

情緒推定用パターン辞書における情緒原因判断条件の改良

滝川晃司 徳久雅人 村上仁一 池原悟

鳥取大学工学部知能情報工学科

{s052032,tokuhisa,murakami,ikehara}@ike.tottori-u.ac.jp

1 はじめに

本稿では、情緒属性付き結合価パターン辞書を用いて、テキストからの情緒推定を行う。本辞書は、日本語語彙大系 [4] に収録されている約 14,000 件の結合価パターンに、情緒が生起する原因への関連性の情報を追加して構築された辞書である [1]。さらに、情緒原因の判定を用言の語義のみに頼っていたため、情緒原因の特徴の一部しか保証されていなかった問題に対して、不足する情緒原因の特徴を補うために、本辞書に「判断条件」が追加された [2]。しかし、判断条件は、情緒主とある事物との間の関係性を示しているが、両者の関係の方向性、すなわち、「接近」と「解離」は示されていなかった。そこで本稿では、「接近」と「解離」を示すように、判断条件の改良を行う。

2 情緒属性付き結合価パターン辞書

これまでの取り組み [2] で作成された辞書を紹介する。

本辞書は、日本語語彙大系の結合価パターンに情緒属性を付与して構築された。入力文と結合価パターンがマッチすることで、入力文の用言の語義が解析できる。その語義の表す事態が情緒の生じる原因と言えるとき、入力文から情緒が推定される。

ここで、パターンに付与されている情緒属性は、情緒名、情緒原因、情緒主、情緒対象、および、判断条件である。判断条件は、「生理」、「心理」、「目標実現」、「対人関係」、「対人関係・近」、「対人関係・離」、「対人関係(上下)・近」、および、「その他」の 8 種類が使用されている (図 3 の下線部を除く部分)。

具体例を図 1 に示す。獲得 は《喜び》の生じる原因の一つである。獲得 は単なるラベル名であり、具体的には「目標実現に必要な物事を努力して手に入れた」という特徴を指す。図 1 のパターンと意味属性制約を用いることで語義が解析されるのだが、その語義は下線部しか対応していない。そこで、下線部以外の要因を大まかにカバーするように判断条件「目標実現 (N1, N2)」が追加された [2]。

結合価パターン: N1 が N2 を N3 に/から/より貰う
意味属性制約: N1(主体)N2(*)N3(主体)
判断条件: 目標実現 (N1, N2)
情緒原因: 獲得
情緒主: N1 情緒対象: N2 情緒名: 《喜び》

図 1 情緒属性付き結合価パターン辞書のレコードの一例

判断条件の示す述語は、与えられる文の外の情報に基づき真偽が問われる。例えば、「子供がおもちゃを貰った。」という文において、判断条件は「目標実現 (子供, おもちゃ)」と単一化される。この文において《喜び》が推定されるには、この述語を「子供はおもちゃを必要としており、強く手に入れようとする傾向がある」と暗に解

釈し、この傾向の強さが常識的に認められる必要がある。

こうした常識は、人間の場合、日常生活で養われている。しかし、計算機の場合、常識を手作業で数値化したり、ブログなどから自動収集したりすることで処理することになる。

3 判断条件の改良

3.1 判断条件の問題点

従来の判断条件は、「対人関係」においてのみ、二者の「接近」と「解離」という方向性を区別するために、「対人関係・近」と「対人関係・離」が設けられていた。しかし、「生理」と「目標実現」においては、その区別がなかった。そのため、判断条件の運用に矛盾が生じるという問題がある。

例えば、「子供がおもちゃを貰った。子供がおもちゃを壊した。」という文章において、1 文目は図 1 のパターンに適合し、2 文目は図 2 のパターンに適合する。

この子供がおもちゃを欲しがっていたとすると、「目標実現 (子供, おもちゃ)」が成立し、1 文目からは《喜び》が推定される。しかし、2 文目においても判断条件が成立するので、「欲しがっていたおもちゃを壊して喜んだ」という推定になってしまい、2 文目の推定は誤りになってしまう。

図 2 のパターンが有効なのは、例えば「子供が気に入らないおもちゃを壊す」のように、不要なものを対象とする場合である。したがって、2 文目に対しては「《喜び》とは言えない」と出力したい。

したがって、図 1 の判断条件の場合、《喜び》が推定されるためには、「子供」と「おもちゃ」が「接近」の方向性であることを確認し、図 2 の判断条件の場合、《喜び》が推定されるためには、「子供」と「おもちゃ」が「解離」の方向性であることを確認する必要がある。

