

パターン言語処理に基づく情緒推定に関する研究

2008年1月

徳久 雅人

目次

第1章 序論	1
1.1 研究の目的と背景	1
1.2 関連研究	3
1.2.1 「情緒」という言葉の定義	3
1.2.2 情緒の種類	5
1.2.3 情緒の計算モデル	6
1.2.4 情緒を表す言語知識	7
1.2.5 情緒推定の方法	8
1.3 本研究のアプローチ	9
第2章 言語表現からの情緒推定	13
2.1 言語表現からの情緒推定とは	13
2.1.1 定義	13
2.1.2 原理	13
2.1.3 情緒の存在場所	14
2.1.4 情緒推定の正解	15
2.2 情緒推定方法の分類	15
2.2.1 情緒主の種類	16
2.2.2 言語表現の解釈の種類	17
2.2.3 情緒の推定方法のタイプ	17
2.3 本研究で実現する情緒推定	20
第3章 人間による情緒推定の共通性	21
3.1 調査の目的と方法	21
3.1.1 目的	21
3.1.2 方法	21
3.1.3 タグの設計	22
3.1.4 評価式	23
3.2 情緒タグ付きコーパスの作成と評価	25
3.2.1 テキスト対話への情緒生起原因タグの付与	25
3.2.2 表情の参照可能なテキスト対話への表情・情緒タグの付与	29
3.2.3 日記文への情緒タグの付与	32
3.3 まとめ	37
第4章 パターン言語処理の基礎と実現	39
4.1 言語表現の生成と理解の過程	39
4.2 意味処理への接近	41
4.2.1 意味解析と意味理解	41

4.2.2 言語表現の概念の分類	41
4.2.3 非線形言語モデル	42
4.2.4 意味解析	43
4.2.5 意味理解	44
4.2.6 意味処理のための枠組み	46
4.3 文型パターンの記述子と照合	46
4.3.1 記述子	46
4.3.2 照合	48
4.3.3 構造照合型文型パターン検索	55
4.4 文型パターン辞書検索の基本性能の評価	56
4.4.1 実験条件	56
4.4.2 実験の実施と結果	57
4.4.3 実験のまとめ	61
4.5 文型パターン記述子に関する評価	62
4.5.1 柔軟性を高める記述子	62
4.5.2 評価パラメータ	64
4.5.3 柔軟性の評価	67
4.6 意味属性制約に関する評価	70
4.6.1 意味属性制約の効果の評価方法	70
4.6.2 意味属性制約付き「AのB」型名詞句パターンの作成	70
4.6.3 「AのB」型名詞句の翻訳における意味属性制約の評価実験	71
4.7 まとめ	72
第5章 情緒推定用の言語知識ベースの構築	74
5.1 情緒生起の概念分析	74
5.1.1 ボトムアップ分析	74
5.1.2 トップダウン分析	75
5.1.3 特徴フレームの形成	76
5.2 文型パターンへの情緒属性の付与	77
5.2.1 情緒属性	78
5.2.2 付与の方法	79
5.2.3 第2版から第3版への改訂の様子	80
5.2.4 付与の結果	80
5.3 構築した知識ベースの内訳	80
5.3.1 情緒生起原因を表す用言の多様性	81
5.3.2 用言から推定される情緒の曖昧性	83
5.4 まとめ	88
第6章 インプリメンテーション	89
6.1 パターン言語処理に基づく情緒推定の構成	89

6.2 文型パターンの変形	90
6.2.1 日本語語彙大系の文型パターンの構造	90
6.2.2 記述子の変形の規則	91
6.2.3 変形の実例	92
6.3 適合パターンの選択	92
6.3.1 選択に用いる特徴量	92
6.3.2 ヒューリスティクス	94
6.4 情緒推定の過程	94
6.4.1 入力文の表層レベルの解析	94
6.4.2 情緒属性付き文型パターン辞書	95
6.4.3 SPM による構造照合型パターン検索	95
6.4.4 情緒推定結果の出力	98
6.5 動作確認	98
6.5.1 語義解析と情緒推定	98
6.5.2 ショートストーリーでの情緒推定	99
6.6 まとめ	100
第7章 実験	101
7.1 目的と方法	101
7.2 実験 7-1：知識ベースの網羅性	101
7.3 実験 7-2：情緒推定の精度	102
7.3.1 情緒推定の様子	102
7.3.2 結果	104
7.4 誤り分析	104
7.4.1 誤り分析の対象	104
7.4.2 誤り分析の様子	105
7.4.3 誤り分析の結果	107
7.5 まとめ	107
第8章 検討	108
8.1 省略解析の原理	108
8.2 省略解析の方法	109
8.3 解析の実例	109
8.3.1 Step-1 情緒推定の実行	109
8.3.2 Step-2 推論規則の生成	110
8.3.3 Step-3 省略解析の実行	111
8.4 情緒の意味情報のバリエーション	112
8.5 先行研究との比較	113
第9章 結論	114

謝辞	117
参考文献	118
研究業績一覧	122
主論文	122
参考論文	122
学術論文	122
国際会議	123
図書・解説	123
研究発表	123
研究資料	129
付録1 文型パターン記述子	130
付録2 情緒生起の特徴	132

第1章 序論

1.1 研究の目的と背景

本研究は、計算機による言語意味理解の基礎技術の1つとして、言語表現から話者や登場人物の情緒を推定する方式を提案し、その情緒推定方式に必要な言語知識ベースの設計・構築、ならびに、提案方式の実験的評価を行う。

従来の自然言語処理の研究では、文法的な情報を基本とした形態素解析、構文解析、照応・省略解析、談話構造解析など多くの技術が研究開発され、大きく発展してきたが[1]、それに比べて、言語表現の持つ意味の扱いは難しく、今後に残された課題が多い。従来、文法情報を基本に研究されてきた照応・省略の解析や談話構造の解析でも、意味的な情報が処理できれば、それらの解析精度が大幅に向上することが指摘されており[2],[3]、意味解析技術の確立は、今後の言語処理の発展を左右する重要な課題だといえることができる。

ところで、意味解析の研究対象には、言語で表現される時間・空間・論理など対象世界の捉え方に関する理知的な内容だけでなく、話者や登場人物の欲求・情緒・性格といった感情的な内容も含まれる。従来、前者を対象とした様々な意味処理の研究が行われてきた。たとえば、シソーラスにおける語彙や類語の関係、格関係などである。これに対して、言語表現で表される感情的な内容に関する意味処理研究の例は少ない。これは、感情は人間の高度に精神的な活動の現れであり、その処理を実現するには、人間そのものに迫る研究が必要とみられること、また、そのため、言語で表現される感情についての適切な処理モデルの構築が難しいことによるものと考えられる。

人間の認知的過程の解明を目指した認知科学においては、感情のなかでも「情緒」^{*1} についてを主要な研究課題の1つとしている。認知科学における情緒研究を概観すると、認知的過程の仮説を実証するために、計算機システムに情緒生起のメカニズムを組み込む研究が見られる。たとえば、Pfeifer は、情緒生起の原因の1つとして目標に向けた処理過程での「割り込み」に注目し、意味ネットワークを用いるプロダクションシステムで情緒生起を実現した[4]。Frijda らは、情緒がシステムの関心事を具現化する機構であると考え、人と機械がインタラクションする知的システムに情緒プログラムを組み込んだ[5]。Elliott は、状況の評価的解釈から情緒が生起すると考えて、フレーム表現された状況を Goals, Standards, Preferences の観点から評価し、エージェントの情緒を推定するシステムを開発した[6]。

自然言語処理において、感情的な内容の意味解析は、「言語表現から話し手や登場人物など

*1 情緒は、与えられた刺激によって身体的表出や行動をともなうような、急激で一時的な気持ちの変化とされること、および、感情はそれを含む広義の意味あるいは主観的感覚のみを指すことが多い。後の第 1.2.1 節で述べる。

の情緒を読み取ること」, すなわち, 「情緒推定^{*2}」と言えるのだが, その実現に認知科学の知見が直ぐに利用できないのは, 言語表現から得られる情報が限られているためである. つまり, 認知科学のアプローチでは, 認知過程が計算機上に展開された上で情緒の過程が計算機シミュレーションされているが, 自然言語処理においては, 話し手や登場人物の認知過程をテキストから読み取らねばならず, そのためには, 少なくとも, 基本的な文の解析技術だけでなく文脈の解析技術も必要となる. このため, 現状では, 感情を表す語彙を利用した方式[7]にとどまってしまうか, あるいは, 感情を直接表さない表現を対象にするとしても, 深層格解析や省略・照応解析の上に成り立つ方式[8],[9]となってしまう. しかし, これらの方式では, 感情的意味を基本的な言語解析に利用することができず, 文法的解析と両輪を保つ意味的解析が効果を発揮しない.

そこで, 本論文の筆者は, 情緒処理に対する問題の枠組みを次のように大きく捉え, 感情的意味の解析の問題に対して将来にわたり段階的に解決していくことにする. 本研究はその第一段階であり, 一部分を重点的に解決する.

(a) 生起過程と反応過程

情緒主が外的あるいは内的な刺激を受けてから情緒を生起させるまでの過程と, 生起した情緒に基づいて反応を行うまでの過程

(b) 定性的と定量的

特に生起過程においては, 情緒原因や情緒状態としてどのような特徴を持つのかという定性的側面と, どれほどの強さを持つのかという定量的側面

(c) 一般性と特殊性

外見上主観的で個人差の大きいと見られる情緒も, 一般性のある部分と個別事情に基づく部分とに分けられる.

本研究では, 上記(a)に関して情緒の生起過程の特徴と言語表現の表す概念との対応が情緒推定の手がかりとなることに着目し, 情緒推定の実現を目指す. ここで, 言語表現から概念を抽出すること, および, その概念から情緒生起特徴までの対応を推論することが, 従来手法では言語の構造解析や心理モデルに依存し問題となっていたが, 本研究では, 言語の表現構造を抽出することで表現のもつ概念が分離できるという「非線形言語モデル」[10]を取り入れることで解決を図る. そのため, (b)における定性的側面と(c)における一般的側面^{*3}から, 情緒原因の特徴を調査するとともに, 言語表現の概念と情緒原因の対応を知識ベース化する.

本研究で取り組む具体的な技術的課題として次の点があげられる:

- 情緒推定方法: 情緒推定に使用する言語表現上の特徴を明示して, 情緒推定の方法とその位置づけを明確にすること
- 評価尺度: 人間による情緒推定において, 人々の間で推定される情緒の共通性を定量的に示すこと
- 知識ベース化: 網羅的かつ客観的判断により情緒推定用の知識ベースを構築すること

情緒推定を人間が行う際に注目する言語表現の特徴は幅が広い. 対話テキストであれば, 言葉

*2 情緒推定において, 情緒の持ち主のことを情緒主, 情緒を推定する行為者のことを推定者とそれぞれ呼ぶことにする.

