

嗜好を考慮した感情判断システムの構築

An Emotion Judgment System Based on Human Preference

土屋 誠司†, ‡
Seiji Tsuchiya

渡部 広一†
Hirokazu Watabe

河岡 司†
Tsukasa Kawaoka

1. はじめに

近年、情報処理システムは急速に発展し、あらゆる分野で使用されている。しかし、どんなに便利であっても、自由に使いこなせるようになるまでに多大な時間と労力が必要であるならば、その恩恵を受けようとし、受けられないユーザが存在する。もし、コンピュータが人間とコミュニケーションを行うことができれば、マニュアルなどを見なくとも、人間と会話をするように誰でも簡単に操作することが可能となる。

人間と会話ができる知的なコンピュータの実現には、コンピュータにも人間と同じような常識的な判断が行える必要がある。そのためには、ある語から概念を想起し、さらに、その概念に関係のある様々な概念を連想できる能力が重要な役割を果たす。

これまで、コンピュータに常識的な判断を行わせるための要素技術として、言語から話者の抱く感情を判断する「感情判断システム」の研究[土屋 02][風間 02][植野 02]がなされてきた。感情判断システムでは、感情という繊細であまいなものを扱うため、人間のような木目細やかで知的な判断が望まれる。

本稿では、人間が行う常識的な判断の一つである感情に関する判断について、ある概念から様々な概念を連想できる「概念連想メカニズム」を基に状況を考慮した感情判断システムを構築し、より人間に近い感情判断を可能にする手法を提案する。なお、「状況」という概念は非常に広範囲におよぶため、本稿では、その中で「ユーザの嗜好」に着目してシステムの構築を行う。

2. 感情判断システム

感情判断システムの概略を図1に示す。

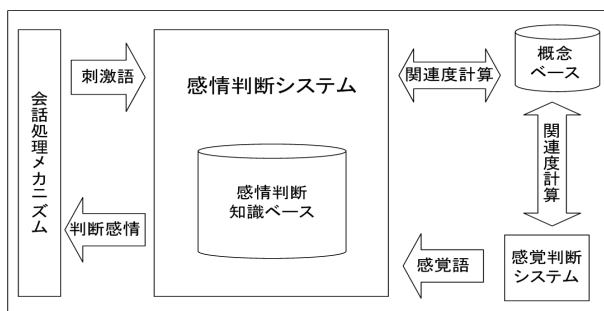


図1. 感情判断システムの概略図

感情判断システムとは、発話文から発話者ならびに発

話文の主体の感情を判断するシステムである。判断結果として扱う感情は「喜び」、「安心」、「悲しみ」、「恐れ」、「怒り」、「落胆」、「後悔」、「恥」、「罪悪感」、「感情なし」の計10種類である。

感情判断システムに入力される語は、人間が日常発している自然な言葉に対して、会話処理メカニズムにより適切な自然言語処理が施された結果であることを前提としている[篠原 02]。それは、感情発生の要因となる主体（以下、主体語と呼ぶ）、修飾語と対象語で構成する感情発生の要因となる物事（以下、目的語と呼ぶ）、感情発生の要因となる行為・作用（以下、変化語と呼ぶ）である（これら感情判断システムに入力される要素を以下、刺激語と呼ぶ）。

感情判断システムでは、主体語、目的語、変化語に関する知識と、それらの組み合わせから判断される感情を登録した感情判断知識ベースを基に処理が行われる。しかし、感情判断システムに入力される主体語、目的語、変化語の数、ならびにそれらの組み合わせの数は膨大であり、すべてを感情判断知識ベースに登録することは現実的ではない。そこで、感情判断知識ベースには各入力要素について代表的な語（以下、代表語と呼ぶ）をいくつかのカテゴリに分類し登録している。そして、感情判断を行う際には、各要素のカテゴリを用いて処理する。例えば、「泥棒がお金を盗る」と入力された場合、適切な自然言語処理の結果、主体語「泥棒」、目的語「お金」、変化語「盗る」が得られる。目的語「お金」は目的語のカテゴリ「貴重な」、変化語「盗る」は変化語のカテゴリ「目的語依存型動詞（継承）」に分類され、発話文の主体の感情は「貴重な」+「継承」=「喜び」と判断される。また、精神的距離（好き嫌い）、感覚的距離（身近さ）、社会的尺度（善悪）を各ABCの3段階で表現する主体語のカテゴリにおいて、主体語「泥棒」は「C, B, C」に分類されるため、発話者の感情は「怒り」、「悲しみ」と判断される（図2参照）。

