

テキストを対象とした情緒推定のための 事象間関係の解析に向けて

古塩貴行 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

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

1 はじめに

言語表現からの情緒推定の1つの方法として、トップダウン的に心的過程を追跡する方法が提案されている [1]。この方法は、入力文の「語義」に基づき情緒を推定する背後で、情緒の生起するための「前提条件」の成立の確認を行うものである。前提条件は、「目標」や「プラン」などの心的構成素の関係に対する条件である。[2]では、語義、情緒、および、前提条件の対応関係を結合価パターン辞書として定義している。しかし、「前提条件」の成立を確認する手順が示されていなかった。そこで、本稿では、前提条件の成立を確認する手順の確立に向けて検討を行うことを目的とする。

従来の意味解析において、情緒主の心的構成素は、各文において解析されうるが、断片的である。前提条件の成立の確認とは、断片的な心的構成素を結び付けることができるか否かを文脈情報や世界知識を介して判定することである。心的構成素の言語表現は事象の表現である。心的構成素の結び付きは、事象と事象の関係である。こうした事象間関係は一般的に因果関係や時間的経緯など様々な関係がある。ここで、因果関係知識を重文・複文の用例集から節間キーワード（接続標識「ため」）に基づきテキストから自動獲得する手法が [3] により示されている。

したがって、前提条件成立の確認手順の確立における問題点は以下の4点が挙げられる。1) 前提条件を列挙すること、2) 用例集を構築すること、3) 前提条件ごとに注目すべき節間キーワードを示すこと、4) 用例集を利用した前提条件の確認を文脈に応じて運用すること。本稿では、例題を用いて上記の問題を検討する。

2 情緒生起の前提条件

前提条件は、情緒生起の原因事態の特徴と深く関係する。本章では、情緒生起の原因事態の特徴を示した後、前提条件の対応関係を示す。

2.1 情緒生起の原因

[2]では、《喜び》/《悲しみ》、《好ましい》/《嫌だ》、《驚き》、《期待》、《恐れ》、《怒り》の8種類の情緒を対象としている。これらの情緒の生起する原因事態の特徴は、階層的に分析されている。大きく〈生理的〉と〈心理的〉に分かれ、〈心理的〉はさらに〈目標実現〉と〈対人関係〉に分かれる。それぞれさらに具体的な特徴に分解される。そして最も具体的な特徴として、〈獲得〉、〈賞賛〉など約120種類に分解される。

(喜び：現状態は前状態よりも好都合である
(生理的 (内的な快； 外的な快))
(心理的
(目標実現
(情報収集 (思惑通り)； 発見)； 判明)
計画 (立案)
実行結果 (獲得； 完遂； 有効)))
(対人関係
(仲間意識 (同意； 同感； 協力； 仲直り)
(優劣関係 (優越； 賞賛； 服従； 保護)))
その他)

図1: 《喜び》の生起特徴フレーム

《喜び》の情緒についての階層性は、図1のようにフレームで表現される。〈目標実現〉や〈実行結果〉という中位の特徴を、〈獲得〉や〈完遂〉という最下位の特徴は継承している。すなわち、「目標実現に有効な物を獲得した」という意味の〈獲得〉の事態が原因となって、《喜び》が生起する。

[2]の結合価パターンには、情緒の種類と原因などの「情緒属性」が対応づけられている。図2に一例を示す。たとえば、「太郎にマウンテンバイクが当たった」という文は、図2の結合価パターンに適合する。文外の情報から、前提条件の成立が確認できると、太郎には、〈獲得〉による《喜び》が生起していることが推定できる。

結合価パターン：
用言意味属性： 所有的移動
パターン： N1がN2にN3で当たる
意味属性制約： N1(景品), N2(人), N3(褒賞)
情緒属性：
情緒主： N2, 情緒対象： N1
情緒名： 《喜び》, 生起原因： 〈獲得〉
前提条件：
条件1： [N2が目標Gを持つ]
条件2： [目標GはプランPで実現可能]
条件3： [プランPにはN1が必要]
条件4： [N1の入手は評価Eが高い]

