理、心情語を導出する処理を心情処理と呼ぶ。また、感情処理メカニズムに心情処理を加えたものを心情判断メカニズムと呼ぶ。

4.1 心情判断メカニズムの構成

心情判断メカニズムの概観は図 2 のようになっている。

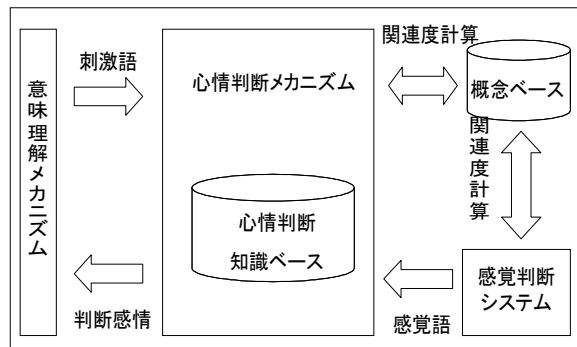


図 2 心情判断メカニズムの概観

心情判断メカニズムでは、入力要素を刺激語と呼ぶ。自然言語で入力された単文を意味理解システムによってフレームに分割する（表 2）。

意味理解システムの結果は心情判断システムにとって不都合のある場合もある。Who フレームには文の主語が入るが、これは心情を持ち得る主体とは限らない。例えば、「台風が来

る」という場合は、Who フレームは台風となるが、台風自体は感情を持っていない。そこで Who フレームを“私”，What フレームを“台風”，用言フレームを“来る”というように補正を行う。つまり心情判断システムは「私のいるところに台風が来る」と理解する。このように補正して得られたフレームに対して心情導出を行う。補正されたフレームを刺激語とする（表 3）。

4.2 カテゴリ化

前述した6つの要素の刺激語を入力として、それらの情報から心情を導出する。しかしこれらの要素の組み合わせは膨大であり、このままで心情判断を行うことは困難である。そのため、各入力要素をそれぞれ幾つかのカテゴリに抽象化し、各カテゴリの組み合わせから心情を判断する。この抽象化をカテゴリ化と呼ぶ。各刺激語に対するカテゴリ化は表 3 のようになる。

表 3 フレームのカテゴリ化

刺激語	分類語
主体語	主体語カテゴリ化
目的語	目的語カテゴリ化
時間語	時間語カテゴリ化
場所語	目的語カテゴリ化
人目的語	主体語カテゴリ化
動作語	動作語カテゴリ化

このカテゴリ化には心情判断知識を保有している知識ベース（図 3）を用いてカテゴリ化を行うものと、外部システムを用いてカテゴリ化を行うものがある。表 3 において、主体語カテゴリ、変化語カテゴリが知識ベースを用いたもので、目的語カテゴリが3.1節の感覚判断システムを用いたカテゴリ化、時間判断システム [10][11]を用いたカテゴリ化である。時間判断システムとは時間を表す語に対して具体的な日時を返答するシステムで、「昨日」などといった入力にも対応するシステムである。

主体語処理 [2] において、各名詞を3つの角度から常識的範疇の内での評価し、各々に A B C のランクを設け、その組み合わせ、即ち計 27

通りに分類した。それらを以下に記述する。

- (1) 精神的尺度 好き嫌い（好き－ A B C －嫌い）
- (2) 感覚的尺度 身近さ（近い－ A B C －遠い）
- (3) 社会的尺度 善悪（善－ A B C －悪）

「父親」，「泥棒」に対して主体語カテゴリ化の処理を用いると、父親は (1, 2, 3) = (A, A, B)，泥棒は (C, B, C) となる。この分類で話者に対する主体語の印象を区別する。主体語と話者との間の感情の“ずれ”の表現は話者感情生成テーブルで定義している。

目的語カテゴリ化には名詞句の意味理解アルゴリズムを使用する [12]。まず品詞毎に名詞は感覚判断システム、形容詞と動詞は目的語知識ベースによって感覚語に置換する。感覚語にはあらかじめ重要度が定義されているので、重要度が最も高いものを目的語カテゴリとする。