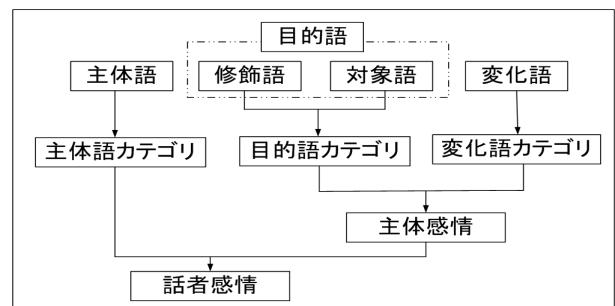


図2. 感情判断システムにおける処理の概略図

なお、各要素に入力される語の多くは、感情判断知識ベースに代表語として登録されていない。そこで、感情判断知識ベースに登録されていない未知の語（以下、未知語と呼ぶ）を代表語に関連付け、未知語を既知語とし

†同志社大学大学院工学研究科
Graduate School of Engineering, Doshisha University
‡三洋電機株式会社 DS 技術開発センターBU
Digital Systems Development Center BU, Sanyo Electric Co., Ltd.

て扱うこと (以下、未知語処理と呼ぶ) により対処している。具体的には、概念連想メカニズムである概念ベース[広瀬 02][小島 02]と関連度計算手法[渡部 01][井筒 02]を用いることにより、意味的な関連性やその強さの度合いを評価し、最も関連の強い代表語に関連付けている。また、目的語のカテゴリ分類において、感覚・知覚判断システム[Horiguchi 02][米谷 03]を用いている。

3. 概念連想メカニズム

3. 1. 概念ベース

概念ベースは、複数の電子化辞書から機械的に自動構築された大規模で汎用的なデータベースである。しかし、概念ベースは機械的に構築されているため、人間の感覚では必要な属性が抜け落ちていたり、明らかに不適切な属性が雑音として含まれていたりしており、必ずしも適切に必要なデータのみで構成されているわけではない。そこで、本研究では、不適切なデータを削除し、必要なデータを追加する自動精練処理を行った概念ベース (概念数約9万) [広瀬 02]を利用している。

概念ベースにおいて、任意の概念 A は、概念の意味特徴を現す属性 a_i と、この属性 a_i が概念 A を表す上でどれだけ重要かを表す重み w_i の対で表現される。概念 A の属性数を N 個とすると、概念 A は以下のように表せる。ここで、属性 a_i を概念 A の一次属性と呼ぶ。

$$A = \{(a_1, w_1), (a_2, w_2), \dots, (a_N, w_N)\}$$

概念 A の一次属性 a_i は概念ベースに定義されている概念としているため、 a_i から同様に属性を導くことができる。 a_i の属性 a_{ij} を概念 A の二次属性と呼ぶ。

3. 2. 関連度計算手法

概念 A, B は、その一次属性を a_i, b_j 、重みを u_i, v_j とし、属性がそれぞれ L 個、 M 個 ($L \leq M$) とすると

$$A = ((a_1, u_1), (a_2, u_2), \dots, (a_L, u_L))$$

$$B = ((b_1, v_1), (b_2, v_2), \dots, (b_M, v_M))$$

と表現できる。概念 A, B の一致度 $MatchWR(A, B)$ は、

$$MatchWR(A, B) = \sum_{a_i=b_j} \min(u_i, v_j)$$

(各概念の重みの総和は1に正規化する) と定義する。

関連度は、対象となる全ての一次属性の組み合わせについて一致度を計算し、一次属性どうしの対応を決定することにより計算する。具体的には、一致する一次属性どうしについては優先的にその対応を決定する。一致しない部分については、その一致度の合計が最大になるように一次属性どうしの対応を決定する。一致度を利用することによって、完全に一致しない一次属性についても関連の度合いを考慮に入れることができる。