図2: 情緒注釈付き結合価パターンの例

2.2 前提条件の対応関係

情緒属性付き結合価パターンは、特徴フレームの階層性に着目して、効率的に前提条件の各条件が付与されている。現段階では、前提条件が図3のように試作されている。紙面の都合で中位特徴名毎に整理したが、生起原因に対応関係がある。たとえば、*印の条件は図2で示したように〈獲得〉に対応する。

記述内容を見ると抽象的な「評価が高い」から具体的な「生理状態を改善」まで異なる粒度の条件がある。これは生起原因に応じて使い分けられる。

<p>〈生理的〉に関する前提条件</p> <p>〔五感に関する状態が悪化〕</p> <p>〔内的/外的な生理状態 P_h が改善/悪化〕</p> <p>〈目標実現〉に関する前提条件</p> <p>〔情緒主 F_r が目標 G を持つ〕* 〔情緒主 F_r がプラン P を持つ〕 〔情緒主 F_r が行動 A_c を実行〕 〔目標 G はプラン P で実現可能〕* 〔プラン P の評価 E が高い/低い〕 〔プラン P は正当/不当〕 〔プラン P には対象 N が必要〕* 〔プラン P に向けて行動 A_c を実行〕 〔対象 N の入手は評価 E が高い/低い〕* 〔対象 N は目標 G の達成に必要な〕 〔対象 N は目標 G の達成を阻害〕</p> <p>〈対人関係〉に関する前提条件</p> <p>〔集団 G_r が存在〕 〔他者 P_r が存在〕 〔他者 P_r は信頼可能〕 〔他者 P_r は情緒主 F_r と仲よし/不仲〕 〔他者 P_r は情緒主 F_r より上位/下位〕 〔他者 P_r は行動 A_c を高く/低く評価 E 〕 〔他者 P_r は行動 A_c を実行〕</p> <p>全般に関する前提条件</p> <p>〔行動 A_c の評価 E が非常に高い/高い/低い〕 〔行動 A_c は情緒主 F_r に有益/無益〕 〔行動 A_c は情緒主 F_r の状態 P_s を改善/改悪〕 〔行動 A_c は内的/外的な生理状態を改善〕 〔行動 A_c は目標 G に有益/無益〕 〔行動 A_c は目標 G の達成を阻害〕 〔行動 A_c はプラン P に有益/無益〕 〔行動 A_c はプラン P の達成を阻害〕</p>
--

図 3: 前提条件の種類

前提条件の成立を確認することを、事象間関係の解析として実行する。そのためには、重文・複文の用例集から適合する表現を検索する必要がある。

たとえば、ある情緒主が「夏休みの宿題をする」という目標を持ち、「図書館へ行く」というプランを持つことまでは既知であるとする。情緒推定にはこれらの関係を確認する必要がある。これは条件〔目標 G はプラン P で実現可能〕を確認することであり、次の構造の重文・複文である用例集から検索することで実現できるだろう。

$CL1_{\text{目標}} \text{ために} CL2_{\text{プラン}}$
 $\Leftrightarrow \text{夏休みの宿題をする}_{CL1} \text{ために} \text{図書館へ行く}_{CL2}$

ここで、 CL は節変数である。「ために」など接続詞あるいは接続詞相当表現（たとえば形式名詞 + 助詞）を本稿では「節間キーワード」と呼び、節間キーワードより前方を「事象 1」、後方の節を「事象 2」と呼ぶ。

さて、上述の手順の実現には、次の問題がある。

- (i) 用例の収集
- (ii) 前提条件に対応する節間キーワードの選択
- (iii) 事象 1,2 の形式への書き換え
- (iv) 前提条件の確認の手順の運用方法

(i) は第 3 章で述べ、(ii) は第 4 章で述べる。(iii) は今後の課題とする。(iv) は第 5 章で述べる。

3 事象間関係を表す用例の収集

事象間関係を解析する根拠とする表現文を単に用例と呼ぶ。本章では用例収集のもととなるコーパス、および、抽出結果を示す。

3.1 重文・複文のコーパス

事象間の関係を、節間キーワードの語義とみなす。上記の例では「ために」は、 $CL1$ が「目標」を表し $CL2$ が「プラン」を表すので、〔目標→プラン〕という語義とみなす。節間キーワードには多義があり「不正をしたために倒産した」のように〔行動→行動〕*1という語義もある。