*3 扱う言語表現が文から文章などへと広がり, 概念が具体的に得られるようになると, 推定される情緒の特殊性が増すと予想される.

尻, 対話相手への要求や対話相手からの返答, 対話者の置かれた状況などが考えられる。日記や物語では, 直接的な情緒の表現, 登場人物の行動の描写, 描かれている環境などが考えられる。また, 情緒主の設定も選択肢の1つであり, たとえば, 物語の登場人物, 物語の書き手, 読者が挙げられる。多岐にわたる言語的特徴の全てを1つの研究で応えることは容易ではなく, かつ, 現在はそれらの統合的手法を議論する時期ではないため, 各研究において, 対象とする情緒推定方法を明確にする必要がある。

評価尺度の導入は工学的立場として不可避である。認知科学では, 情緒の計算モデル, および, それに基づくシステムが構築されたが, 情緒の生じる過程の仮説に対して計算機により検証することが重視されていた。そして, そのシステムの性能の評価という議論はなされていなかった。しかし, 工学的アプローチでは, 有効性を問うために評価しなければならない。人間による情緒推定の共通性に向けて, 計算機による情緒推定が接近できた度合いを評価することが適当であろう。たとえば, 音声認識における情緒推定に関しては, κ 値や一致率という評価尺度が用いられている [11],[12]。テキスト言語理解における情緒推定においても同様の手法が可能であると思われる。

知識ベース化は, 工学的な自然言語処理における重要な課題の1つである。自然言語処理の技術は, 日本語語彙大系や EDR 概念体系辞書に代表されるシソーラス [13],[14], 新聞記事などに統語的・意味的情報などを付与したコーパス [15] が資源となり, 各種解析ツールを駆動させて, 新たな領域での言語処理や, より高精度な処理が実現されてきた。言語的資源とツールの相互の質的・量的な向上が自然言語処理では重要である。情緒推定を基礎技術として研究する上でも同様であり, 既存の資源から情緒推定を実現する立場と, 新たな言語資源を構築する立場とがありうる。ここで, 後者の立場では, 情緒を扱う知識ベースの構築が課題である。さらにその構築のための客観的な基準が必要である。知識ベースは, 通常, 改良を繰り返す。そのためには, 複数の分析者が携わる。客観的な基準があるならば, 分析者の主観を抑制することができ, 分析者に依らず安定した構築が可能であり, また, 基準の調整による知識ベースの修正・改良という構築手順が成り立つ。

そこで, 本研究では, 次のことを行う。(1) 日本語の言語表現からの情緒推定の方法を鳥瞰し, 現在開発の遅れている知識ベースの種類を特定するとともに, 対象とする情緒推定方法を示す。(2) 人間による情緒推定を試みて, 問題領域の感触を得るとともに, 情緒推定に対する評価基準を示す。(3) 情緒推定方法と言語理解の関係を考察し, 言語知識ベースの形式を設計する。(4) 情緒の生起についての概念分析を行い, 知識ベースの構築のための客観的な判断基準を設ける。(5) その基準に基づき知識ベースを構築する。(6) 情緒推定実験により知識ベースの達成状況を評価する。

1.2 関連研究

1.2.1 「情緒」という言葉の定義

「嬉しい」や「悲しい」という気持ちに対して, その気持ちの状態だけでなく, 気持ちの生じる原因や気持ちに対する反応という動的な側面を「感情的な意味情報」として言語理解タスクに提供することを目指している。ここで, 気持ちを表す用語として, 日常的には, 「感情」, 「情緒」, 「情動」, 「情操」, 「情感」, という言葉が使われている。また, 心理学や認知科学において, これらの語は区別されているが, あまり安定していない。そこでまずは諸文献における用語の定義を確認しておく。

1969年の文献[16]では次の説明がある:

- **感情 (feeling):** 「外界の状況にゆり動かされて起こるところの主観的な状態が, 感情の基であって, 外界を認知する知的機能とは異なるものである. それは, 機能や作用ではなくて状態なのである. 」
- **情緒・情動 (emotion):** 「より一般的に主観のゆり動かされる状態であると言われる. 」
- **情操 (sentiment):** 「科学, 道徳, 芸術, 宗教などの文化的価値をもったものに対して起こる感情的な統合的全体を意味するといわれる. 」

1994年の文献[17]では, 1982年~1991年の文献[18],[19]をベースとして社会的認知研究の標準的な定義を次のように紹介している:

- **感情 (affect):** 「以下の「好み」「評価」「ムード」「情動」のすべてを含む一般語である. 」
- **好み (preference):** 「快・不快を表す比較的マイルドな主観的反応である. 」
- **評価 (evaluation):** 「好みの中でも「対人的評価」の語を社会心理学者はよく用いてきた. つまり他者に対する単純な正負の反応である. それは内容的には, 対人魅力, 対人的好嫌, 偏見などを含む. 」
- **ムード (moods):** 「特定の対象がなく, 現行の思考過程のコンテキストを形作る感情である. 」
- **情動 (emotion):** 「単によい/悪い感じ(feeling)を越えた複雑な感情の組み合わせで, しばしば生理的喚起(覚醒)を伴った表出反応を含む. 恐怖や怒り, 悲しみ, 喜びなどがある. しばしば「感情」と同義に用いられる. 」

そして, 注意書きとして, 「一般的にはこのように定義可能であるが, 研究の系統によっては感情の定義の範囲や, 感情と認知の関連性をどう定義するかという点でかなりの差がある」([17], p.174)という記述がある.

1999年の文献[20]では, 次の説明がある:

- **感情 (emotion):** 「(前半省略) 英語圏においては emotion が, 日本語圏においては「感情」が, 感情にかかわる最も包括的な用語として用いられていること, また 1993年に発足した日本感情心理学会において感情 = emotion を学術用語として対応させたこと, さらに feeling としての「感情」研究は欧米においても日本においても少ないことから, 感情と emotion を相互に対応する用語として扱う. なお日常語として用いられる「感情」は, 感情の意識化された主観的成分を強調して用いられる場合が多く, emotion というより affect ; affection ; feeling といった英語に近い意味で用いられる場合が多いので注意されたい. (後半省略)」
- **情動 (emotion):** 「感情(emotion)の動的側面が強調される場合に用いられてきた用語であり, 急激に生起し, 短時間で終結する反応振幅の大きい一過性の感情状態または感情体験をさす. 」
- **情緒 (emotion):** 「情動と同義. 」

自然言語処理の分野では, 文献[21]において「情緒」が使われたが, 近年の研究発表においては「感情」の方が多く見られる. 英語においても, emotion よりも affect の方が使用されている[22]. 感情に対する反応動作や表出の伴うことが必ずしも議論の対象ではないためであると思われる. また, 書評や口コミのような「評判分析」は, “sentiment analysis” と翻訳される.

なお, 感性情報処理の分野では, 情報処理を, 物理レベル, 論理レベル, 感性レベルに階層的

に分類し、音楽・絵画・表情・仕草を通じてとらえられる主観性についてを「感性」として扱っている[23]。情緒・感情などを意識した情報処理が課題とされているものの、言語を論理レベルに位置づけている。そのため、自然言語処理の分野からは、主観性のある情報を扱うために「感性」という言葉を使うことはあまり見られない。

本研究では、**情緒 (emotion)** という言葉を使用する。将来的には反応動作や表出との関係を議論の対象とすることを狙うためである。**感情 (affect; feeling)** という言葉については、広い範囲での気持ちや感想を表すものとして、とりたてて使用しない。

1.2.2 情緒の種類

情緒の種類として、直感的・日常的には、「喜び」、「悲しみ」、「怒り」などが挙げられるが、学術的には情緒の種類が固定されていない。さらには、そもそも、情緒の種類として、カテゴリ分けすることが適当であるのかさえ問題とされ、快-不快、注目-拒否、睡眠-緊張を軸とする3次元モデルが提唱されている。情緒の種類に何があるのかという問いは、容易に答えられるものではなく[20]、あえて議論を避けて、計算機処理の可能性を考察することさえある[24]。本研究は、言語表現からの情緒推定を行うので、その結果として、情緒の種類を表す表記が必要となる。そこで、先行研究において提案されている情緒の種類の中から、言語表現や言語処理として有効と思われる情緒セットを使用する。そのために、先行研究における情緒セットを概観する。

情緒セットとして単純なものは、快-不快、正-負、あるいは、+と-のように2値のもの、あるいは中立を含めた3値のものである。たとえば、Lehnert は、物語理解において、各事象に+ (正の情緒)、- (負の情緒)、M (中立)の情緒状態を割り当て、それらの関係を、a リンク(実現)、m リンク(動機)、t リンク(終結)で表した[25]。そして、情緒状態とリンクによるネットワークのトポロジに対して、「競争」や「復讐」というラベルを与えて物語の要約を行った。情緒の種類は少ないのだが、事態間の関係をリンクで表すことで、物語の構造を捉えることができる。

情緒をよく表すものとして顔の表情が古くから注目されている。心理学において Ekman は6種類の表情として、「驚き」、「恐怖」、「嫌悪」、「怒り」、「幸福」、「悲しみ」を挙げている。この分類は、文化や地域に依存しないと述べられている[26]。そのため、マンマシンインタフェースの応用に使われることが多い。たとえば、情緒は、学習者のモチベーションを高めることに有効であるとして、擬人化エージェントによる児童向け教育システムに、情緒の機能が導入されている[27]。ここでは、後述する OCC モデルに基づく情緒の体系を使用しながらも、児童に分かりやすい情緒として Ekman の分類の中から選択した情緒を使用している。

心理学において、情緒の種類を上げるために、内観や表情が根拠とされていたが、その他の根拠として、情緒を表す語彙に着目した分類がある。Plutchik は、情緒の強さに応じて語彙が存在すること、類似する情緒に応じた語彙が多様に存在すること、および、相反する情緒や行動を表す語彙が存在することを指摘して、多因子分析法に基づき語彙分類を行った[28],[29]。その結果、基本的情緒として、joy (喜び)/sadness (悲しみ)、acceptance (受容)/disgust (嫌悪)、fear (恐れ)/anger (怒り)、surprise (驚き)/anticipation (予期)という対極する8つの情緒を示した。複雑な情緒として、たとえば、joy + acceptance = love (愛)、acceptance + fear = submission (服従)、fear + surprise = awe (畏怖)のように基本的情緒の合成で説明した。情緒の強さに応じた分類としては、たとえば、apprehension (不安)-fear (恐れ)-terror (恐怖)や、annoyance (いらだたしさ)-anger (怒り)-rage (激怒)のように情緒的語彙が並べられた。なお、Plutchik の基本情緒の分類には、情

