

コンピュータを用いた感情判断における名詞句修飾の意味理解

Semantic Understanding of The Noun Phrase in Emotion Judgment System

齋藤 安彰
Saito Yasuaki

風間 勇志
Kazama Yuji

渡部 広一
Watabe Hirokazu

河岡 司
Kawaoka Tsukasa

*1 同志社大学工学研究科

Department of Knowledge Engineering and Computer Sciences, Graduate School of Engineering, Doshisha University

In order to make computer friendlier for humans, it is thought that the construction of a Commonsense Judgment System, which deals with commonsense like humans, is necessary. In this paper, Emotion Judgment System, which judges the human's emotion from text sentences, is constructed. Especially, this paper treats the text sentences including noun phrase modifications.

1. はじめに

近年、コンピュータは普遍的なものになり、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意図を適切に汲みとる知的能力を持たせることが必要となる。その知的能力には、まず質問の意味を理解し、その質問に対して常識的な判断を下す能力が含まれる。判断には、時間や物の大小、速さ等を判断する量的判断、暑い寒いといった五感に関する感覚判断[Horiguchi 02]、そして喜怒哀楽等の感情判断があり、それぞれにおいて柔軟な判断を行うことが望まれる。

前述の通り、コンピュータと人間が会話するためには、コンピュータに人間の感情を推測させることが、大切な要素となる。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[土屋 02; 風間 02; 植野 02]。本稿では、この感情判断メカニズムにおける一つの要素として、修飾語の処理の方式について提案する。

しかし、小説のような文章の感情を判断させるシステムを考える場合、人は文章を短文の集合として捉えているのではなく、物語の展開や、印象に残った場面を元に判断を行っていると考えられる。そのような判断システムには、これまでとは異なった手法を用いて判断させる必要がある。つまり、文章の判断には、1文処理の拡張として考えるのではなく、文章をひとまとまりとして考えるという考え方で処理を行う必要があると考えられる。本稿では文章の判断に用いることの出来る処理手法を提案する。

2. 感情判断システム

感情判断システムでは、会話処理メカニズムから適切な自然言語処理の結果として主体語、目的語(修飾語+対象語)、変化語を入力として得る事を前提としている(図 2.1)。これらの入

力要素を刺激語と呼ぶ。そして刺激語から発話者の感情を判断結果として会話処理メカニズムに返す。会話処理メカニズムでは感情判断システムによって得られた発話者の感情情報を元に応答処理により、返答文やジェスチャー等を生成する事が出来る。また、感情判断システムは、感情判断のために必要な判断知識ベースを持つ。この判断知識ベースを感情判断知識ベースと呼ぶ。感情判断システムで扱う感情の種類は「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。

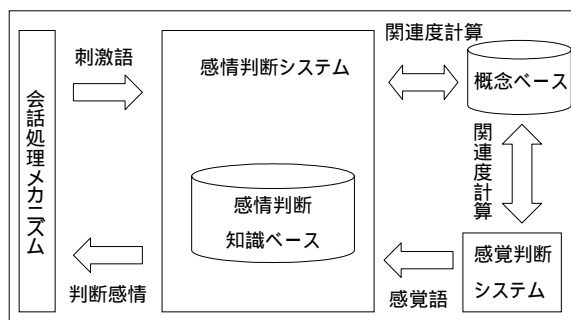


図 2.1 感情判断システムの概観

3. 感情判断システムの処理の流れ

前述した4要素を入力として、それらの情報から話者の感情を判断するが、これら4要素の入力情報の組み合わせは膨大である。そのため、各入力要素をそれぞれ幾つかのカテゴリに代替する事で組み合わせ数の問題を解消している。そして、各カテゴリ(及び主体感情)の組み合わせから感情を判断することになる。図 3.1 に感情判断システムの処理の流れを示した。