そこで、重文・複文を約 15 万文収録している日英対訳コーパス [4] から用例を収集する。本来このコーパスは日英機械翻訳や言い換え技術の開発のために構築が進められたものである。上述の節間キーワードが約 150 種類に分類され、各用例に分類コードが付与されている。将来的にこうしたコードを利用して節間キーワードの多義の問題が回避できるので、本コーパスを対象に事象間関係の用例を収集する。

3.2 用例の抽出と DB 化

重文・複文のコーパスには、2 つの節で構成される文と、3 つの節で構成される文がある。本稿では前者を対象に用例を抽出する。そのため、98,999 文が抽出可能な文となる。

まず、抽出した用例は、節間キーワードの前後で、事象 1 と事象 2 に分ける。次に事象検索のヒット率を高めるために事象 1 および事象 2 をそれぞれ 1 つの格要素と述語の組（分割事象と呼ぶ）として分割する。最後に、事象 1 および事象 2 からそれぞれ得た分割事象を節間キーワードで組み合わせ（分割事象対と呼ぶ）、事象間関係 DB に登録する。なお、分割事象対は、事象 1・2 の順序について正順と逆順の 2 通り用意する。

以下に具体例を示す。

- 1) 抽出用例: 「理科の懸賞作品募集に応募して、彼女は特別賞を得た。」
 - 2) 事象 1: 「理科の懸賞作品募集に応募する」
事象 2: 「彼女は特別賞を得た」
 - 3) 分割事象 1: 「理科の懸賞作品募集に-応募する」
分割事象 2a: 「彼女は-得た」
分割事象 2b: 「特別賞を-得た」
 - 4) DB1: 『理科の懸賞作品募集に-応募する』
→ 『～て、彼女は-得た』
DB2: 『理科の懸賞作品募集に-応募する』
→ 『～て、特別賞を-得た』
DB3: 『彼女は-得た』
→ 『て～、理科の懸賞作品募集に-応募する』
DB4: 『特別賞を-得た』
→ 『て～、理科の懸賞作品募集に-応募する』
- 用例 1 つから分割事象対 4 件を DB に登録した。98,999 件文の用例から DB 化を行った結果、162,294 件の分割事象対を登録した。節間キーワードごとにラン

*1 ある行動からある行動が導かれている。

ダムで1)の用例を10件選択し(節間キーワード「の」は4件),各用例から作成した4)の分割事象対を登録内容から取得して,適切さを検査したところ,90%(229/255)は適切であった.適切さの検査項目には統語的な分割の成功の他に,意味的な項目も含まれる.

4 前提条件と節間キーワードの対応

用例集と前提条件の照合を行う手続きを確立するためには,「前提条件のタイプごとに,どのような節間キーワードを用いて用例を検索すればよいか」という問題を解決しなければならない.そこで,具体例で検討する.

4.1 節間キーワードを求めるための検索実験

日記の文集[5]から情緒推定対象文をランダムに選び結合価パターンを適合させ,前提条件を列挙する.各条件に対応する分割事象対を手で作成して,DBの検索をする.節間キーワードは任意としておき,検索結果から収集する.そして,検索結果を分析し,節間キーワードの利用可能性を評価する.すなわち,検索結果と日記の状況をゆるやかに解釈しながら比べて,次の成立を検査する.

ポイントA:分割事象の意味が状況に一致する

ポイントB:因果関係など節間キーワードの語義が一致する

ポイントA・Bの組み合わせから次の評価値を検索結果に与える.

○:AおよびBを満たす

△:Aを満たすが,Bを満たさない

×:Aを満たさない

?:判断保留

対象文数は15文,事象間関係DBの検索結果は1,674事象対となった.図4に,具体例を示す.