緒の過程(emotion sequence)の概念が取り込まれている。その過程とは、「刺激となる事象」、「認知」、「感じている状態」、「表立った振る舞い」、「効果」である。たとえば、「価値のある物の獲得」という事象に対して「所有」を認知し、「喜び」を感じる。そして、情緒に反応して、「保持・繰り返しの振る舞いにより、「資源の獲得」の効果を指すという過程が説明されている。

認知科学において、情緒の過程における「事前評価(appraisal)」に重点を置いた分類として、OCC モデルがある[30]。OCC モデルでは情緒の生じる状況を event, action, object の3つの観点から評価することを基準に、情緒を分類している。Elliott はその分類に改良を加えて、26 種類に分類している。それは、joy, distress, happy-for, gloating, resentment, jealousy, envy, sorry-for, hope, fear, satisfaction, relief, fears-confirmed, disappointment, pride, admiration, shame, reproach, liking, disliking, gratitude, anger, gratification, remorse, love, hate である[6]。

以上の他にも様々な種類の情緒が挙げられており、その解説をする文献がある(たとえば, [31], [32])。こうしたなか、本研究では、次の理由から、Plutchik の分類を参照する:

- 基本情緒の組み合わせで複雑な情緒が説明できる点は、計算機処理に都合が良い。
- 語彙分類に基づき体系が作られているため、言語処理と相性が良いと思われる。
- 言語表現には時制の概念が強く取り込まれている。Plutchik の分類には fear や anticipation のように未来や予期に関する情緒が基本的情緒として挙げられている。
- 言語表現には感嘆文や強調文の概念が取り込まれている。Plutchik の分類には surprise (驚き)が基本的情緒として挙げられている。
- 情緒の強さや情緒の過程を説明できる体系である。

1.2.3 情緒の計算モデル

情緒の計算モデルは、認知科学の分野において示された。情緒の生じる過程に対する認知的な仮説を、計算機によるシミュレーションにより検証することが目的であった。Pfeifer は目標実現過程における割り込みとして情緒を位置付け、プロダクションシステムを用い情緒の過程を再現するシステム FEELER を作成した[4]。Frijda らは関心実現機構として情緒を位置付け、対話システム ACRES に情緒的な機能を持たせた[5]。両者は、Goal・Plan の達成状態を要因とした情緒生起の計算モデルである。その後、情緒の生起する状況を event, action および object に対して評価することで情緒生起が捉えられるとして、OCC モデルが示された[30]。Elliott は、これら3つの状況分類に対して goals, standards および preferences の観点で評価することフレーム表現により実現した[6]。OCC モデルは仮想世界上のエージェントやロボットの内部表現との親和性が高いことから、認知科学や人工知能の分野では、アプリケーションの情緒的機能として広く利用されている[27],[33],[34]。

一方、言語理解に関して、Dyer は深い物語理解の中で情緒を取り扱った[7]。計算機上での知識表現に対して議論がなされた。状況を記述するために、CD (conceptual dependency)^{*4} を素性として、ゴールとプラン、スクリプト、および、TAU(Thematic Abstraction Unit)^{*5} が導入された。物語理解では、登場人物の行動の因果関係を説明することが課題とされていた。たとえば、誰が行動をしたのか、どこで行動したのか、なぜ行動を起こしたのかなどである。そのため、行動の断片から、

*4 文をプリミティブな行為に分割して意味を表すための系

*5 エピソード記憶の構造

ゴールを求める問題(プラン認識)や隠された行動を求める問題のためめに、プランやスクリプトという知識表現が必要とされた。情緒に関しては、感情語彙 (affective lexicon) という知識を導入した。この語彙知識は、STATE (2値の情緒)、CHAR (情緒主)、G-SITU (ゴールに関する状況)、TOWARD (情緒対象)、SCALE (情緒の定性的な強さ)、E-MODE (予期の有無)で構成される。解析対象のテキストに、感情語彙を検出すると、これらの構成要素を利用して関連する文から行動の関連性を推論する。こうして、情緒状態が判明したときに深い言語理解を行う過程を示す研究である。

ここで、言語の表す事柄を、プラン知識やフレーム表現のような計算機用の知識表現で記述することは、ロバスト性の確保が難しい。そこで、物語の対象を狭く深くした上で開発した技術を徐々に広げていくアプローチと、対象を広く浅くした上で開発した技術を徐々に深めていくアプローチに分かれている。

前者は、Mueller が現在も続けているアプローチである[35]。Mueller の方法は、言語の形態素・構文解析・照応解析などが完了し、物語が event calculus と呼ばれる一種の論理形式で記述済みであることが前提となっている。そのための論理形式の述語の設計が議論の対象となっている。後者は、Riloff らの研究がある[36]。Riloff らの方法は、情報抽出の技術にならない、パターンにより言語表現から必要な情報を抽出すること、および、統計的な信頼性に基づきパターンの適合要素を判定する方法である。

後者の方法に関して、情緒を推定する方法が Liu らによって提案されている[9]。Liu らは Web 上のドキュメントから、Positive/Negative を表す単語と共起する表現を収集し、Positive/Negative の判定に有効な格関係を抽出する方法である。この方法は、ドキュメントの分類の問題としては有効に働くのだが、情緒主と情緒対象の関係、Positive/Negative な情緒と他の事態との関係など、物語理解において要請されている関係までは解析できない。

1.2.4 情緒を表す言語知識

本研究は日本語を対象として網羅的な言語知識ベースの構築を目指しているので、日本語における情緒を表す言語表現について、関連研究を示す。

日本語の品詞には、動詞、形容詞・形容動詞、名詞、副詞、連体詞、接続詞、感動詞、助動詞、助詞、接辞がある。それぞれにおいて、情緒を表す単語や熟語(相当表現)がある。一般のシソーラスにおいて、単語の語義ごとにカテゴリが設けられている[37],[38]。そこには、「感情」のカテゴリが示されており、上記の動詞から副詞まで、感情を表す単語が示されている。また、感情に関する語のみを収集した辞典もある[39]。形容詞・形容動詞、副詞、擬音語・擬態語、および、それらの相当表現については、主たる語義に加えて感情的なイメージが伴うことがあることから、その感情的なイメージの有無をまとめた辞書も出版されている[40],[41],[42]。

言語表現から情緒を推定する上で、情緒の種類だけでなく、情緒主と情緒対象も出力することが望まれる。日本語における感情表現文は主語が1人称であり、情緒主が話し手であると言われている[43]が、実践的に日本語を解析する上では、情緒を表す単語が、条件節や引用節などで使用される場合には情緒主が話し手になるとは限らない[44]。したがって、辞書には、情緒主と情緒対象についての情報を付随させる必要がある。松本らは、日本語の結合価に着目して、情緒を明示する表現に対して、情緒主や対象の解析ができる辞書を作成した[45]。

一方、助詞・助動詞、および、これらを含む付属語的な表現に着目した辞書もある。グループ・ジ

ヤマシイは、「～てしまった」などの表現を、日本語文型辞典としてまとめ、感情的なニュアンスを解説した[46]. 助詞・助動詞による表現要素は、主体的表現と呼ばれる。日本語文は、客体的表現と主体的表現で構成されており、客体的表現に対する話し手の立場が主体的表現に表されるという[47]. このことから、日本語文型辞典に示されている感情の解説は、話し手を情緒主として知識ベース化が可能であると思われる。

1.2.5 情緒推定の方法

先行研究における情緒の計算モデルや言語知識に基づき、言語表現から情緒を推定する方法についてまとめる。

「情緒の過程」は、「情緒を生じさせる原因の発生」、「情緒を持つ状態」、および、「情緒の反応」で構成される。言語表現に情緒的過程のいずれかの事態が表されていることを根拠にして、情緒を推定できる。この関係を図 1.1 にまとめる。

Dyer の方法では、言語表現が「原因」を表すものとしては、ゴールなどの非情緒的な意味までを解析する。Elliott の Affective Reasoner は、言語を直接扱うものではないが、非情緒的な意味の解析結果に相当するものが、フレーム表現で得られることが前提となっており、それを基に情緒の推定が可能になっている。両者の方法を組み合わせることができれば、深いレベルの言語理解を通じて情緒推定が可能になるとと思われる。

既存の辞書(第 1.2.4 節)を用いると、「情緒状態」を表す言語表現から直接的に情緒的意味の解析が行われる。その解析は、述語だけの判断ではなく述語に後続する表現との組み合わせにより情緒を確定させる[45]. また、Dyer の方法は、情緒状態を表す表現(感情語彙)を推論の起点として、原因となる状況に向けて解析を深める。その結果、上述の情報(ゴールの状況など)との関係が解析される。

「反応」を表す言語表現から情緒を解析する方法は、Elliott の Affective Reasoner では、用例ベースに基づく方法が述べられている。しかし、反応は、表情など身体状態を表すものでない限り、情緒推定においては根拠が弱い。たとえば、図 1.1 では、魚釣りのエピソードの一部であるが、反応の「明日もまた来よう。」だけで魚釣りのエピソードを理解しようとすると、「全く魚が釣れず、意地になっている」という推論も成り立つ。

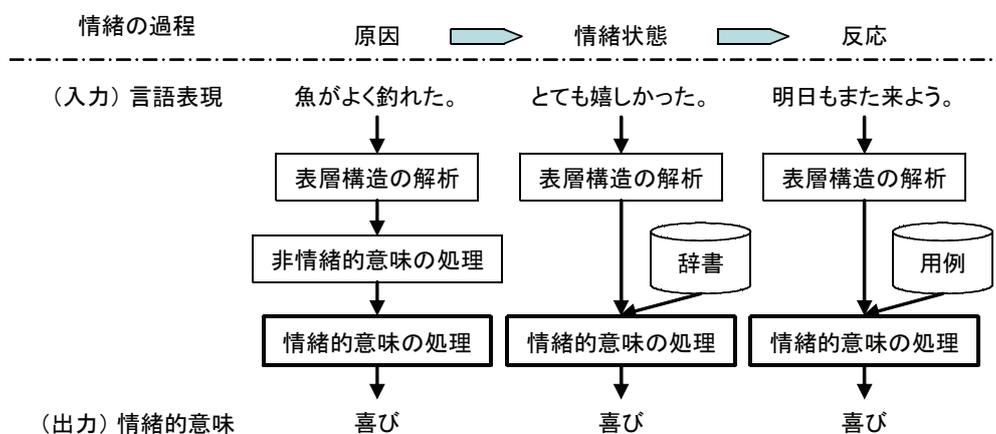


図 1.1 先行研究に見られる情緒推定の位置付け(その1)