一次属性どうしが完全一致するものがない場合、概念 A, B のうち属性数の少ない概念を $A (L \leq M)$ とし、概念 A の一次属性の並びを固定する。

$$A = ((a_1, u_1), (a_2, u_2), \dots, (a_L, u_L))$$

概念 B の各一次属性を対応する概念 A の各一次属性との一致度 ($MatchWR$) の合計が最大になるように並べ替える。

$$B_x = ((b_{x1}, v_{x1}), (b_{x2}, v_{x2}), \dots, (b_{xL}, v_{xL}))$$

このように対応を決めると概念 A, B の関連度 $MR(A, B)$ は、

$$MR(A, B) = \sum_{i=1}^L MatchWR(a_i, b_{xi}) \times (u_i + v_{xi}) \times \frac{1}{2} \times (\min(u_i, v_{xi}) / \max(u_i, v_{xi}))$$

となる。すなわち、関連度は対応する一次属性の一致度と、それらの属性の重みの平均および重みの比に比例すると考える。

一次属性どうしが完全一致するものがある場合 ($a_i=b_j$) は、別扱いにする。これは概念ベースには約9万の概念が存在し、属性が一致することは稀である。従って、属性の一致の扱いを別にするにより、属性が一致した場合を大きく評価するためである。具体的には、対応する属性の重み u_i, v_j の大きさを重みの小さい方にそろえる。このとき、重みの大きい方はその値から小さい方の重みを引き、もう一度、他の属性と対応をとることにする。例えば、 $a_i=b_j$ で $u_i=v_j+\alpha$ とすれば、対応が決定するのは (a_i, v_j) と (b_j, v_j) であり、 (a_i, α) はもう一度他の属性と対応させる。このように対応を決め、対応の取れた属性の組み合わせが T 個の場合、

$$A' = ((a'_1, u'_1), (a'_2, u'_2), \dots, (a'_T, u'_T))$$

$$B' = ((b'_1, v'_1), (b'_2, v'_2), \dots, (b'_T, v'_T))$$

となる。

概念 A, B の関連度 $MR(A, B)$ は

$$MR(A, B) = \sum_{i=1}^T MatchWR(a'_i, b'_i) \times (u'_i + v'_i) \times \frac{1}{2} \times (\min(u'_i, v'_i) / \max(u'_i, v'_i))$$

となる。

4. 感覚判断システムと知覚判断システム

感覚・知覚判断システムは、ある語に対して人間が常識的に抱く一般的な印象を判断するシステムであり、感覚・知覚に関する暗黙的な知識を読み取ることができる。本稿において「感覚」とは、人間が五感 (視覚・聴覚・嗅覚・味覚・触覚) で感じる印象 (ex.: 赤い, 丸い, うるさい, 臭い, 甘い, 痛い), 「知覚」とは人間が五感以外で感じる印象 (ex.: めでたい, 可愛い, 危ない) と定義する。

感覚・知覚判断システムを実現するためには、ある語とその特徴である感覚や知覚との関係を表す知識が必要である。しかし、2章の感情判断システムでも述べたように、すべての論理関係を記憶させることは困難であり効率が悪い。そこで、日常よく使われる語を代表語として判断知識ベースに登録し、代表語から想起される感覚・知覚を登録する (ex.: 「林檎」と「赤い」, 「夏」と「暑い」)。また、代表語を、シソーラス[NTT 97]を基に分類し、感覚と知覚を与え木構造化することによって、感覚と知覚の継承が可能となる (図3参照)。しかし、感覚・知覚判断システムにおいても、日常会話で使用される多くの語は判断知識ベースに登録された代表語ではないため、概念連想メカニズムを用い、意味的な関連性やその強さの度合いを評価し、最も関連の強い代表語を決定する未知語処理を行う必要がある。