図 3.1 に示したように、主体語は主体語カテゴリ、修飾語+対象語は目的語カテゴリ、変化語は変化語カテゴリにそれぞれ代替する。これらの処理を分類処理と呼ぶ。そして、それぞれのカテゴリのうち、目的語カテゴリと変化語カテゴリの組み合わせから主体の感情を導出する。この感情を主体感情と呼ぶ。主体感情は文の主体である主体語の感情を指す。例えば、「泥棒+お金+得る」という入力情報であった場合、主体語である泥棒の感情が主体感情に当たり、この場合「喜び」の感情が導出される。次に、この主体感情と主体語カテゴリとの組み合わせから、発話者の感情(この場合、「怒り」)が導出される。この発話者の感情を話者感情と呼ぶ。

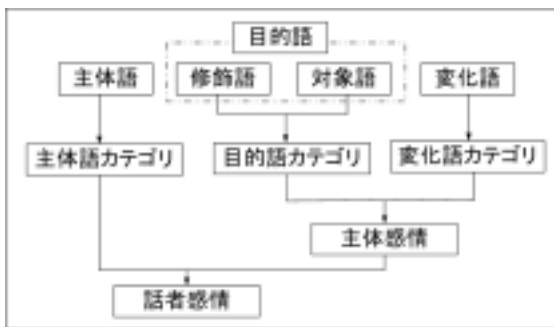


図 3.1 感情判断システムの処理の流れ

3.1 感情判断知識ベース

感情判断システムが保有する感情判断知識ベースは、図 3.1 に示した流れのうち、各刺激語の情報をそれぞれのカテゴリに分類するために用いる“分類処理知識”と、カテゴリから感情を導出する際に用いる“感情生成テーブル”と呼ばれる知識からなる。図 3.2 に感情判断知識ベースの内容を示した。

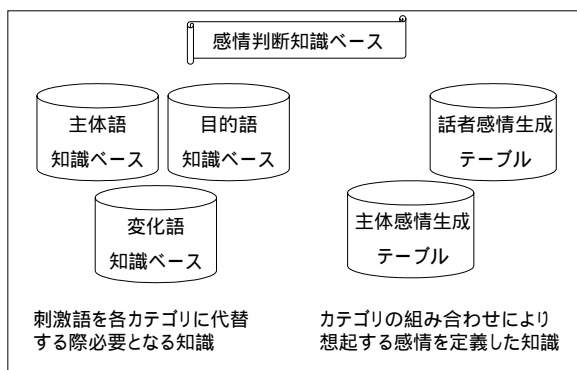


図 3.2 感情判断知識ベース

図 3.2 における【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】については、主体語カテゴリと主体感情の組み合わせにより想起する感情を定義したテーブルである。

3.2 感覚判断システム

対象語のカテゴリ化には感覚判断システム[Horiguchi 02; 米谷 03]が使われている。感覚判断システムとは、ある語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。人間が五感で感じる印象を“五感感覚語”、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。

3.3 目的語の処理

図 3.1 において、目的語のカテゴリ化は修飾語、対象語の組み合わせで行われる。修飾語として形容詞・形容動詞の品詞を受け付けている。修飾語は大きく4つに分類され、分類を一意に決めてしまう修飾語を“直接修飾型”、修飾する対象語によって分類を変化させるものを“依存修飾型”、対象語の分類に変化を与えない語を“無修飾型”、後ろの対象語に関する強さ(程

度)を表すもの“程度表現型”と定義し、それぞれの形容詞・形容動詞が知識ベースに登録されている。

3.4 変化語の処理

変化語として入力される語は動詞、サ変接続名詞が多い。この変化語は“感情一意想起動詞”と“目的語依存型動詞”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よらず分類を決めるものであり、目的語依存型とは“継承”、“逆転”など目的語の分類により、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。

4. 概念ベースと関連度計算

ある単語 A をその単語と関連の強いと考えられる単語 a_i と重み w_i の対の集合として定義する(式 1)。

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

ここで、 A を概念表記、 a_i を1次属性と呼ぶ。このような集合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ある概念表記の集合に含まれているものとする。従って、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。