浅く広く言語理解を行うアプローチの場合、「原因」から情緒を推定する際、深いレベルの非情緒的意味の解析を行わずに、情緒的意味の解析を行う(図 1.2)。目良らの方法[8]、Liu[9]らの方法である。しかし、目良らの方法は、深層格フレームの得られることが処理の前提条件であるため、照応・省略解析が必要である。Liu らの方法も同様に単語の省略された場合には正しく解析できない。たとえば、「太郎が花子と付き合う。」という表現から話し手の情緒を推定する際、話し手と花子の関係が恋愛関係であり、話し手と太郎の関係が恋敵である場合は、かけ算的な好感度の計算や表現の共起性より理想通りの処理が可能であるが、「花子と付き合う。」のように格要素が省略されている場合には省略解析をしなければ理想通りの処理はできない。したがって、省略・照応の解析などの言語理解の実現が前提であり、その解析技術に対して感情的意味を提供することには利用し難い。

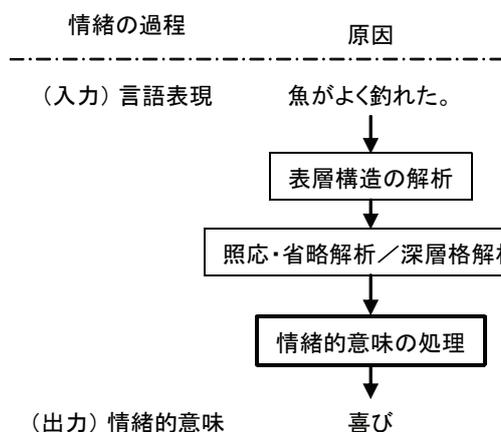


図 1.2 先行研究に見られる情緒推定の位置付け(その2)

1.3 本研究のアプローチ

先行研究において、情緒推定は、情緒の過程のうち、「原因」、「情緒状態」、「反応」のいずれかを表す言語表現を解析することで実現してきた。

「情緒状態」を対象とする解析方法は、現状では一定のレベルに到達していると言える。たとえば、言語知識ベースが、日本語の基本的な単語を網羅していること、さらに解析により得られる情報として情緒名、情緒主、情緒対象などが含まれていることが理由として挙げられる。

「反応」を対象とする解析方法は、情緒推定が不確定となりがちであるため、情緒状態を対象とする方法、および、原因を対象とする方法の両方と協調した方法とするなどの工夫が必要である。そのため、原因を対象とする解析方法が可能となった後の課題と思われる。

「原因」を対象とする解析方法は、ゴールやプランなど非情緒的意味という深いレベルの解析を前提とする方法、および、照応・省略解析や深層格解析という中レベルの解析を前提とする方法がある。ここで、本研究では、言語理解の基本的なタスクに、情緒という意味的情報を提供することを狙っているのだが、従来の方法には次の2つの問題がある。第1は、言語理解と情緒推定方法の関係である。照応・省略解析や談話解析などの中レベルの解析を前提とする情緒推定の方法では、情緒的意味をそれらのレベルの解析に提供できない。第2は、言語知識ベースの充実度である。たとえば、目良らの方法は、深層格に関する言語知識を網羅的に準備する必要があるのだが、現在のところ網羅的に実現されていない。また、Liu らの用例に基づく方法に使用する知識には、

情緒名, 情緒主, 情緒対象など関連する情報を抽出するための機構が備わっていない。

そこで, 本研究では, 次の点に取り組む:

- 表層的な言語情報から, 情緒の生起する原因を判断し, 情緒を推定する方法を示す.
- そのための言語知識ベースを, 日本語の基本的な用言の範囲で網羅的に構築する.

図 1.3 に本研究の情緒推定方法の位置付けを示す. 先行研究の方法(図 1.2)と比べて, 中レベルの解析による情報を使用しない点が異なる. 表層構造から直接的に情緒的意味の解析を行うことに, 技術的な飛躍がある. 本研究は, 池原の「非線形言語モデル」[10]に基づく「パターン言語処理」を応用することで, それを実現する. 非線形言語モデルによる言語表現の解析の特徴は, 「言語表現の構造を抽出すること」, および, 「表現構造に対応付けられた意味的情報を解析結果として出力すること」である. したがって, 非線形言語モデルにより情緒的意味を解析するためには, 次の2点が副課題となる:

- 表層構造を抽出するための言語知識の記述形式を設計する.
- 表現構造が情緒生起の原因を表すかどうかを判定し, 情緒的な情報の対応付けをする.

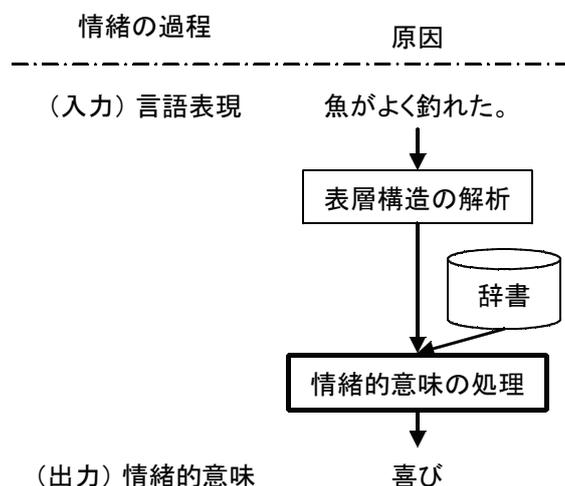


図 1.3 本研究の情緒推定の位置付け

前者については, 日本語語彙大系において有効性が確かめられたことを受けて, 文型パターンを基礎とする形式を採用し, 本研究ではさらに柔軟性を高めた記述子の導入を行う. 後者については, 人手による分析作業を基本とするのだが*, どのような事態が情緒生起の原因を表すのかについての判断基準が必要である. 情緒生起の原因について, OCC モデルでは, event, action, object に対する goal, standard, preference の成立に着目することが示されているが, それは抽象度が高く, 言語表現に対する分析の基準としては不適當である. そこで, さらなる副課題として, 次のことがあげられる:

- * 情緒の生起する原因の特徴を列挙する.

*6 大規模コーパスとみただ Web ドキュメントを情報源として, 感情表現語と依存関係にある事象の表現を収集することで, 情緒の生起する状況の収集を行う試みがある[48]. 情緒の種類とその状況との組は自動的に得られるが, 情緒主と情緒対象までは自動的に得られない. 現段階では, 人手による分析の方が直ぐに実現できる.

情緒の生じた状況を描いた物語を分析し、さらに、心理学や認知科学で指摘されている事柄を導入することにより、情緒生起原因の特徴を体系的にまとめる。

以上により構築した言語知識ベースについて、正しく作成できたのか、また、どの程度の精度で推定できたのかについて評価する必要がある。言語表現からの情緒推定とは、「与えられた言語表現から、ある人物の情緒を読み取ること」と本研究では定義するので、情緒推定の正解というものは、「与えられた言語表現から、複数の人が共通的に読み取った情緒」として定めることができる。計算機が人間の代行として言語理解の処理を行うとき、正解が不明確なものに対しては、多くの人に信じられる事を計算機が出力すべきという考え方や、少数でも信じられる事を逃さず計算機が出力すべきという考え方ができる。

本研究は、前者の考え方から定まる正解を第一目標として考える。そこで、人間が言語表現から情緒を推定する際、そもそも、どの程度共通した情緒が推定されるのかについて、基礎的な知見が必要である。そのために次の課題がある。

- 人間による情緒推定の共通性を定量的に評価する。

そのために、本研究では、「情緒タグ付きコーパス」の構築を行う。このコーパスは、テキスト対話や日記文を収録したものであり、これらには「情緒タグ」を付与する。情緒タグは、人間がテキストから読み取った情緒を表す注釈である。同一箇所複数の者が付与したタグの一致を集計することで、情緒推定の共通性を評価する。

以上について、本論文では、図 1.4 の構成で議論を進める。第2章では、言語表現からの情緒推定について定義を述べ、情緒推定方法の分類を示す^{*7}。第3章では、コーパス構築を通じて人間による情緒推定の共通性を調査する^{*8}。第4章ではパターン言語処理の基礎と実現について述べる。特に前半では、表層的な言語情報からの意味処理について、基礎的な思想を述べ、後半では、文型パターンの記述とその運用方法を示す^{*9}。第5章では、文型パターンを基本形式として、情緒推定のための知識ベースの構築を行う。そこでの問題は、情緒と関わる文型パターンであることを客観的に判断する基準を設けることである。そこで、第5章の前半では、情緒の生起する原因に対する概念分析を行う^{*10}。第5章の後半ではその分析結果に基づき知識ベースの構築を行う^{*11}。第6章では、パターン言語処理に基づく情緒推定の方式を実装し、第7章において、その性能を実験的に評価する。第8章では、本提案方式で情緒的意味を提供することの可能性について、簡単な照応解析の例題を用いて示す。

*7 主論文 8 において、情緒の位置付けを述べた。先行研究ではあまり明確な議論がなされていなかった。

*8 自然言語処理の分野では、人間による情緒推定の様子や精度が扱われていなかった。主論文 3, 7 および 8 はコーパスの構築を通じてそれらを具体的に扱った。

*9 自然言語処理においてパターンを使用する方法は古くから採用されてきた。しかし、数万から数十万のオーダーでの使用は例がない。その理由としてパターンの記述言語仕様がなかったことが挙げられる。主論文 4 および 6 により、意味的・統語的な考慮をしたパターン記述言語仕様の作成に結び付いた。