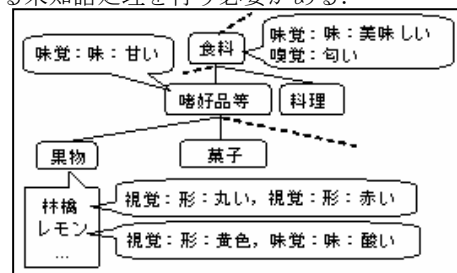


図3. 感覚・知覚判断システムにおける知識表現の例

5. 嗜好を考慮した感情判断システム

感情判断システムでは、目的語のカテゴリ分類に「感覚・知覚判断システム」を利用し、人間が常識的に想起する印象を基に「常識的な観点」で感情判断を行っている。そのため、ユーザごとに印象が異なるような対象語に対しては、感情判断を適切に行うことができない。例えば、辛い食べ物が苦手なユーザ（発話者）が「キムチを食べた」と発話した場合、発話者は「悲しみ」の感情を抱いていると思われる。しかし、従来の感情判断システムでは、「キムチ」=「食料」=「重要な」であり、結果として「喜び」と誤った感情を判断してしまう。そこで、ユーザ個人の嗜好の状況を考慮した感情判断（以下、嗜好感情判断と呼ぶ）を実現する方法を提案する。

5. 1. 嗜好学習

ユーザによって印象が異なる語に対しては、あらかじめユーザごとの印象（嗜好）をシステムが把握しておく必要がある。嗜好感情判断を行うために、発話内容からユーザの嗜好情報（どのユーザがどんな語に対してどんな印象を抱くかの情報）を取得する「嗜好学習」が不可欠である。

発話文の中には「私は甘い食べ物が好き」のようにユーザの嗜好を表現した文（以下、嗜好文と呼ぶ）がある。嗜好文であるかどうかは、変化語によって判断できる。

変化語が「名詞・形容詞・形容動詞」の場合、例えば「私は牛乳が嫌い」という発話文の変化語「嫌い」から目的語「牛乳」の目的語カテゴリが「嫌い」であると判断できる。

また、変化語が「動詞」の場合、例えば「私は牛乳を嫌う」という発話文から、変化語「嫌う」の変化語カテゴリが表す感情と同じ印象であると判断できる。このように、嗜好情報を蓄積し、嗜好感情判断を実現する。

5. 2. 概念マッチング処理

5. 1 節の嗜好学習により、嗜好文から取得した嗜好情報を基に嗜好感情判断を行うためには、入力された刺激語に嗜好学習した内容が含まれるかの判定をしなければならない。この判定を「概念マッチング処理」と呼ぶ。

概念マッチング処理は、単に取得した嗜好情報と入力された刺激語との表記一致を検索するだけでなく、語と語の関係や意味を考慮したマッチングを行わなければならない。以下では、3 種類の概念マッチング処理を提案する。なお、複数の概念マッチング処理を適応する場合は、節の若いものから順に行うものとする。

5. 2. 1. シソーラスを用いた概念マッチング

例えば、「花」に対して好意を持つユーザは、その集合の一つである「向日葵」や「蒲公英」に対しても好意を持つと考えることができる。このような判断には、シソーラス[NTT 97]を用いて概念マッチングを行うことができる。シソーラスでは、一般名詞の意味的用法を表す意味属性（ノード）の上位-下位関係、全体-部分関係を木構造で表現している。ノードの下には、それに属する名詞（リーフ）が登録されている。