対象文:「理由を尋ねる」
 情緒候補:〈協力〉による《期待》
 前提条件
 条件1:[他者 P_r が存在]
 条件2:[他者 P_r は信頼可能]
 条件3:[行動 A_c は情緒主 F_r に有益]
 評価例
 条件2:[他者 P_r は信頼可能]
 語義:[他者→信頼]
 試作した分割事象対:「レジの人」→「信頼できる」
 検索結果
 ○の例:
 『この店の人たち_{他者}は明るい』→『～て,親切_{信頼}だ』
 ×の例:
 『店の人が無愛想だ』→『～ので,お客があまり来ない』

図4: 検索実験の様子

4.2 利用可能な節間キーワード

評価○となった検索結果について,前提条件と節間キーワードの対応関係を調べたところ,表1のようになった.15種類中12種類の節間キーワードと7種類の前提条件との対応を得ることができた.この対応より,たとえば

前提条件が「プラン P には対象 N が必要」のときには,節間キーワードが「て」、「と(条件)」、「ても」、「が」または「よう」である分割事象対を検索すればよいということが分かる.また,節間キーワード「て」はどの前提条件とも対応しており,かつ,他の節間キーワードに比べ正解数が多い一方で,「よう」、「こと」および「の」といった節間キーワードは前提条件との対応が得られなかった.この差には,事象間関係DBに対する節間キーワードごとの登録数の差が起因している.

5 前提条件の成立確認の手順について

前章により,前提条件と節間キーワードの関係に検討がついた.本章では情緒推定の場面を想定した例題を使い,前提条件の成立を確認する手順を考察する.

5.1 確認手順

次の手順で確認する.1)例題に用いるテキストから情緒推定の対象となる文を決める.2)結合価パターンを適合させ関連する前提条件を列挙する.3)前提条件の変数に,日記から分かる情報を代入する.不明な部分は変数のまま残す.4)前提条件から分割事象対を手作業で作成する.節間キーワードは第4章に従い代入する.5)事象間関係DBから4)の分割事象対と一致する登録内容を検索する.3)で変数として残されていた部分については,適合した用例の表現が代入される.6)検索結果を1つずつ確認し,適切な適合結果を残す.

5.2 例題検討

図5の日記を対象に検討を進める.

日記 数ヶ月前に雑誌の懸賞に応募した.今日,マウンテンバイクが当たったという通知が届いた.マジで?って感じ!今週末に配達してくれるということだった.めちゃくちゃ嬉しい!

図5: 日記の一例

- 1) 下線部を推定対象とする.
- 2) 結合価パターンは,図2が適合する.
- 3) 前提条件に日記の文脈情報を代入する.
 条件2のプランは,先行文脈より動作の文を代入した.条件4の評価は,後続の文脈より評価表現を代入した.その他は推定対象文の格要素を代入した.
 条件1:[(日記の著者)が目標 G を持つ]
 条件2:[目標 G は(懸賞に応募する)プランで実現可能]
 条件3:[(懸賞に応募する)プランには(マウンテンバイク)_{対象}が必要]
 条件4:[(マウンテンバイク)_{対象}の入手は(うれしい)_{評価}]
- 条件1は条件2から得る.条件2は事象間関係を確認する必要がある.条件3および4は成立している.
- 4) 分割事象対を作成する.
 条件2:『CL1』→『て～,懸賞に-応募する』
- 5) 事象間関係DBを検索する.
 r1:『彼女は-得た』

表 1: 前提条件と節間キーワードの対応の分布

前提条件 節間キーワードの語義	て	と 引用 ()	と 条件 ()	ば	ので	ても	が	ように	よう	こと	ことが	ことは	の	のが	のは	合計
【目標GはプランPで実現可能】 【目標→プラン】	40	3	2		5		2				1			2	1	56
【プランPには対象Nが必要】 【プラン→対象】	30		1			1	1	1								34
【対象Nの入手は評価Eが高い】 【プラン→評価】	25		6	5	3	2	8	1				1		3	3	57
【対象Nは目標Gの達成を阻害】 【対象→阻害】	13	1		1	1		1	2								19
【他者Prは信頼可能】 【他者→評価】	7				4		1					2				14
【行動Acは情緒主Frに有益】 【行動→評価】	7			2	3			2				1		2		17
【行動の価値Eが非常に高い】 【行動→評価】								2				1				3
合計	122	4	9	8	16	3	13	8			1	5		7	4	200

- 『て～, 理科の懸賞作品募集に-応募する』
 r2:『特別賞を得た』
 →『て～, 理科の懸賞作品募集に-応募する』
 r3:『落選した』
 →『て～, 懸賞小説に-応募する』

6) 検索結果を確認する.