*10 情緒の生起についての概念分析を主論文 1 および 2 において行った。

*11 主論文 5 において構築方法とその結果について示した。

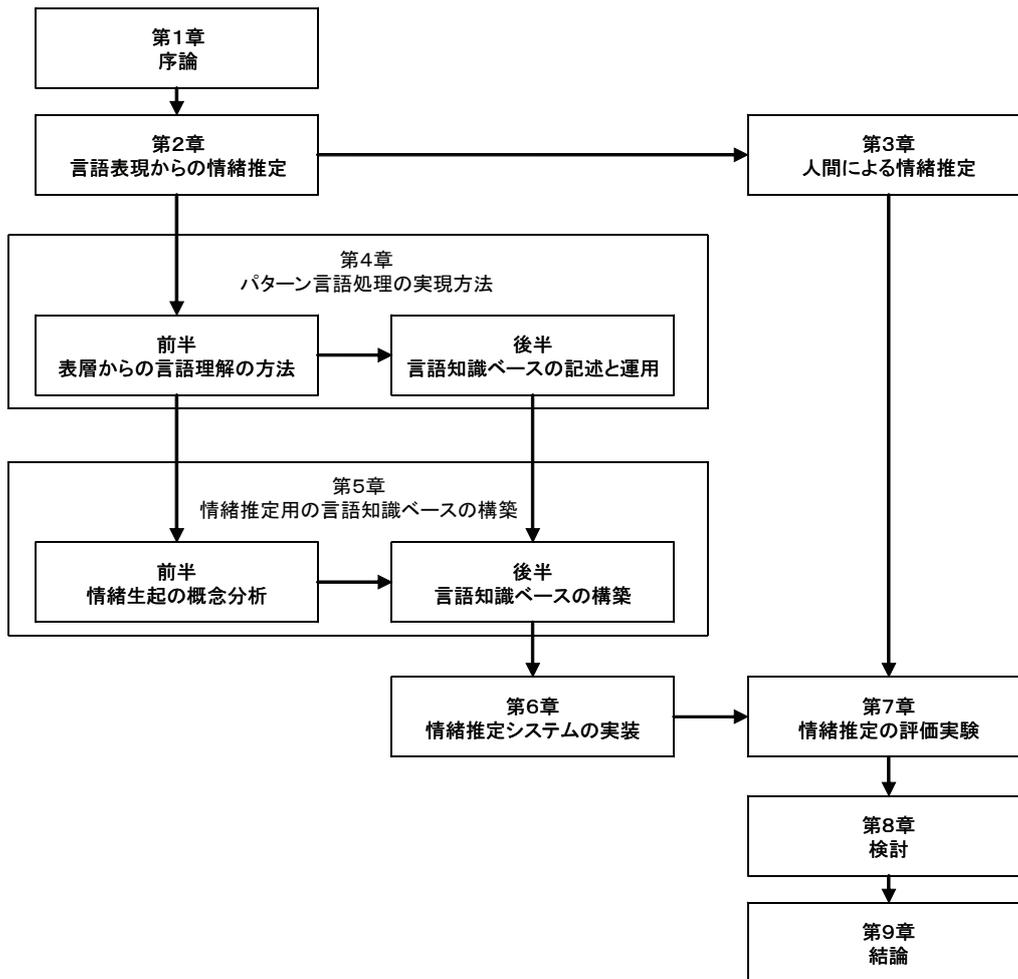


図 1.4 本論文の構成

第2章 言語表現からの情緒推定

本章では、「言語表現からの情緒推定」を定義するとともに、その方法の分類を示す。人間がどんな情緒を持つのか、あるいは、言語表現から人間がどんな情緒を推定するのかは、各者の主観的要素であるため一見すると科学的方法が成り立たないように見える。そこで、工学的立場から、情緒推定で扱う情緒の位置付けを定義することで、客観的に扱えるものであることを示す。次に、情緒推定の方法は、「情緒の過程」、「情緒主」、および、「言語表現の解釈の種類」に着目して分類する。関連研究を、この分類に対応付けて、現時点で開発の遅れている情緒推定の方法を明らかにする。

2.1 言語表現からの情緒推定とは

2.1.1 定義

「言語表現からの情緒推定」とは、「与えられた言語表現から、ある人物の情緒を読み取ること」である。情緒推定を行う人物のことを「推定者」と呼び、読み取られた情緒の持ち主のことを「情緒主」と呼ぶ。

2.1.2 原理

言語表現から情緒を推定する根拠は、「言語表現が情緒主の情緒の過程と関わるか否か」である。

「情緒の過程」とは、既に第 1.2.5 節で示したが、次の3つの段階である：

- 情緒生起の原因の発生
- 情緒の生起した状態
- 情緒反応の発生

「情緒生起の原因」とは、「望みが叶った」や「不当な扱いを受けた」など「喜び」や「怒り」などを引き起こす状況である。「情緒の生起した状態」とは、情緒主が、「喜び」や「怒り」などの情緒を感じていることを指す。「情緒反応」とは、情緒により表情が変化することや体がビクリと動くことのような身体的な反応の他に、情緒による新たな動機付け、振る舞いの割り込み、あるいはこれらの気持ちの強弱の変化というような心理的な反応がある。

「(情緒の過程と)関わる」とは幾つかの考え方がある。1つには、言語表現が直接的に情緒の過程を表すことである。たとえば「僕は望みが叶った」と言えば情緒原因を直接述べることになる。別のものとしては、言語表現が間接的に情緒の過程を表すことである。たとえば、「バーベキューをする」という予定があるときに、「明日は晴れる」という表現からは、情緒原因の成立している状況であることがわかる。

このように、言語表現からの情緒推定は、原理に対して方法が幾通りかありうる。後の節では、言語表現の解釈の仕方に注意を払いながら、情緒推定の方法を整理する。

2.1.3 情緒の存在場所

情緒推定により読み取られた情緒は、推定者が信じる情緒であって、情緒主が本当に持つ情緒ではない。このことは当然の事柄であるが、言語表現と情緒の関係を考察する際に混乱しやすいので、情緒推定で扱うことのできる情緒について確認しておく。

情緒主が持つ情緒は、その情緒主だけが知っている。情緒主から正直に情緒が伝えられない限り、情緒主以外の人物は情緒主の持つ情緒を知り得ない。つまり、表情、発話音声、言葉尻などに情緒の表れる機会はあるものの、情緒主はこれらを意識して操作でき、情緒主の持つ情緒とは異なる情緒が表される可能性があるためである。

このことに関して、情緒には2つの存在場所があり、各場所の情緒を次のように呼ぶことにする：

- **真の内在する情緒：**情緒主の内部に存在する情緒
- **表出する情緒：**表情・音声・文字などの表現に存在する情緒

「表出する情緒」は、情緒主が表現に認めている情緒と、推定者が表現から解釈する情緒の2面性がある。たとえば、情緒主にとって「平常」を表す顔が、推定者にとって「不機嫌」を表す顔と見えることがある。Ekman の指摘するとおり表情の解釈の共通性は高いのだが、社会的に顔の形を規定するほどのものではないので2面性がある。一方、言葉の表現で「嬉しい」といえば、情緒主も推定者も表出する情緒はともに「嬉しい」ということになる^{*12}。このように、2面性の無い場合もある。

情緒推定において、推定者は、主として真の内在する情緒を追求して推定を深め、表出する情緒はその途中で解析される。たとえば、幸福の表情で「嬉しい」と情緒主が発言すれば、推定者は表出する情緒としては、幸福や嬉しいという解釈をするだろう。しかし、情緒主がアイドルタレントであり推定者がゴシップ記者であったとき、両者の社会的関係や情緒主の性格をはじめとする様々な事柄を総合し、情緒主が笑顔を見せたとしても本当のところはそれほど喜んでいないのではないかと推定するかもしれない。このことから、情緒にはもう1つの存在場所があり、その場所の情緒を次のように呼ぶことにする：

- **推定上の内在する情緒：**情緒主の内部に存在すると、推定者が信じている情緒

表出する情緒と推定上の内在する情緒の区別について考えてみよう。言語表現についていえば、「嬉しい。」のように文字通りの解釈が情緒を表す場合には、表出する情緒として推定者は解釈する。一方、「褒められた。」のように「喜び」の可能性はあるが、その情緒が明言されていない場合には、表出する情緒としては推定者は何も解釈せず、その可能性のある情緒は推定上の内在する情緒とみなす。

以上の関係を図にすると図 2.1 のようになる。この図では、表現の主を情緒主として情緒推定を行っている。情緒主と推定者はそれぞれ信念を持っている。推定者は信念の中に情緒主の信念を持っている。言語表現・音声・表情による表現を解析した結果から、表現に表出する情緒を、推定者は内部の情緒主の信念に留めている。情緒主は、表現に認めている表出する情緒を自信の信念に留めている。次に、推定者は、表現内容・情緒主の性格・その他の根拠から、情緒主の真の内在する情緒を目指して情緒を推定し、推定上の内在する情緒として信念に留めている。

*12 情緒主は、皮肉を込めて「嬉しい」と言う際、この表現に《嬉しいという気持ち》は込めていないが、この表現が《嬉しいという気持ち》を語義として表すことを認めている。

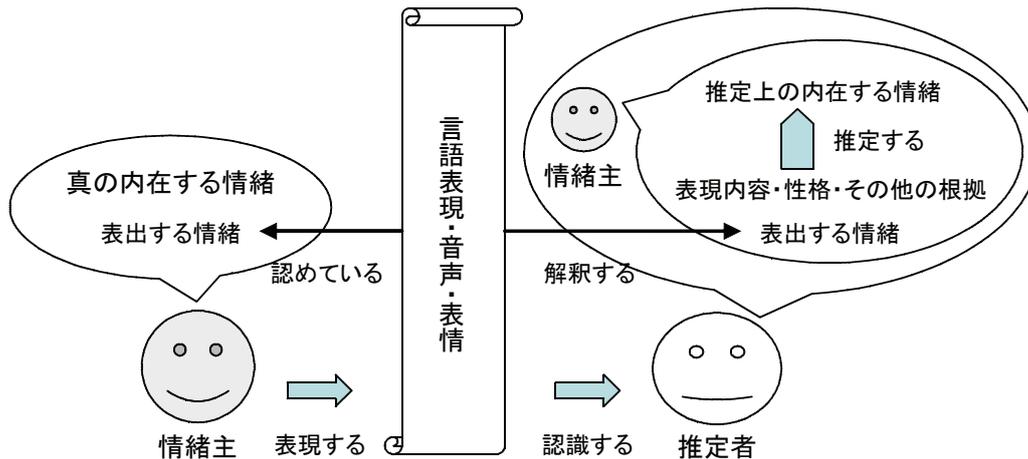


図 2.1 情緒の3つの存在場所

2.1.4 情緒推定の正解

言語表現からの情緒推定を計算機で実施する際、本研究では、工学的な立場から「推定上の内在する情緒」を追求することにする。そして、その推定の正解を、複数の推定者が共通的に推定する情緒とする。