具体的な処理では、入力された刺激語の上位ノードを参照し、そのノードが取得した嗜好情報と一致すれば、概念マッチングが成功したといえる。

5. 2. 2. 概念ベースを用いた概念マッチング

例えば、「私は運動が苦手」という嗜好文から目的語「運動」の嗜好情報を取得した場合を考える。「私はスポーツをする」という発話文があると、目的語「運動」と目的語「スポーツ」は意味的に極めて密接な関係であり、目的語「運動」に対する印象は目的語「スポーツ」に対するものと同じだと判断するのが自然である。しかしこれらは、シソーラスでは関係を表現できない。そこで、概念連想メカニズムを用いて概念間の意味的な関連性の強さを判断し、概念マッチングを行う。

具体的には、関連度計算手法を用い、取得した嗜好情報と刺激語との関連度を算出する。算出された関連度が、ある閾値以上であれば意味的な関連性が強いと判断し、概念マッチングの成立とする。

この閾値の設定には、関連度の平均値を用いる。概念ベースから無作為に抽出した概念 X に対して「女性-夫人」, 「山-丘」などの極めて密な関係, 「山-川」, 「夕焼け-赤い」などの密な関係, 「山-机」, 「電車-眼鏡」などの疎な関係にあるデータを人手で作成し、各関係における関連度の平均値を求める。概念 X と極めて密な関係, 密な関係, 疎な関係との関連度の平均は, 0.34, 0.13, 0.02 であり, 各関係間には関連度に十分有意な差が見られる。そこで, 極めて密な関係と密な関係の中間値である関連度の平均 0.23 を意味的な関連性の強さを判断する閾値として採用する。なお, 人手で作成したデータは, 4 人が 4 人共にその関係が正しいと判断したものの 200 セット (計 600 データ) を用いている。

5. 2. 3. 感覚判断を用いた概念マッチング

例えば「私は辛い食べ物が苦手」という修飾語を含んだ嗜好文があり「私はキムチを食べる」と発話された場合、キムチは「辛いもの」であり、かつ、「食べ物」であるということ認識しなければならない。つまり、語の感覚・知覚的な印象を獲得できなければ処理することができない。そこで、感覚・知覚判断システムを用いる。対象語に対して感覚・知覚判断を行い、取得した嗜好情報と一致した場合、概念マッチングが成功したとする。

6. 評価

評価方法はアンケートにより収集した評価用例句を用いて、システムが判断する感情を評価する。評価用例句は、主体語を「私」と限定し「嗜好文」と「その嗜好情報を刺激語として含んだ文」の 2 文を 1 セットとする。また、評価用例句のセットに対して、嗜好感情判断システムの結果として出力され得る主体の感情 10 種類について、「常識的（常識度）」(○), 「非常識（非常識度）」(×), 「非常識でない」(△) の 3 段階であらかじめ判断している。なお、評価用例句数は 300 文、計 150 セットを用意した。表 1 に評価用例句セットの一例を示す。

用意した評価用例句セットは、嗜好文とその嗜好情報を刺激語として含んだ文との対応関係から大きく 3 つに分類することができる。それは、単純な表記一致、全体部分の関係、語の意味や連想を考慮するパターンである（表 2 参照）。このことから「語の意味や連想を考慮した処理」が嗜好感情判断において重要であるといえる。

次に、嗜好感情判断システムにおいて、判断結果として出力される感情の常識度・非常識度の評価結果を示す。

表3に評価の際に使用した概念マッチング処理の種類を示し、それに基づく評価結果を図4に示す。

表1. 評価用例文セットの一例

番号	修飾語	対象語	変化語	喜び	安心	悲しみ	恐れ	落胆	...	なし
1		運動	嫌						...	
		野球	いする	×	×	○	×	○	...	×
2	甘い	物	好き						...	
		ケーキ	食べる	○	×	×	×	×	...	×

表2. 評価用例文データの内訳

対応パターン	例文1	例文2	評価用例文セットの割合
表記一致	隣人が嫌い	隣人に会う	15%
全体部分関係	運動が得意	野球をする	13%
語の意味や連想を考慮	肉が嫌い	ステーキを食べる	72%
	甘い物が好き	ケーキを食べる	