本研究では、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[笠原 97]を基に、概念の追加や質の向上を目的とした精練操作(属性の追加・修正など)を施し、ルールにより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[広瀬 02]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、関連度計算[井筒 02]とは、その度合いを定量的に評価するものである。“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。

5. 修飾処理の拡張

極めて多種多様な入力に対応するために、修飾処理の拡張を行った。知識ベースの変更を行い、修飾語として重要だと考えられる動詞、名詞による修飾処理の拡張を行った。

5.1 目的語分類語の処理

目的語分類語とは、目的語カテゴリに導出される語である。目的語分類語は33種類と“なし”の34種類であったが、刺激語の要素である対象語の処理に、感覚判断システムが導入されたために、目的語カテゴリで扱える34種類以外の語の処理が必要となった。感覚判断システムの出力を目的語カテゴリで扱える34語に置き換えることは、適当な語がないために困難であるために、感覚判断システムの出力をそのまま分類語として追加し、処理をする。

5.2 分類語の重み付け

従来の感情判断システムでは、刺激語として修飾語が入力されれば、原則的に修飾語の分類を優先する処理を行っていた。

しかし、複数品詞による修飾語の処理にそれは適用することができない。そこで複数品詞による修飾語の処理において、感情の発生の核となる部分を判定させるために、分類語に重みを定義した。これは分類語が感情を想起し易い語と、感情を想起しづらい語があるためである。

その度合いを定量的に評価するために関連度計算を用いる。感情語と分類語の意味的関連が近ければ感情を想起しやすいと考えられる。しかし、感情判断システムで扱っている感情は10種類と数が少なく、プラスの感情語よりもマイナスの感情語の数が多いために、関連度がマイナスに偏ってしまう可能性がある。そこで、情緒の系図[九鬼 91]の感情語を利用する。これは人の感情を計略立てたものであり、感情語は43語で比較的偏りが少ないと考えられる。この43語の感情語と203語の関連度をとり、関連度が最大となった数値を比較した。

この結果を元に重みを決定する。関連度が最大値の1となっている語があったが、これは目的語分類語が43種類の感情語に含まれていたということである。この語に関しては感情を想起し易いと判断することが出来る。判断基準として、まず関連度をそのまま重みと設定することを考える。これは全てのカテゴリに差をつけるということである。しかし、感情を発生しやすい語を完全に数値化することはできないので、ある程度の区切りをつけてランク付けを行う。ここで3つに区切りランク付けをした。これは、閾値の設定が容易であり、それ以上に細かく分類することが困難であるためである。

5.3 動詞修飾知識ベースの作成

既に動詞を扱える変化語の処理の利用を考える。変化語カテゴリは12種類である。この12種類から203種類にカテゴリ化するのは困難であるため、動詞から203種類の目的語分類語を直接導出する。このために動詞知識ベースに目的語分類語を追加する。目的語分類語の付加にはこの動詞知識ベースで使用されている動詞ソーラス構造を利用する。このソーラスはカテゴリ、小カテゴリ、語彙の3層構造となっている。小カテゴリごとに目的語分類語を定義する。一意想起型動詞には目的語分類語をそのまま付加、“継承”、“逆転”の依存型動詞に関しては、一意に分類語を付加することはできないので、依存型の分類をそのまま付加した。この作業は人手で行った。

5.4 動詞分類処理

刺激語として、動詞修飾語、対象語が入力された場合、動詞修飾語が一意想起型の場合は、動詞修飾語、対象語の分類語をそれぞれ導出し、分類語の重みによって優先する分類語を決定する。動詞修飾語が依存型であった場合、分類語が“継承”であるなら、対象語の分類語を目的語分類語とする。分類語が“逆転”であるなら、別に作成した逆転知識ベースを元に対象語の分類語を反転させる処理を行う。逆転知識ベースとは、目的語分類語にそれぞれ反対となる目的語をあらかじめ対応付けたものである。反対語の目的語分類語が明らかに対応しないものは、“なし”を対応させた。