認知科学や心理学においては、「真の内在する情緒」が主たる議論の対象である。そのため、表情、汗、心拍、生体電位など情緒状態と直接関連するような生理的な情報が参照されている。これに対して、工学的な立場では、「真の内在する情緒」が議論の対象となることもあるし、「推定上の内在する情緒」が議論の対象となることもある。たとえば、「真の内在する情緒」は、カウンセリングや物語朗読のようなタスクにおいて、患者の言い尽くせない気持ちを理解する上で追求されるべきであり、登場人物の物語性を表す上で追求されるべきである。一方、「推定上の内在する情緒」は、実際的な問題の捉え方である。たとえば、コールセンターなどでのクレームの解析において、計算機が人間の受け付け係りの代理をする際、受け付け係りが推定する情緒を、計算機も推定できなければならない。また、電子メールや Web におけるテキスト・コミュニケーションにおいて文章の感情的な誤解を読み手に与えないように配慮するための支援ツールを想定すると、文章の話し手本人の情緒として、聞き手がどのように推定するのが関心の的になる。このように推定上の内在する情緒を追求することも有益な言語処理技術の一つである。

「推定上の内在する情緒」は、推定者の主観的な判断に強く依存するため、正解が定め難い。しかし、計算機を人間の推定者の代理とすることや、言語理解により標準的な情緒の意味情報を提供することを想定すると、まず第一に実現すべきは、複数の推定者に共通的に推定される情緒を推定することである。それができた上で、少数意見として推定される情緒を漏らさず推定することが要請されるであろう。

2.2 情緒推定方法の分類

言語表現からの情緒推定では、言語表現に関わる人物が情緒主となり得るので、情緒推定方法を分類する上で、人物の種類を列挙しておく必要がある。また、情緒推定の原理では、情緒過程と言語表現の関わりに着目するので、言語表現の解釈の種類を確認しておく必要がある。そこで、本節では、情緒主の種類、および、言語表現の解釈の種類をまとめた上で、情緒推定の方法を分類

する.

2.2.1 情緒主の種類

言語表現を中心に、関連する人物を列举すると次の4種類が挙げられる:

- **話し手**: 言語表現を作成する人物
- **登場人物**: 言語表現に現れる人物
- **聞き手**: 言語表現の受け手のうち、直接の効果が生じる人物
- **傍観者**: 言語表現の受け手のうち、直接の効果が生じない人物

「話し手」は認識したものを言語表現で表す人物である。話し手の認識に、ある人物が含まれており、言語表現に表されることがある。この人物を「登場人物」という。次に、言語表現を受け取る人物として、2種類がある。1つは言語表現の効果が直接生じる人物である。たとえば、命令文で書かれた言語表現を受け取ると、その人物はその命令に従うべきであるかどうかを判断する。本研究では、このような人物を「聞き手」と呼ぶことにする。一方、言語表現の直接の効果が生じない人物は、話し手に対して第三者である。本研究では、このような人物を「傍観者」と呼ぶことにする(図 2.2)。

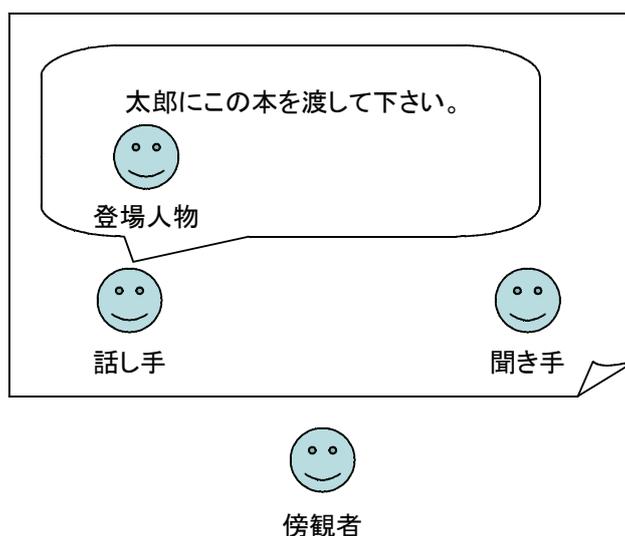


図 2.2 言語表現に関わる4種類の人物

ここで、これらの人物の分類は、固定的でないことに注意しよう。たとえば、文章において、話し手は1人称の代名詞「僕」や「私」により登場人物となる。同様に、聞き手は2人称の代名詞「きみ」、「あなた」により登場人物となる。また、引用節などにおいては、話し手は、文中の節に対して傍観者の立場をとることがある。たとえば、「事務所に戻ったら電話をくれと伝言を頼まれた」という文において、話し手は、伝言をするかどうかの判断だけが発生し、「電話をする」ことへの命令についての判断は発生しないので、この引用節に対しては傍観者である。

2.2.2 言語表現の解釈の種類

言語表現の解釈には、言語表現の文字通りの解釈と、追加的な推論による解釈の区別があるという[49]。本研究ではこれらを次のように呼ぶことにする：

- **明示的な解釈：** 言語規範に基づく言語表現の解釈
- **暗示的な解釈：** 使われた場に依存する言語表現の解釈

明示的な解釈に関して、日本語の表現は、「客体的表現」と「主体的表現」に分けられ、前者により言語表現の対象が表され、後者により対象に対する話し手の主体性が表されるという[47]。このことより、言語表現のうち客体的表現からは、登場人物の状況が解釈され、主体的表現からは、話し手の情緒などが解釈されると予想される。

一方、暗示的な解釈は、文章や対話による一連の表現を理解することである。その結果、登場人物や話し手の置かれた状況、および、話し手や聞き手の対話状態が得られる。これらから、登場人物、話し手、聞き手についての情緒が推定されると予想される。また、傍観者の立場からは、文脈や対話状態を物語の理解としてとらえ、登場人物、話し手、聞き手の心情に共感したり、ほくそ笑んだりすることが考えられ、このことから傍観者の情緒が推定できると予想される。

2.2.3 情緒の推定方法のタイプ

情緒の推定方法は、情緒主、言語表現の解釈、および、情緒的過程という3つの軸で区別できる。表 2.1 に関係をまとめる。表中の C,S,R は、それぞれ、情緒生起原因、情緒状態、情緒反応を指し、言語表現の解釈がこれらと関係することにより情緒が推定できることを表している。

幾つかの注意点を説明する。第1に、客体的表現から話し手の情緒が推定されるとしていることについて、話し手が「私」という代名詞により言語表現に登場する場合については、話し手は登場人物として扱われる。しかし、ここの注意点は、話し手はそれ以外にも言語表現に関わりを持つことから逃れられないことである。たとえば、「来る」という動詞や「ここに」という副詞において、話し手の視点を除去することができない。

第2に、主体的表現に対して聞き手の情緒推定を対応づけていない。たとえば、フォーマルな場で親しくない人が敬語を使わないとき、聞き手があまり良い感情を持たないことがある。これは、たとえば、主体的表現として用いられる「～ですよ。」という表現は、馴れ馴れしく不快を引き起こすことのある表現の一例として挙げられる。この表では、こうした話し手と聞き手の関係を対話性の範疇で扱うことにしている。

第3に、物語性から傍観者の情緒を推定する上では、「C : 情緒生起原因」しか手がかりにならない。それは、言語表現に傍観者の情緒状態や情緒反応が出現し得ないためである。

表 2.1 情緒推定方法に関連する要素の関係

解釈 情緒主	客体的表現	主体的表現	文脈※	対話性	物語性
登場人物	C/S/R	-	C/S/R	-	-
話し手	(C/S/R)	C/S/R	C/S/R	C/S/R	-
聞き手	-	-	-	C/S/R	-
傍観者	-	-	-	-	C

※ 重文・複文を含む

本研究では、主として言語表現の解釈の種類に基づき、情緒推定の方法を次の5タイプに分類する。

- **タイプ1**：客体的表現に基づき、登場人物の情緒を推定する方法
- **タイプ2**：主体的表現に基づき、話し手の情緒を推定する方法
- **タイプ3**：文脈に基づき、登場人物・話し手の情緒を推定する方法
- **タイプ4**：対話性に基づき、話し手・聞き手の情緒を推定する方法
- **タイプ5**：物語性に基づき、傍観者の情緒を推定する方法

2.2.3.1 タイプ1

客体的表現が情緒的な過程と対応することで、客体的表現の登場人物の情緒を推定する方法がタイプ1である。情緒的過程の3種類についての例を見てみよう。

- 太郎が大学に合格する。(情緒生起の原因の発生)
- 太郎が喜ぶ。(情緒生起の状態)
- 太郎が万歳をする。(情緒反応の発生)

登場人物は「太郎」である。太郎を情緒主とすると、各文は、情緒的過程を表しているので、太郎の情緒が推定できる。

まず、原因を表す表現について、「合格する」という事象は、通常、努力の末に得られる結果である。目良らの手法では、情緒主(太郎)が大学に対して「+」の好感度を持つとき、「合格する」という事象に対する好感度は、「動作主自身に対する好感度 * 対象に対する好感度」により定まるとしている[8]。土屋らの方法では、大きな規模の常識推論により事象に対する感情が推論されるとして、たとえば、「交差点で転ぶ」という事象は、「交差点は人の多い場所として連想され」、「その場所での失敗は恥である」という推論がなされる[50]。しかし、これらの推論を行うための用言の網羅的な知識ベースは構築されていない。また、こうした事象と情緒の関係に対する知識の収集は、Liu らや遠藤らの方法があるのだが、「情緒主」、「情緒対象」、「情緒の種類」という情緒的な属性について議論されていない[9],[48]。

次に、情緒生起の状態について、「喜ぶ」や「頭にくる」のように情緒状態を明示する単語や慣用表現が存在する。これは一般のソーラス(たとえば、[13],[37],[38])において、「感情」のカテゴリの下に網羅的に配置されている。さらに、情緒主や情緒対象といった属性についての知識ベース化が既に進められている[45]。

情緒の反応について、表情など身体的な状態や変化を伴う場合については、情緒状態を明示する表現としてソーラスに収録されることが多い。たとえば、「(恐れにより、)顔色を変える。」のような表現は反応であるが、感情状態・感情動作の下位に配置される。これは、言語学的に語彙を考える上で、感情に関わるか否かが関心の的であり、感情の原因であるか、感情自体であるか、あるいは、感情の結果であるかについての分類に関心が無いためである。一方、常識的には感情を伴った行為であったとしても語彙として感情に関わらない場合には、ソーラスに感情のカテゴリ下に収録されることはない。たとえば、「(怒って、机を)蹴飛ばす。」のような表現は感情の下位には配置されず、身体動作の下位に配置される。このように、現在のところ、情緒の反応については、網羅的な知識ベースが存在しない。