動作語カテゴリ化では入力される語は動詞、サ変接続名詞が多い。この動作語は“感情一意想起動作語”と“目的語依存型動作語”の大きく2種類に分類される。感情一意想起型とは“喜び型”，“悲しみ型”などといった目的語に依らずカテゴリを決めるものであり、目的語依存型とは“継承”，“逆転”など目的語のカテゴリにより、そのイメージを受け継ぐか、覆すかといった分類に定義されている。入力語を元に、知識ベースを参照しカテゴリ化を行う [3]。

時間語のカテゴリ化には時間判断システムを用い、現在との時間差を算出する。

4.3 心情判断知識

心情判断システムが保有している知識ベースは大きく2種類に分類され、4.2節で述べたカテゴリ化に用いる知識ベースと感情生成に用いる知識ベースである（図 3）。【主体語知識ベース】・【目的語知識ベース】・【変化語知識ベース】は、入力情報からそれぞれカテゴリへ代替する際に用いる知識を格納しており、【主体感情生成テーブル】【話者感情生成テーブル】は、各カテゴリ（話者感情生成テーブルについては、主体語カテゴリと主体感情）の組み合わせ

せにより想起する感情を定義したテーブルである。

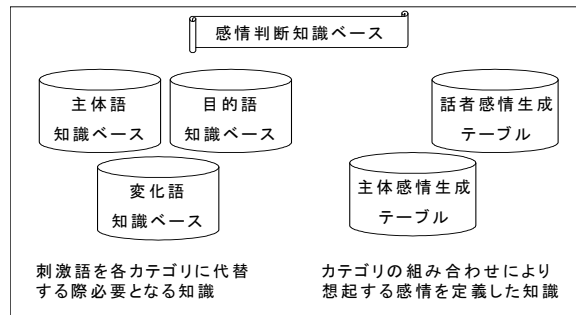


図 3 心情判断知識ベース

5. 感情導出

感情導出は大きく通常感情処理, 人目的感情処理, 場所条件感情処理の3つに分類される。それぞれ図 4に示したアルゴリズムによって条件分けされる。3つに分類された感情処理を経て, 最終的に心情強度を加え, 感情語と心情語が出力される。

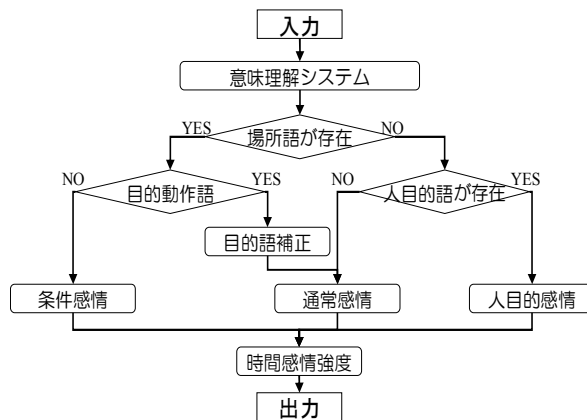


図 4 心情判断システムの流れ

5.1 通常感情処理

処理の流れを図 5に示した。通常心情処理は, 主体語カテゴリ, 目的語カテゴリ, 動作語カテゴリの組み合わせから心情を導出するものである。ここでは目的語カテゴリと動作語カテゴリから主体感情生成テーブルを用いて主体感情を導出している。

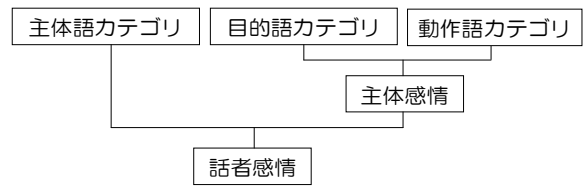


図 5 通常心情処理の流れ

5.2 人目的感情処理

人目的感情処理とは人を目的とした動作の場合, 主体とその目的との関係を捉えて心情を導出する必要とするものがある。そのために, 同じ手法を用いてカテゴリ化した主体語と人目的語のカテゴリから新たに相対カテゴリを導出する。相対カテゴリとは, 主体語カテゴリと人目的語カテゴリの差をとったものである。それを用いた処理の流れを図 6に示した。