5.5 名詞修飾処理

名詞修飾には対象語のカテゴリ分類に使われている感覚判断システムを適用する。名詞修飾語、対象語それぞれの分類語を感覚判断システムによって導出し、その重みによって優先する刺激語を決定する。

6. 複数品詞による修飾処理

入力文の局所的な修飾部分には感情発生に関する核となる部分が存在することから、分類語の重みを利用して、これまでの処理を拡張し、複数品詞による修飾処理を行う。処理の流れを図6.1に示した。

まず各修飾毎に分類語を導出、知識ベースを参照し重みを出す。刺激語がない要素の分類は“なし”となる。そしてまず動詞修飾語の分類を判定する。これは、逆転というカテゴリが、イメージを反転させるという働きを持っているためである。逆転以外の場合はすべての分類語に対し重みを比較、優先する重みを決定する。逆転であった場合は、動詞修飾語の分類語以外の分類語を比較し、決まった分類語を前述の逆転知識ベースで反転処理を行う。継承であった場合には継承を重み0とし、重みによる優先処理を行う。

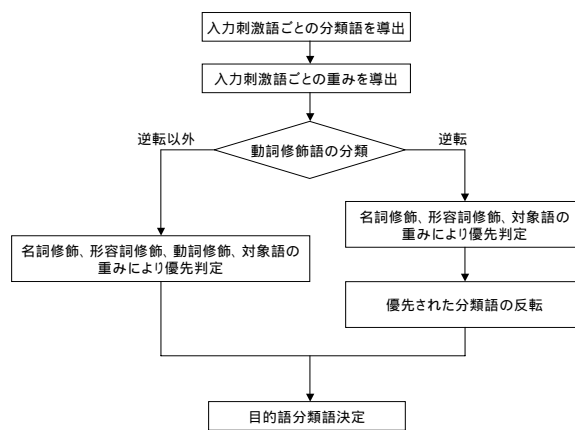


図 6.1 複数品詞による修飾処理の流れ

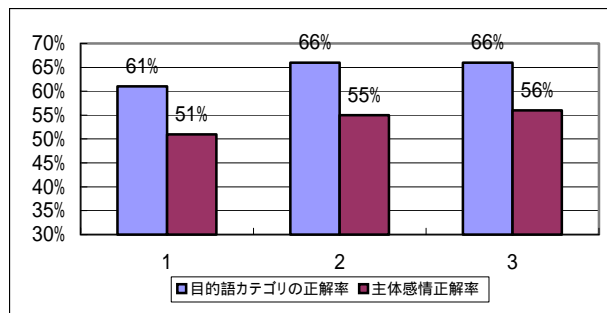


図 7.1 重みの評価

7. 評価

アンケートによって集められたテストデータ100セットでこのシステムの評価を行った。このテストデータは、感情を発生しやすいと思われる1つの文からなるもので、主体語、修飾語、対象語、変化語のうち、主体語を“私”と限定し、修飾語を名詞、形容詞・形容動詞、動詞のうちの2つ以上を含む文と限定した。修飾語、対象語のセットのうち判断の核となる部分を人が判断してある。修飾語の評価として、核となる部分が正解であるかどうかで評価を行った。全体の評価として、主体感情の“常識的”(○)、“非常識”(×)の2つの区分で記入してもらった。出力されても非常識とは言えないという感情は空欄になっている。優

先された分類語がどの修飾語を元に決定されたかを評価する。感情全体の評価は、常識的と判断されている“ ”の感情が複数ある場合は“ ”の感情が最低ひとつ出力され、非常識と判断されている“×”の感情が出力されない状態を正解とする。

7.1 重みの評価

図 7.1 において、1 が関連度で優先を決定したもの、2 が対象語を含まず、修飾語を重みにより優先を決定したもの、3 が対象語を含めて重みにより優先を判定したものである。重み優先では精度が向上している。