2.2.3.2 タイプ2

主体的表現が情緒的な過程と対応することで、話し手の情緒を推定する方法がタイプ2である。

ここで、主体的表現と客体的表現との境界を明確にしようとする諸説ある(たとえば、述語に後続する文末表現の分類[51])。本研究は、その境界を明確にすることは目的としないが、目安として、話し手の持つ内的な基準により表現が選択されるものを主体的表現とみなしている。たとえば、助詞・助動詞は主体的表現ではあるが、さらに、修飾語句についても程度の判断が話し手の基準に基づく場合には主体的表現として扱う。以下に例を示す。

- 雨に降られてしまった。(文末表現)
- そのことは嬉しい。(形容詞述語)
- あいにく課長は留守です。(修飾表現)

いずれも話し手の「情緒生起の状態」を表す。「～れてしまった」は被害や迷惑を表す。感情を表現する形容詞は感情の直接的表出や感情的品定めという分類が既になされている[44]。また、副詞や形容詞・形容動詞の連用形により、述語に対する話し手の態度も伺える。

ただし、形容詞述語や修飾表現の位置付けは、場合によってはタイプ1で扱うことがある。なぜならば、これらの属性を判断したものが登場人物である場合があるからである。たとえば、「娘は、そのことが嬉しかったので、家に帰って直ぐに親に話した。」という文では、「嬉しい」と判断したのは「娘」であり、話し手であるとは限らない。

副詞、形容詞、擬音語・擬態語という単語についての感情的なイメージは、既に相当表現も含めて網羅的に示されている[40],[41],[42]。形容詞について、評判分析の分野では、評価対象に依存して、PositiveとNegativeが判定されるので、ドメインに依存した辞書作りの方法が提案されている[52]。また、文末表現による情緒的な判断の変更はヒューリスティクスで対処する方法が提案されている[45]。付属語や接続表現に関しては、文型辞典として解説がある[46]。

2.2.3.3 タイプ3

タイプ1とタイプ2が、明示的な解釈に対するものであった。タイプ3からタイプ5は、暗示的な解釈も考慮に入る。文章や対話の理解では、文脈に整合性のあることが前提となっており、対話の発話文や文章中の各文の解釈は蓄積されて、全体が解釈される。その解釈が情緒的過程と対応することを根拠に、情緒を推定する。

タイプ3では、重文・複文の関係、および、文と文の関係も考慮に入れる。以下に例を示す：

文1：僕は大学の入学試験に合格した。

文2：でも、それは第3志望の大学だった。

文1だけを参照して情緒を推定するものがタイプ1であった。タイプ3では、文1と文2の両方を参照して情緒を推定する。文2の冒頭の接続詞「でも」が自然だと感じる理由は、文1だけからは登場人物の喜びが推定され、文2においてその喜びが最高のものではないことが想像されることに依る。

重文・複文や文脈を扱う上では、言語理解の過程で得られる情報の内部表現(representation)が計算機処理の上で重要になると考えて、直接、内部表現の問題に取り組んだ研究として、前述した[7],[35]の研究が挙げられる。また、文脈の一貫性と接続表現による評価極性の変化に着目し、好不評を表す語彙を収集する方法が提案されている[53]。

なお、接続表現そのものが情緒を表す場合がある。たとえば、「～したにもかかわらず～」というとき、努力に相応しい結果が得られていないことが伺える。これは、タイプ2と位置付けることもできるが、前後の文脈(あるいは、従属節と主節の関係)の解釈により情緒推定が補強されることから、タ

タイプ3と位置付けることもできる。このように境界が不明瞭な場合がある。

2.2.3.4 タイプ4

タイプ4では、対話を扱う。対話者の意図を1文単位で取り出したり、一連の発話文から取り出したりするため、計算機処理の方法としてはタイプ1～3と重なる部分がある。

以下に対話の状態から情緒を推定する例を示す：

1A：明日ピクニックに行くんだ。

2A：天気は良さそうだよ。

3A：一緒に行かないか？

4B：わあ、いいねえ。

A者の始めの2つの発話1Aと2Aから、「ピクニックをする場所の天気が良い」ということが解釈される。A者の計画の評価が良いので、A者に情緒生起の原因が成立しており、「期待」を生起させていると推定できる。続いて3Aにより対話相手B者の対話の状態として「誘われていること」、「楽しい計画が発生すること」が成立し、これが原因となってB者に「喜び」や「期待」が生起すると推定できる。4Bでの「わあ」はその気持ちを直接的に表した表現である。すなわち、タイプ2として情緒生起の状態を表したものである。むしろ、タイプ3としては、4Bの「わあ」という発声の生じたこと、および、同意する行為自体が情緒反応の発生と、とらえられる。

対話の状態は、対話プランニングとして扱われる。Rosisらは、医療説明対話におけるプラン知識に、対話相手に生じる情緒を組み込んで、対話相手の情緒の推定とそれに応じた対話生成を行った[54]。

2.2.3.5 タイプ5

対話や文章を傍観者の立場でとらえ、その解釈が傍観者の情緒的な過程に関わることを基にして情緒を推定する。

タイプ4の対話例で、「傍観者」がA者やB者に肩入れしているならば、2人の成り行きに共感し、わくわくしながら読み進めるかもしれない。つまり、物語として理解した状態において情緒生起の原因が発生するかどうかにより、傍観者の情緒が推定できる。

物語の感情的な構成について、Lehnertは情緒的な+/-/m(中立)の属性を持つplot unitにより「復讐」など感情的な観点で物語を要約する方法を示している[25]。傍観者はそうした構造に対して自身の感情を抱く。ここで、物語の登場人物の情緒に対して傍観者が共感するかどうかは、登場人物と傍観者の関係によるもので、目良らや松本らは2人の人物の情緒を+と-で表し、その掛け算的な計算で説明した[8][55]。また、感情移入も考慮してより詳しく分類する議論もある[56]。

2.3 本研究で実現する情緒推定

第1章で述べたとおり、本研究で目指す情緒推定では、言語の表層的な構造を情緒推定の手がかりとする。これは、言語理解の基本タスクに、情緒的な意味的情報を提供するという目的があるからである。この方式は、5つのタイプのうち、タイプ1あるいは2に該当するものである。

前述のとおり、タイプ1と2において、言語知識ベースの整備が遅れているのは、「原因」に関するタイプ1の情緒推定のための知識ベースである。そこで、本研究では、原因に基づくタイプ1の情緒推定の方法の確立とそのための網羅的な言語知識ベースの構築を目標とする。

第3章 人間による情緒推定の共通性

人間による情緒推定の能力を調査する。第2章で述べたとおり、情緒推定の正解は、人々の間で共通的に推定される情緒であるとした。そこで、本章では、実際に複数の人間に情緒推定を行ってもらい、どの程度共通的に情緒が推定されるのかについて定量的に調査する。具体的には、情緒タグ付きコーパスの構築という方法により、情緒タグの一致率を用いて評価値を得ることを提案する。そして、テキスト対話と日記文において調査を実施する。

3.1 調査の目的と方法

3.1.1 目的

言語表現から読み取られる情緒は、読み取る人によって異なることもあれば、人々の間で共通して読み取られることもある。本研究の計算機による情緒推定は、共通して読み取られる情緒を求めようとしている。しかし、共通した情緒の存在する割合については明確ではなく、計算機による情緒推定においてどれだけの精度が得られれば妥当であるのかも分かっていない。そこで、本章では、人手による情緒推定の共通性を定量的に調査する。

さらに、この調査を通じて、共通性の得られない理由について検討する。予想される理由として、(i)言語理解の過程に違いがあること、(ii)言語表現だけでなく音声や表情などの非言語情報も必要であること、(iii)言語表現の表す内容に依存すること、などが考えられる。そこで、タグ付きコーパスを作成し、タグの比較・分析により共通性の度合いと相違の理由を明らかにする。

3.1.2 方法

下記の3種類のコーパスを作成する：

- (コーパス1) 情緒生起原因を含む情緒タグを付与したテキスト対話コーパス
- (コーパス2) 表情タグと情緒タグを付与したテキスト対話コーパス
- (コーパス3) 情緒タグ付き日記コーパス

コーパス1は、寓話におけるテキスト対話を題材として、言語表現から情緒を推定する過程をタグで記述する。具体的には、まず、1人の推定者がタグ付与を行う。対話文の発話意図、および、非情緒的な心的状態を表すタグを付与する。次に、それらから情緒生起の原因を見い出すことで情緒を推定し、情緒とその原因を表すタグを付与する。このことより、情緒とその原因を表すタグがあれば情緒推定の過程が明らかになることを確認する。次に、別の推定者が情緒とその原因を表すタグを付与し、両者の共通性を調べることで情緒推定の過程の違いを分析する。

コーパス2は、漫画におけるテキスト対話を題材として、表情と情緒のタグを付与する。2人の推定者が付与した情緒タグの一致を調べることで、表情参照時の情緒推定の共通性を計測する。次に、表情を参照せずに情緒タグを付与した場合の共通性を計測する。表情参照の有無により共通性がどの程度変化するかにより、表情参照の重要性を分析する。

コーパス3は、日記文を題材として、日記文の節の単位に、複数の推定者が情緒タグを付与する

とともに、各節の表現に対して、情緒的過程(情緒原因・情緒状態・情緒反応)の分類をする。3つの事態の表現される頻度、および、その表現における情緒推定の共通性を調査する。

3.1.3 タグの設計

3つのコーパスにおいて、情緒タグ、表情タグ、心的状態タグ、対話行為タグを使用する。ここでタグの意味を説明する。

1) 情緒タグ

情緒タグは、以下に示すような9分類系、5分類系、および、3分類系の3系統とする。本研究では9分類系を基本とする。5分類系と3分類系は、情緒の種類に対する粒度の問題を考察することや、関連研究と比較することに使用する。

(9分類系) 9分類系の情緒タグは、次の9種類とする：

《喜び》, 《悲しみ》, 《好ましい》, 《嫌だ》, 《驚き》, 《期待》, 《恐れ》, 《怒り》, 《なし》

Plutchik の基本情緒[29]を参考にした8種類と、情緒の無い状態《なし》である。

(5分類系) 5分類系の情緒タグは、次の5種類とする：

《P》, 《N》, 《A》, 《S》, 《なし》

9分類系の微妙な差異を吸収してタグを分析するために、5分類系を用意する。《喜び》, 《好ましい》, および、《期待》は《P》に、《悲しみ》および《恐れ》は《N》に、《嫌だ》および《怒り》は《A》に、《驚き》は《S》に、そして、9分類系の《なし》は《なし》に、それぞれまとめた。このまとまりは、日本人が情緒を認識する際、類似するものとして認識されることがいわれており[57]、本研究はそれに倣って定めた。