- ・ r1 は妥当 (「当たる」の用言意味属性と一致).
- ・ r2 は用言は妥当だが格要素が不適當 (「特別賞」は文脈上にない).
- ・ r3 は不適當 (用言意味属性が不一致).

3') 条件 1 に代入する.

条件 1 は条件 2 より得た情報で満たされた.

条件 1 : [(日記の著者) が (- 得る)_{目標} を持つ]

以上により前提条件の全てが検索できた. 同様の方法で, 8 件の日記から選択した 15 件の推定対象文のうち, 10 件について前提条件の成立を確認できた.

6 考察

6.1 前提条件の不成立

前提条件の成立が確認できない場合があった. 用例集に存在しない分割事象対を検索しようとしたためである. 人間は何でも知っているという訳ではない. 検索に失敗したときは新しい情報であるかもしれない. 前提条件の不成立をどのように判断するかが問題である.

6.2 省略表現の補完

前提条件の確認の過程は, 文章の中の各文の省略表現を補うことが並行する. 図 5 の日記では, 「めちやくちや嬉しい」という表現がある. 嬉しさの対象の 1 つが, 「マウンテンバイク」であることを前提条件 4 が仮説として挙げている (条件 4 「評価 E」に「嬉しい」を代入). しかしその根拠に「嬉しい」の対象格の解析が先行するのかどうか, 検討の余地がある. この例題の場合, 「当たる」と「嬉しい」は情緒状態が一致して辻褄が合うので, 仮説を信じながらテキストの理解が進められる.

7 おわりに

テキストからの情緒推定というタスクは, 文脈情報や世界知識との関連が見逃せない. 心的過程をトップダウン的に理解することを基本的枠組とする情緒推定的方式において, 文脈情報や世界知識を推定対象と結び付ける手順の確立が課題である. その手順を実現するための要点は, 1) 用言と情緒の対応関係に対して, 文脈情報・世界知識との接点を設けること, 2) 世界知識の中で「事象間の関係」を見出して 1) の接点を通じて文脈情報や対象文に結び付けること, である.

本稿では, まず, 1) については [2] で試行された「前提条件」を使用し, 次に, 2) については [3] の考え方を参考にして, 重文・複文の用例集から事象間関係を DB 化を行い, そして, 「前提条件」と「事象間の関係」の結び付けの手順について例題を用いて検討した. 約 1,700 件の事象間の関係を参照し, 節間キーワード (接続標識) との適合性を検査した. そして, 8 件の日記から 10 件の推定対象文について前提条件の成立を確認できた. 結び付けの手順について, こうして 1 つの解決の道筋が示すことができた.

今後の課題は, テキスト中から仮説として前提条件に代入する表現を得る方法を開発すること, 前提条件と事象間関係の対応関係をより詳細に分析することである.

謝辞 本研究は科研費 (若手 (B):17700151) の下で行いました. 重文・複文のコーパスに関しては, 科学技術振興事業団における CREST 研究課題「セマンティックタイポロジーによる言語の等価変換と生成技術」の支援によります.

参考文献

[1] 徳久, 村上, 池原:情緒推定の手がかりについて, 言語処理学会第 11 回年次大会発表論文集, S3-2, pp.652-655, 2005.
 [2] 田中, 徳久, 村上, 池原:情緒生起情報付き結合価値パターン辞書の開発, 言語処理学会第 12 回年次大会発表論文集, E5-5, 2006 (予定).
 [3] 乾, 乾, 松本:接続標識「ため」に基づく文章集合からの因果関係知識の自動獲得, 情報処理学会論文誌, 45(3), pp.919-933, 2004.
 [4] 池原, 阿部, 徳久, 村上:非線型な表現構造に着目した日英文型パターン化, 自然言語処理, Vol.11, No.3, pp.70-95, 2004.
 [5] 石原:英語で日記を書いている, ベレ出版, 2002.