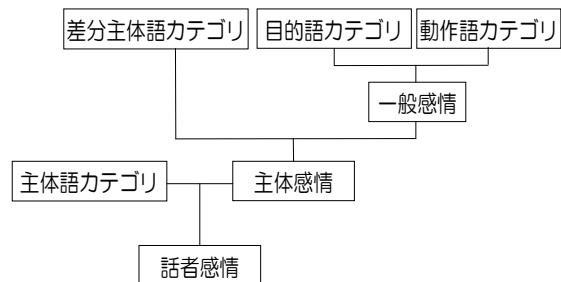


図 6 人目的語処理の流れ

図 3 心情判断知識ベースに示した話者感情生成テーブルは, 主体の感情を話者の感情に変換するための知識ベースである。具体的には主体語カテゴリと感情語の組から感情語が導出される。例えば「泥棒 (C, B, C) +喜ぶ」から「怒り」が導出される。この主体語カテゴリとは話者の視点から見た基準であり, これを相対カテゴリ, つまり主体から見た基準に置き換えることで, 主体感情の導出にも応用できる。この処理を行った後に通常の基準で話者感情生成テーブルを用いることで話者感情を導出する。

5.3 条件感情処理

条件心情処理とは場所語がある場合の処理である。この場所語は動作語によって場合分け

される。これは大きく2つに分類され、動作語が目的動作語であるかで判断される。

- (1) 私は遊園地へ行く
- (2) 私は賑やかな場所で転ぶ

(1) の例文は主体語+目的語+動作語の構成の文型になるので、これは通常感情処理での処理が適当であると考えられる。しかし (2) の例文は、主体語+目的語+動作語という形式には当てはまらない。「私は転ぶ」という文に“賑やかな場所で”という条件が加わることによって『恥』という感情が発生すると考えられる。

(1) のような目的語をとる動作語を目的動作語とし、(2) のような条件になる動作語を条件動作語と呼ぶ。この判断は NTT シソーラス^[13] を用いて作成した動作語の知識ベースで判定している。処理の流れを図 7 に示した。

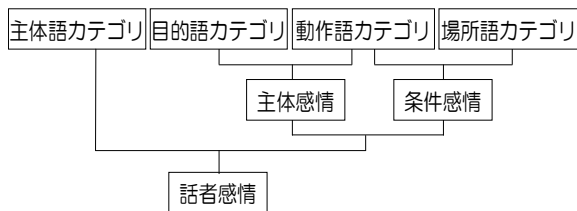


図 7 条件感情処理の流れ

まず通常感情処理と同様に、目的語カテゴリと動作語カテゴリから主体感情を導出する。それとは別に場所語カテゴリと動作語カテゴリから条件感情を生成する。条件感情は主体語の感情であるために、話者感情を導出する前に考慮する必要があるため、主体感情に条件感情を加えたものから最終的な話者感情を導出する。

6. 心情導出

心情判断システムは発話者の心情を判断するシステムである。この発話の時制は現在ということになる。発話内容の時間語は様々な役割を果たすが、ここでは発話内容の起こった時制を表すものについて対象とする。具体例を以下に示した。

- (3) 昨日、父が死んだ
- (4) 10年前、父が死んだ

このように昨日起こったことについて発話した時点での感情、心情を判断することを考える。このような場合、感情としてはどちらも“悲しみ”が適当であると考えられるが、その強さに差があると考えられる。この強さを表現するために心情語を用いることとする。人の心情というものは非常にあいまいであり、厳密に強さを定義することは困難である。よって正の心情である『快』と負の心情である『不快』の2つを用いて表現する。そのために話者感情生成テーブルを利用する。表 4 に話者感情生成テーブルの例を示した。

表 4 話者感情生成テーブルの構造(例示)

精神的 尺度	感覚的 尺度	社会的 尺度	主体 感情	喜び	安心	...
A	A	A	喜び	10	7	...
A	A	B	喜び	8	5	...
B	A	C	怒り	7	3	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮

この知識ベースによって話者感情が導出されるが、主体語カテゴリと主体感情の組み合わせ以外にも、感情語が複数導出される場合の優先度も定義されている。これは最大値 10 であり、サンプルデータを用いて算出されている。この数値を用いて心情強度を導き、感情の種類から『快』、『不快』を判定する。つまり感情導出結果が“喜び、強度 10”だとすると、“快、強”となる。この数値は実験的に求めた数値を元に決定した。