7.2 マイナスイメージの優先

7.1 節の評価を考察した結果、人はマイナスのイメージを優先させる傾向があると考えられる。感覚語にはプラスとマイナスの情報があらかじめ付加してあるので、それを利用しマイナスのイメージの重みを増やしてやる必要がある。これは実験により 1.5 という数値に設定した。

7.3 対象語重みの調整

修飾語は対象語の印象を変化させるべくつけられたものである。対象語の重みを下げて評価を行った(図 7.2)。この結果により、-1 が適当であると考えられる。

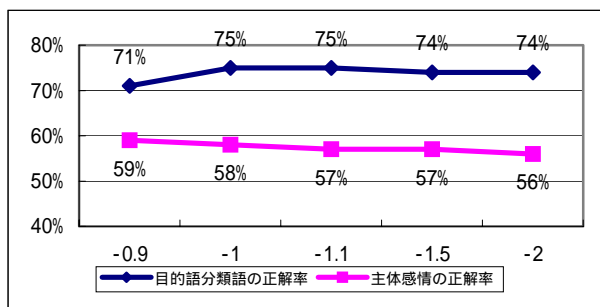


図 7.2 対象語重み調整

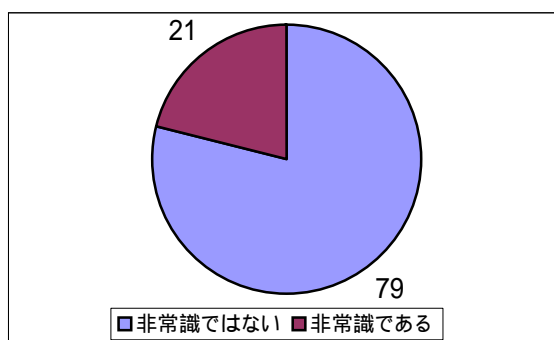


図 7.3 分類語の評価

7.4 精度

7.3 では人が優先しているものと同様のものを判断できるかどうかという評価であったが、優先された刺激語を元にカテゴリ化された分類語が正しいかどうかの評価を行った(図 7.3)。

感情の核となる刺激語を 75% 優先可能だが、その分類語が非常識なものではないかというものを更に評価するために、分類語を評価した。その結果 79% が非常識でない出力をすること

ができた。この結果が 7.3 節の 75% を上回っている理由について考察を行った。“綺麗な夜景の見える場所”という名詞句があった場合、人は“綺麗な”という刺激語を感情の核として考えるとされる。この刺激語から“美しい”という分類語を導出することが出来る。しかし、“夜景”という刺激語からも“美しい”という分類語を導出することが出来る。分類語重みが等しい場合は、関連度を用いて優先語を決定しているために、“夜景”という刺激語が優先され、分類語の結果は非常識でないとなる。

このように、優先する語以外からも非常識でないカテゴリ化が行われているために、多少正答率が向上している要因となっている。

8. おわりに

本稿では極めて多種多様な入力に対応するために、名詞句による修飾処理を目指した。そのために処理できる品詞を拡張し、感情を発生しやすいカテゴリに対し重み付けを行い、感情発生の際となる要素を判別することによってそれを実現した。この処理は入力語情報のみで行うもので、構文情報の利用によって精度向上が見込まれる。しかし、構文情報も使用することなく 79% は判断可能であるということを示せた。

今回はこのような局所的な部分に適用して、評価を行ったが、このような手法を用いることによって、一文からの判断ではなく、文章の判断にも利用できると考えている。

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「知能情報科学とその応用」における研究の一環として行った。

参考文献

[土屋 02] 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.

[風間 02] 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理, 情報科学技術フォーラム FIT2002, E-28, pp.137-138, 2002.

[植野 02] 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.

[笠原 97] 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.

[広瀬 02] 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.

[井筒 02] 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.

[Horiguchi 02] 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.

[米谷 03] 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.

[九鬼 91] 九鬼周造: 「いき」の構造, 岩波書店, 1991.