結合価パターン: N1 が N2 を壊す
意味属性制約: N1(*)N2(具体物)
判断条件: 目標実現 (N1, N2)
情緒原因: 完遂
情緒主: N1 情緒対象: N2 情緒名: 《喜び》

図 2 判断条件が問題となる例

3.2 判断条件の追加

図 3 の下線部は、本稿で新しく追加する判断条件の種類である。「生理」と「目標実現」の後に「近」と「離」を追加し、「接近」と「解離」を確認できるようにする。「不要」は、結合価パターンの語義だけで、ほぼ情緒原因の特徴を示す場合に付与する。

図 1 と図 2 のレコードを改良した結果を図 4 と図 5 に示す。

生理
生理・近
生理・離
心理
目標実現
目標実現・近
目標実現・離
対人関係
対人関係・近
対人関係・離
対人関係(上下)・近
その他
不要

図3 判断条件の種類

結合価パターン：N1がN2をN3に/から/より貰う
意味属性制約：N1(主体)N2(*)N3(主体)
判断条件：目標実現・近(N1,N2)
情緒原因：獲得
情緒主：N1 情緒対象：N2 情緒名：《喜び》

図4 目標実現・近に改良したレコード例

結合価パターン：N1がN2を壊す
意味属性制約：N1(*)N2(具体物)
判断条件：目標実現・離(N1,N2)
情緒原因：完遂
情緒主：N1 情緒対象：N2 情緒名：《喜び》

図5 目標実現・離に改良したレコード例

3.3 修正作業

本稿では、文献 [2] で「目標実現」が付与された 5,932 件に対して、手作業で修正を行う。作業の基準を以下に示す。

1. 両者の関係の方向性が「接近」であることが、情緒の成立する前提であるならば「目標実現・近」と修正する。
2. 両者の関係の方向性が「解離」であることが、情緒の成立する前提であるならば「目標実現・離」と修正する。
3. 結合価パターンの用言のみで解析される語義が情緒原因を示している場合は判断条件を「不要」とする。

3.3.1 目標実現・近への修正

図6のレコードの場合、《喜び》を推定するためには、N1とN2の方向性が「接近」であることが前提となる。よって、判断条件を「目標実現・近」へ修正する。

図6で情緒の推定される例文としては、「サラリーマンが給料を受け取る」が挙げられる。「サラリーマン」と「給料」は「接近」の関係にあると言えるからである。

結合価パターン：N1がN2をN3から/より受け取る
意味属性制約：N1(主体)N2(具体物 抽象物 精神)N3(主体)
判断条件：目標実現・近(N1,N2)
情緒原因：獲得
情緒主：N1 情緒対象：N2 情緒名：《喜び》

図6 目標実現・近へ修正したレコード例

3.3.2 目標実現・離への修正例

図7のレコードの場合、《喜び》を推定するためには、N1とN2の方向性が「解離」であることが前提となる。よって、判断条件を「目標実現・離」へ修正する。

図7で情緒が推定される例文としては、「彼が不安を払拭する」が挙げられる。「彼」と「不安」は「解離」の

関係にあると言えるからである。

結合価パターン：N1がN2を払拭する
意味属性制約：N1(主体)N2(抽象)
判断条件：目標実現・離(N1,N2)
情緒原因：完遂
情緒主：N1 情緒対象：N2 情緒名：《喜び》

図7 目標実現・離へ修正したレコード例

3.3.3 不要とした例

図8の結合価パターンの用言(慣用句)「腕が上がる」は、情緒原因 有効 をほぼ示していることから、判断条件を「不要」とする。

図8で情緒が推定される例文としては、「彼は仕事の腕が上がる」が挙げられる。

結合価パターン：N1はN2に/の腕が上がる
意味属性制約：N1(主体)N2(抽象物 人間活動)
判断条件：不要
情緒原因：有効
情緒主：N1 情緒対象：N2 情緒名：《喜び》

図8 不要へ修正したレコード例

3.4 修正結果

「目標実現・近」が 2,727 件、「目標実現・離」が 513 件、「不要」が 1,916 件となった。以下に用言(慣用句は格要素を併記)の具体例と情緒を示す。

「目標実現・近」に修正した用言
貰う(喜び) 受け取る(喜び) 落とす(悲しみ)
実践する(期待) 溜める(喜び) 届く(喜び)