7. 評価

通常心情処理については既に報告されている^[2]。ここでは、人目的心情処理と条件心情処理について評価を行った。元となるテストデータが 655 セットある。しかしこれは主体語、目的語、動作語から構成されるテストデータで

あり、これを文に修正を行ったテストデータを用いる。その中から各判断に分類されるテストデータを抽出し、評価を行った。

テストデータセットの作成においては非常識ではない感情の導出を目的としているため、文に対して非常識であると判断された感情、つまり“×”と定義されている感情語が導出されれば不正解、それ以外の感情語が導出されれば非常識とはいえないということで正解とした(表 5)。

表 5 テストデータのセット

文	喜び	安心	悲しみ	恐れ	...
私は可愛い子供を見る	○		×	×	...
私は好きな女性を失う	×	×	○		...
⋮	⋮	⋮	⋮	⋮	⋮

7.1 人目的感情処理の評価

テストデータのうち、目的語に人を取るテストデータは 106 セットあった。その 106 セットで評価を行った結果を図 8 に示した。ここでは人目的語を考慮しないで通常感情処理を行ったものの精度と比較した。

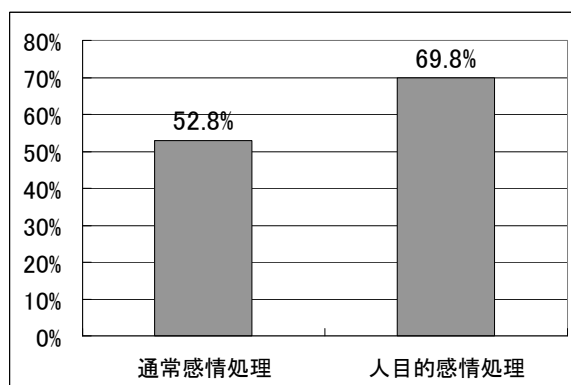


図 8 人目的感情処理の結果

ここでは全てを通常感情処理で処理する場合と比較して約 17%の精度向上が見られた。これは相対カテゴリを用いることにより、主体から見た人目的語の印象が反映されていることが原因である。更に、感覚判断システムでは、

一般的な名詞に関してはかなりの精度で感覚語を出力するが、人に関する語、“親”、“親戚”などの語に対しては対応できていない。そこで、主体語カテゴリ化の処理を行った結果である人目的語カテゴリから感覚語を導出する補正を行った結果である。

7.2 条件感情処理の評価

テストデータのうち、場所語が存在するテストデータは 64 セットであった。動作語が目的動作語であるか、条件動作語であるか野判定を行って初めて処理が分割できるので、これを合わせたものである。場所条件を考慮することによって約 13%の精度向上が見られた(図 9)。

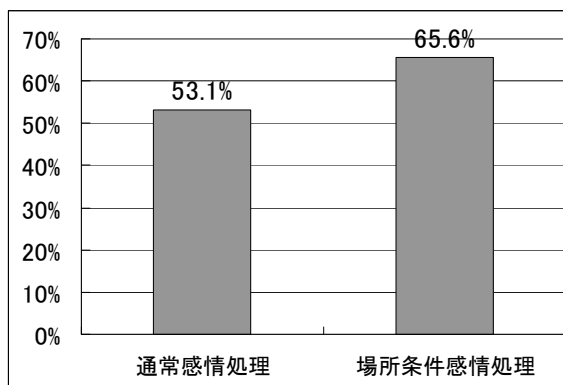


図 9 条件感情処理の結果

7.3 感情判断システムの評価

655 セットの評価データの意味理解システムの精度は、適切に補正可能な分解がされるかという基準で約 97%であった。7.1, 7.2 節のテストデータのセットを含む補正可能であった 640 セットのデータを用いてシステム全体の評価を行った。話者感情に非常識なものがないものを正答とした(図 10)。

感情処理(単語区切り)は元の区切りテストデータを用いてすべて通常感情処理を用いたものである。つまり文を分解するという操作を行っていない。心情判断が図 4 に示したアルゴリズムによって感情導出処理を分岐させた評価である。これは入力を自然言語文としているのにもかかわらず、多少であるが精度向上に成功している。