表3. 評価の際に使用した概念マッチング処理の種類

評価番号	概念マッチング処理
①	概念マッチングを用いない (嗜好情報知識を使わない)
②	「シソーラス」を用いた概念マッチング
③	「概念ベース」を用いた概念マッチング
④	「シソーラス」と「概念ベース」を用いた概念マッチング
⑤	すべての概念マッチングを用いる

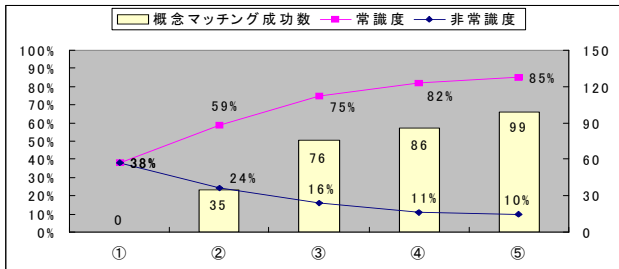


図4. 評価結果

嗜好状況を考慮しないシステムでは、判断された感情の常識度と非常常識度はともに 38%であり、②, ③, ④, ⑤と適応する概念マッチング処理を増やすことによって判断される感情の常識度ならびに概念マッチング数が増加している。すべての概念マッチング処理を組み合わせた⑤が最も良い結果となり、評価用例文セット全体の 85%に対して発話者 (主体) の感情を常識的に判断でき、150セット中 99 セットの評価用例文セットに対して概念マッチングが成功している。また、5. 2. 2節で述べた概念ベースを用いた概念マッチングが、嗜好感情判断において特に有効に機能しているといえる。

本稿で提案した概念マッチング処理を用いても非常識な感情を判断結果とした例は、「私はコーヒーが好き」→「私は喫茶店に行く」といった評価用例文セットである。これは、「喫茶店」には「コーヒー」があるという「場所とそこに存在する物の連想」が必要な例である。この

ような評価例文セットには、常識判断システムのうち場所判断システム[杉本 03]を利用すれば対応できる。本稿の嗜好感情判断システムでは、5. 2. 3節で述べたように感覚・知覚判断システムを利用した概念マッチングを行ったが、今後、場所判断システムや時間判断システム[野村 03]といった常識判断システムを利用し、さまざまな観点で連想を行うシステムに拡張する予定である。

7. おわりに

本稿では、人間が行う常識的な判断の一つである感情に関する判断について、ある概念から様々な概念を連想できる「概念連想メカニズム」を基に、ユーザの嗜好を考慮した感情判断システムを構築し、より人間に近い木目細やかな感情判断手法を提案し、その有効性を実験により評価した。

今後、場所判断システムや時間判断システムといった常識判断システムを利用し、さまざまな観点で連想を行うシステムに拡張することにより、より人間に近い判断が可能であると思われる。

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, 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.
- [篠原 02] 篠原宜道, 渡部広一, 河岡司: 常識判断に基づく会話意味理解方式, 言語処理学会第 8 回年次大会, B6-2, pp.651-654, 2002.
- [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.
- [小島 02] 小島一秀, 渡部広一, 河岡司: 連想システムのための概念ベース構成法—属性信頼度の考え方に基づく属性重みの決定, 自然言語処理, Vol.9, No.5, pp.93-110, 2002.
- [渡部 01] 渡部広一, 河岡司: 常識的判断のための概念間の関連度評価モデル, 自然言語処理, Vol.8, No.2, pp.39-54, 2001.
- [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, 3C1-07, 2003.
- [NTT 97] NTT コミュニケーション科学研究所監修: 日本語語彙体系, 岩波書店, 1997.
- [杉本 03] 杉本二郎, 渡部広一, 河岡司: 概念ベースを用いた常識場所判断システムの構築, 情報処理学会自然言語処理研究会資料, 153-11, pp.81-88, 2003.
- [野村 03] 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, 情報科学技術フォーラム FIT2003, E-047, pp.191-193, 2003.