「目標実現・離」に修正した用言
壊す(喜び) 伝染する(悲しみ) 巻き込む(嫌だ)
取る(嫌だ) 嵌まる(嫌だ) 引き離す(怒り)

「不要」に修正した用言
腕が上がる(喜び) 雲行きが悪い(恐れ) 欺く(嫌だ)
ピンチだ(恐れ) 無事だ(好ましい) 不自由する(嫌だ)

4 判断条件の運用

第4.1節では、判断条件を用いた情緒推定を行うために、ブログから、判断条件の示す二者の関係についての知識ベースを構築する。第4.2節では、この知識ベースを用いた情緒推定の手順を示す。

4.1 判断情報知識ベースの構築

判断条件を用いて情緒推定を自動的に行うためには、判断条件が示す二者の関係をあらかじめ学習させておく必要がある。そこで、2008年8月から2008年10月分のブログの本文データ(13.1GB)を学習用データとして使用し、判断条件の示す二者の関係を学習させた知識ベースを構築する。この知識ベースを「判断情報知識ベース」と呼ぶこととする。

本知識ベースの構成要素を以下に示す。

1. 判断条件名
2. 判断条件の第一引数の格要素
3. 判断条件の第二引数の格要素
4. ブログ文にマッチしたレコードの情緒に同意できた割合
5. ブログ文にマッチしたレコードの情緒に同意できなかった割合

6. プログ文の情緒が判断できなかった割合

4~6の各割合は4~6の各条件に当てはまるプログ文数 / 4~6の各条件に当てはまるプログ文数の和で求める。

4.1.1 判断情報知識ベースの構築方法

以下の手順で判断情報知識ベースを構築する。

1. 方向性を知りたい二つの格要素を指定する。なお、第一引数が一般的な人である場合は ϕ とし、省略できることとする。
2. プログデータより、1で指定した格要素が含まれている文を抽出する。抽出の際、短すぎる文や長すぎる文を除くために、20byteから200byteの文字数制限をかける。
3. 抽出した文に対して、本辞書とのパターンマッチングを行う。結合価パターンに最もふさわしくマッチし、かつそのパターンのレコードに付与されている判断条件が「目標実現」である文を抽出する。
4. 判断条件「目標実現」の引数に、1で指定した格要素が含まれている文を抽出する。
5. 4で抽出した文に対して、マッチしたレコードの情緒に同意できるかの判定を手作業で行う。マッチしたレコードの情緒が明らかに成立しているとは同意できる場合は T 、明らかに成立していない場合は F 、常識的に考えて、どちらとも言える場合は T と F 、情緒を示す情報が不足している場合は $?$ とカウントし、結果を判断情報知識ベースへ蓄積する。

4.1.2 判断情報知識ベースの構築の様子

以下に、第4.1.1節の作業の様子を示す。

- 図9の入力文の場合、「渋滞にはまっ」に、「 $N1$ が $N2$ に嵌まる」のパターンが意味属性制約を満たしながら適合する。出力される情緒は《嫌だ》であるが、「渋滞にはまって動けない」の「動けない」から、情緒《嫌だ》という出力に、人手の判断で同意することは自然であるため、 T とカウントする。

入力文：途中から渋滞にはまってしまい動かない
 適合パターン： $N1$ が $N2$ に嵌まる
 意味属性制約： $N1$ (主体)動物乗り物) $N2$ (場所 場 計器...)
 判断条件：目標実現・離 (ϕ , 渋滞)
 情緒原因：制限
 情緒主： ϕ 情緒対象：渋滞 情緒名：《嫌だ》

図9 入力文の情緒が成立していると推定した例

- 図10の入力文の場合、「後悔した」から、《喜び》が生起しているとは言い難いため F とカウントする。

入力文：携帯番号を交換したのを、後悔した。
 適合パターン： $N1$ が $N2$ を $N3$ と交換する
 意味属性制約： $N1$ (主体) $N2$ (*) $N3$ (主体)
 判断条件：目標実現・近 (ϕ , 携帯番号)
 情緒原因：獲得
 情緒主： ϕ 情緒対象：携帯番号 情緒名：《喜び》

図10 入力文の情緒が成立していないと推定した例

- 図11の入力文の場合、「障害物競争」に出場することに対して、《期待》が生起しているか、していない

かどちらとも判断できるため T と F の両方をカウントする。

入力文：障害物競争に出場したのですが、
 適合パターン： $N1$ が $N2$ に出場する
 意味属性制約： $N1$ (主体) $N2$ (競争 スポーツ)
 判断条件：目標実現・近 (ϕ , 障害物競争)
 情緒原因：成算
 情緒主： ϕ 情緒対象：障害物競争 情緒名：《期待》