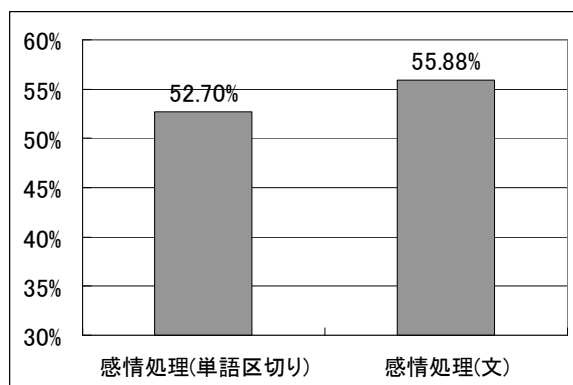


図 10 話者感情の評価結果

8. 考察

意味理解システムを用いることによって、目的語の意味を理解し、それに応じた抽象化をすることで精度向上につながった。目的語を統一した手法だけでなく、場所語、人目的語と理解することで、補正を行い更なる精度向上に成功した。しかし、目的語カテゴリを導出できない語の場合、主体感情が『なし』となってしまう。条件感情条件に適合すれば感情が発生するが、それ以外の処理では感情が発生することはない。よって処理の初期段階で感情がないと判断されてしまうと感情導出が困難になる。それを避けるためにカテゴリ化の手法の精度向上が今後の課題となる。

9. おわりに

本稿では、自然言語入力における話者感情、心情の導出手法を提案した。具体的には入力を自然言語の単文とすることによって、より様々な入力に対応することができるようになった。また文の構成によって感情導出のパターンをわけることによって、より精度の高い判断システムにつながると考えられる。

謝辞

本研究は文部科学省からの補助を受けた同志社大学の学術フロンティア研究プロジェクトにおける研究の一環として行ったものである。

参考文献

- [1] 土屋誠司, 小島一秀, 渡部広一, 河岡司: 常識的判断システムにおける未知語処理方式, 人工知能学会論文誌, Vol.17, No.6, B, pp.667-675, (2002)
- [2] 風間勇志, 植野義孝, 渡部広一, 河岡司: 常識的感情判断と主体語処理, 情報科学技術フォーラム FIT2002, E-28, pp.137-138, (2002)
- [3] 植野義孝, 風間勇志, 渡部広一, 河岡司: 常識的感情判断システムにおける動詞の処理, 情報科学技術フォーラム FIT2002, E-29, pp.139-140, (2002)
- [4] 笠原要, 松澤和光, 石川勉: 国語辞書を利用した日常語の類似性判別, 情報処理学会論文誌, Vol. 38, No. 7, pp.1272-1283, (1997)
- [5] 広瀬幹規, 渡部広一, 河岡司: 概念間ルールと属性としての出現頻度を考慮した概念ベースの自動精練手法, 信学技報, NLC2001-93, pp.109-116, (2002)
- [6] 井筒大志, 渡部広一, 河岡司: 概念ベースを用いた連想機能実現のための関連度計算方式, 情報科学技術フォーラム FIT2002, pp.159-160, (2002)
- [7] A. Horiguchi, S. Tsuchiya, K. Kojima, H. Watabe, T. Kawaoka: Constructing a Sensuous Judgment System Based on Conceptual Processing, Computational Linguistics and Intelligent Text Processing (Proc. of CICLing-2002), Springer, pp.86-95, (2002)
- [8] 米谷彩, 渡部広一, 河岡司: 常識的知覚判断システムの構築, 第17回人工知能学会全国大会論文, 3C1-07, (2003)
- [9] 篠原宜道, 渡部広一, 河岡司: 常識判断システムを用いた会話意味理解方式, 情報処理学会自然言語処理研究会資料, B6-2, pp.89-96, (2002)
- [10] 小畑陽一, 渡部広一, 河岡司: 時間の理解に関する常識的判断メカニズム, 情報処理学会第59回全国大会講演論文集
- [11] 野村理樹, 渡部広一, 河岡司: 時間の常識的判断システムにおける未知語処理方式の検討, FIT2003, pp.191-193, (2003)
- [12] 齋藤安彰, 風間勇志, 渡部広一, 河岡司: 感情判断システムにおける名詞句修飾の意味理解, 第18回人工知能学会全国大会論文集, 2D2-06, (2004)
- [13] NTT コミュニケーション科学研究所 日本語語彙体系, 岩波書店 (1997)