図11 入力文の情緒がどちらとも言えると推定した例

- 図12の入力文の場合、携帯電話を取り出した人間の情緒を、この文から推定するには、情報が不足しているため？とする。

入力文：もうすぐ入院するからとケータイを取り出し
 適合パターン： $N1$ が $N2$ を $N3$ から/より取り出す
 意味属性制約： $N1$ (人) $N2$ (具体物) $N3$ (具体物)
 判断条件：目標実現・近 (ϕ , ケータイ)
 情緒原因：獲得
 情緒主： ϕ 情緒対象：ケータイ 情緒名：《喜び》

図12 入力文に情緒を示す情報が不足している例

4.1.3 判断情報知識ベースの構築結果

第4.1.1節の手順に従って構築した知識ベースを表1に示す。

表1 判断情報知識ベース

#	第一引数	第二引数	判断ラベル	T	F	?
1	ϕ	ツバメ	目標実現・近	73	27	0
2	ϕ	キズ	目標実現・近	18	82	0
3	ϕ	障害物競争	目標実現・近	43	46	11
4	ϕ	テスト	目標実現・近	61	39	0
5	ϕ	渋滞	目標実現・離	84	12	4
6	ϕ	実技試験	目標実現・近	100	0	0
7	ϕ	彼	目標実現・近	62	38	0
8	ϕ	携帯番号	目標実現・近	52	42	4

$T, F, ?$ の単位は%

4.2 判断情報知識ベースを利用した情緒推定

情緒推定は以下の手順で行う。

1. テキストと結合価パターンを照合し、引数に具体的な格要素の入った判断条件を得る。
2. 1で得た判断条件と引数をキーとして、判断情報知識ベースを参照する。
3. $T > F$ ならば情緒が成立していると判断し、パターンに付与されている情緒を出力する。 $T < F$ ならば情緒が不成立であると判断し、情緒《なし》を出力する*1。

5 情緒推定実験

5.1 実験の目的

判断条件を用いた情緒推定の有効性を評価する。実験には、情緒タグ付き日記コーパス [5] を使用する。本コーパスは、文献 [6] の日本語部分を、節の単位で収録している。節ごとに情緒タグが付与されている。付与者は情緒推定経験者6名である。

本コーパスに収録されている節に対して、判断条件を用いない情緒推定を行った際、情緒タグ通りの情緒を推定できなかった節が幾つか存在した。その節に対して、判断条件を用いた情緒推定で、情緒タグ通りの情緒を推定

*1 正確に言えば、付与されている情緒が成立するとは言えないと出力するべきである

できるか実験を行う。

また、判断条件を用いない情緒推定でも、情緒タグ通りの情緒を推定できた節に対して、判断条件を用いた情緒推定で、情緒タグ通りの情緒を出力できるか実験を行う。

5.2 実験方法

第 4.2 節の手順に従って、情緒推定を行い、本コーパスの情緒タグと比較を行う。

実験対象は、判断条件未使用時の情緒推定で、本コーパスの情緒タグと一致しなかった 36 節の中から以下の条件を満たす 7 節、および、判断条件未使用時の情緒推定でも情緒タグと一致した 89 節の中から以下の条件を満たす 6 節の、計 13 節を対象とする。

1. 結合価パターン辞書のパターンとマッチする。
2. マッチしたパターンのレコードには「目標実現」が付与されている。

5.3 実験結果

正答率 を 正答数 / 節数 で求める。

判断条件未使用時の情緒推定の正答率は 46%(=6/13) である。判断条件使用時の情緒推定の正答数は、さらに 3 節増加した。よって、正答率は 69%(=9/13) となった。情緒タグと一致した 3 節を以下に示す。

表 2 情緒タグと一致した節と情緒推定結果

#	節	情緒タグ	情緒推定結果
1	車のサイドにキズを発見し	驚き	驚き
2	障害物競争に参加する	なし	なし
3	携帯を落とす	悲しみ	悲しみ

6 考察

判断条件を使用した情緒推定において、本コーパスの情緒タグと一致しなかった節について考察する。表 3 に節と情緒推定結果を示す。

表 3 情緒タグと一致しなかった節と情緒推定結果

#	節	情緒タグ	情緒推定結果
1	ツバメを観察するのが	なし	期待・恐れ
2	運転免許の実技試験を受けた	なし	期待・恐れ
3	歴史のテストを受けた	嫌だ	期待・恐れ
4	いつもの店で彼と待ち合わせる	なし	期待

(1) 情緒推定結果が一概に悪いとは言えない (#1, #2)

#1 の節にマッチしたパターン「N1 が N2 を観察する」のレコードより「目標実現 (φ, ツバメ)」を得る。また、#2 の節にマッチしたパターン「N1 が N2 を受けた」のレコードより「目標実現・近 (φ, 実技試験)」を得る。知識ベースを参照すると、共に $T > F$ となっており、レコードに付与されている情緒が成立することを示している。よって、情緒《期待》、《恐れ》が出力された。結果として、本コーパスの情緒タグとは一致しなかったが、方向性が「接近」である「ツバメ」を観察する際や、「運転免許の実技試験」のような自ら希望して試験を受ける際は、《期待》や《恐れ》が生起することは自然である。

したがって、#1 と#2 の情緒推定結果が、一概に悪い結果であるとは言えない。

(2) 情緒属性の追加、および、判断条件の改良が必要 (#3)

#3 の節にマッチしたパターン「N1 が N2 を受けた」のレコードより、「目標実現・近 (φ, テスト)」を得る。

本知識ベースを参照すると、共に $T > F$ となっており、レコードに付与されている情緒が成立することを示している。よって、情緒《期待》、《恐れ》が出力された。

ここで、「歴史のテスト」は、#2 と違い、自ら希望して受けているテストではないと考えられる。このようなテストは、「接近」の関係ではなく「解離」の関係であり、情緒も《期待》、《恐れ》ではなく、《嫌だ》が出力されるべきである。

しかし、現在本辞書において、「N1 が N2 を受けた」のパターンのレコードには、「目標実現・離」や、情緒《嫌だ》の付与が見られない。そこで、今後二つの改良が必要となる。

一つは、本辞書において「目標実現・近」が付与されているレコードに対し、「目標実現・離」ならばどういった情緒が生起するのかを調べ、見落とししている情緒がないか調査を行うことである。

もう一つは、#2 の「運転免許」や、#3 の「歴史」も情緒推定の考慮に入れられるような、判断条件の拡張を行うことである。

(3) N1, N2 以外も考慮に入れるべき (#4)

#4 の節にマッチしたパターン「N1 が N2 と N3 で待ち合わせる」のレコードより、「目標実現 (φ, 彼)」を得る。本知識ベースを参照すると、 $T > F$ となっており、レコードに付与されている情緒が成立していることを示している。よって、情緒《期待》が出力された。

「彼と待ち合わせる」ことに対して、《期待》が生起するのは自然であると考えられるが、本コーパスの情緒タグに《なし》が付与されている。これは、「いつもの店」から、待ち合わせる本人にとって、特に特別な行動ではないと判断したためと考えられる。本稿の情緒推定では、「いつもの」のような表現は加味していないため、今後、このような表現も加味する情緒推定が必要である。

7 おわりに

本稿では、判断条件に両者の関係の方向性を示す情報を追加した。そして、判断条件を利用する情緒推定が、情緒推定精度の向上に繋がることを示した。

今後の課題は、本辞書内で見落とされている情緒の調査や、判断条件の改良、そして、N1, N2 以外も考慮に入れる情緒推定の実装、ならびに、判断情報知識ベースの自動構築である。

参考文献

- [1] 田中努, 徳久雅人, 村上仁一, 池原悟: “結合価パターンへの情緒生起情報の付与”, 言語処理学会第 10 回年次大会発表論文集, pp.345-348, 2004.
- [2] 吾郷裕昭, 徳久雅人, 村上仁一, 池原悟: “情緒生起原因を表す結合価パターン辞書への判断条件の補強”, 電子情報通信学会 2008 年総大会講演論文集, 基礎・境界, A-13-1, p.232, 2008.
- [3] 徳久雅人, 岡田直之: “パターン理解の手法に基づく知能エージェントの情緒生起”, 情報処理学会論文誌, 39(8), pp.2440-2451, 1998.
- [4] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦: “日本語語彙大系”, 岩波書店, 1997.
- [5] 徳久雅人, 松浦大樹, 村上仁一, 池原悟: “情緒タグ付き日記コーパスの作成と評価”, 第 5 回情報科学技術フォーラム, 一般講演論文集, 第 2 分冊, pp.235-236, 2006.
- [6] 石原真弓: “英語で日記を書いてみる 表現集編”, ベレ出版, 2002.