(3分類系) 3分類系の情緒タグは、次の3種類とする：

《Positive》, 《Negative》, 《なし》

現在、多くの研究において使用されているものは、ポジティブとネガティブであり、中立や平静、もしくは、ポジティブとネガティブの判断のつかない場合(つまり不明や不定)を「なし」とすることが多い[58]。本研究では、《喜び》, 《好ましい》および《期待》を《Positive》, 《悲しみ》, 《嫌だ》, 《恐れ》および《怒り》を《Negative》, 《驚き》および9分類系の《なし》を《なし》とそれぞれまとめた。《Positive》と《Negative》の両方が認められる場合は、複数のタグを付与する。

(原因ラベル)

情緒タグには、情緒生起原因のラベルを補足することがある。原因ラベルは第5章で定義する。簡単な例を挙げる。《喜び, 獲得》という9分類系のタグに「獲得」という補足をしているとき、「目標を実現することに必要な物を手に入れた」ということが原因となって喜びが生じたことを、このタグは表している。

2) 表情タグ

表情タグは、次の8種類とする：

〈幸福〉, 〈嫌悪〉, 〈悲しみ〉, 〈驚き〉, 〈恐れ〉, 〈怒り〉, 〈背後〉, 〈なし〉

はじめの6種類は Ekman の分類に基づく表情である[26]。残りの1種類の〈背後〉は、後に扱う表情が漫画における表情であるため用意した。漫画では、後ろ姿の描写であっても、「青ざめ」や「冷や汗」などを伴って描かれることがある。これは表情に相当するほど登場人物の情緒を表すので、表情タグでカバーすることにした。

表情タグは、顔の形状が対応する情緒を表すタグである。情緒主は、意識下では真の内在于る情緒が表情に素直に表出するが、意識して表情を選択することもできるので、表情に表出する情緒は真の内在于る情緒とは必ずしも一致しない。なお、素直な表出とは、《喜び》、《好ましい》、《期待》に対して〈幸福〉が対応し、《嫌だ》に対して〈嫌悪〉が対応する。〈悲しみ〉、〈驚き〉、〈恐れ〉、〈怒り〉については文字通りに対応する。

3) 心的状態タグ

情緒の生起する過程を追跡する上で、情緒以外の心的な状態をタグとして明示する必要がある。本研究では心的状態タグとして、次の12種類を用意する：

〔生理〕、〔欲求〕、〔情緒〕、〔目標〕、〔プラン〕、〔予測〕、〔評価〕、〔記憶〕、〔認識〕、〔行動〕、
〔言語〕、〔信念〕

心的状態の種類は、心のモデル[59]で用いられている主要な構成要素である。これらの心的状態の活性化連鎖により、「イソップ物語」の「きつねとぶどう」のシナリオが再現できる。このことから、情緒の生起する過程を追跡するためのタグとして有効性がある。

ここで、典型的な活性化連鎖について説明する。「喉の渇き」という〔生理〕の状態変化により、「喉の渇きを潤す」という〔欲求〕が生じ、それが〔目標〕として設定される。〔目標〕を達成するために〔プラン〕が幾つか候補として作られる。〔プラン〕の作成には過去の状況を〔記憶〕から取り出して使用する。〔プラン〕の実行時の様子を〔予測〕して、〔評価〕を行う。〔評価〕に基づき〔プラン〕の採否が決まる。採用された〔プラン〕は、周囲の状況などの〔認識〕を行いながら〔行動〕を1つずつ実行する。これらの心的状態の変化は、〔情緒〕の生起に結び付く。一方、〔言語〕は、対話における発話文に対応し、さらに、意図や対話状態に対応する。〔信念〕は、対話相手のモデルをとらえるために用いる。

4) 意図タグ

発話意図を表すタグは対象タイプと行為タイプの組で定義する。たとえば、「お腹が空いたよ。」という発話文は、〈生理の伝達〉という意図タグが対応する。具体的に各タイプは次の種類のものとする。

対象タイプは、次の11種類とする：

「生理」、「欲求」、「情緒」、「目標」、「プラン」、「予測」、「評価」、「記憶」、「認識」、「行動」、
「その他」

行為タイプは、次の9種類とする：

「質問」、「伝達」、「確認」、「肯定」、「否定」、「要求」、「受諾」、「拒否」、「その他」

対象タイプにより、発話文が表す心的状態を明示する。そのため心的状態タグの要素で構成する。行為タイプは、対話解析において一般的に用いられるものである(たとえば[60])。

3.1.4 評価式

コーパスに付与されたタグの分布を評価するために、目的やタグ付与の条件に応じて、幾つかの式が使用されている。本研究では、関連研究と比較できるように、カッパ値(κ)、一致率、同意率、再現率、および、適合率を用いる。

1) カッパ値(κ)

2名のタグ付与者によるタグの一致する割合を評価する式として、Cohen の κ 値がある[61]。単純にタグの一致する確率ではなく、偶然に一致する確率を差し引いている。主観的な判断によるタ

グの一致の評価にしばしば使用される[11]. κ 値は次式で計算する:

$$\kappa = (P_o - P_c) / (1 - P_c) \dots\dots\dots(3.1)$$

P_o は 2 名の付与者によるタグの一致数の割合である. P_c は偶然の一致の期待値の割合である.
 たとえば, 2名のタグ付与者 A 者と B 者が, 3種類のタグ a, b, c を1つの対象につき各者1つずつ付与した結果が表 3.1 のように集計されたとする. たとえば, 付与者 A がタグ a をつけたところに付与者 B が a をつけた回数は8である. このとき, 一致数の割合は, $P_o = (8+7+9) / 34 \doteq 0.705882$ であり, 偶然の一致の期待値の割合は, $P_c = (11*12/34 + 12*10/34 + 11*12/34)/34 \doteq 0.332180$ である. したがって $\kappa \doteq 0.559585$ となる.

表 3.1 タグの集計結果の例

		付与者 B			
		a	b	c	計
付与者 A	a	8	2	1	11
	b	3	7	2	12
	c	1	1	9	11
	計	12	10	12	34

2) 一致率

κ 値は1つの対象に対して1つのタグを付与する際の評価指標である. 1つの対象に複数のタグを付与する場合には, そのまま用いることはできない. そこで次の評価式を用いる場合がある[62]:

$$\langle \text{一致率} \rangle = \frac{2 N(S_a \cap S_b)}{N(S_a) + N(S_b)} \times 100 (\%) \dots\dots\dots(3.2)$$

ここで, S はタグの集合を表し, $N(S)$ は集合 S の大きさを表す. $S_a \cap S_b$ は付与対象が同一のところで同一となるタグの集合を表す. 2つのタグ集合が完全に一致する場合に一致率が1となり, 全く異なる場合に一致率が0となる.

3) 同意率

情緒のように主観的に決定するタグの場合, 2名の付与者が独立にタグを付与し一方の者が他者のタグを参照したとき, 最初に自身が付与したタグと一致しなくても, その他者のタグも適切であるといえることがある. このとき他者のタグにおいて同意の得られた割合が評価の対象となる. そこで次の評価式が使用できる:

$$\langle \text{同意率} \rangle = \frac{N(A(S))}{N(S)} \times 100 (\%) \dots\dots\dots(3.3)$$

ここで, $A(S)$ は, 集合 S において同意の得られたタグを部分集合として抽出する関数である. 同意の得られるタグの集合があらかじめ決定できるならば, 同意率は以降に説明する適合率と同じ評価値となる. しかし, 適合率と異なる名称で評価式を与えているのは, 同意の得られるタグの集合があらかじめ決定し難いためである. さらに, 同意を得る判定が厳しくないため, 適合率よりも高い値を示し易い傾向がある. したがって, 異なる名称を与えておくことが望ましい.

4) 再現率と適合率

1)から 3)まででは、タグ付与者間の一致に対する評価式を示した。一方、正解のタグが与えられるとき、正解タグと付与したタグの一致する割合について次の2つの評価式を使用する。

$$\langle \text{再現率} \rangle = \frac{N(S_c \cap S_i)}{N(S_c)} \dots\dots\dots(3.4)$$

$$\langle \text{適合率} \rangle = \frac{N(S_c \cap S_i)}{N(S_i)} \dots\dots\dots(3.5)$$

ここで、 S_c と S_i はそれぞれ、正解タグの集合と付与したタグの集合である。

3.2 情緒タグ付きコーパスの作成と評価

3.2.1 テキスト対話への情緒生起原因タグの付与

3.2.1.1 設定

従来の情緒の計算モデルでは、情緒主の心的状態としてゴールやプランの達成状況が注目されていた。このモデルを参考に情緒推定を行うならば、対話の発話文からプランなどの非情緒的な心的状態を抽出し、情緒生起の原因の成立を検出することで、情緒が推定できると考えられる。したがって、発話文から情緒が推定されるまでの間を、意図タグ、非情緒的な心的状態タグ、および、情緒タグで明記してコーパスを作成する方法が考えられる。

情緒生起の原因を見分ける基準について、詳細は第5章で説明するので、ここでは簡単に具体例を示す。図 3.1 の if-then ルールは、「目標実現に有効な(評価の高い)プランを得た」ことを条件に「喜びが生起する」ことを導く。こうした判断のためのルールは約 120 種類用意している。

if 情緒主(F)が目標(G)を持つ、かつ、 情緒主(F)が目標(G)を実現するためのプラン(P)を持つ、かつ、 プラン(P)の評価(E)が良い then 情緒主(F)に情緒(《喜び、立案》)が生起する
--

図 3.1 情緒を見分けるための if-then ルール(人間用)の一例

ここで、このルールを熟知した者は、情緒生起の原因を表すキーワード(たとえば、《喜び、立案》のうちの「立案」の部分)が明示されているならば、意図タグおよび非情緒的な心的状態タグは、逐一付与されていなくても把握でき、情緒推定の過程を分析することができる。すなわち、原因を表すキーワードと情緒のタグを付与することが最も重要である。

そこで、コーパスの作成は、ルールを熟知した2名の大学院生が行う。1人目は、基本通りに意図タグ、非情緒的な心的状態タグ、および、情緒タグの付与を行いながらコーパスを作成する。その結果より、基本的な情緒推定の手順が成立するか否かを確認する。2人目は、原因と情緒名のタグを付与することとして、1人目がその結果に対して共通性を評価する。

さて、タグ付与の対象とする対話は、「勧誘の対話場面」である。ここでは「相手に行為をしてもらうこと」を「勧誘」と呼んでいる。対話をする人物は「勧誘者」と「被勧誘者」という立場になる。情緒推

定は、どちらか一方の立場の人物を情緒主として解析を進める。

タグの付与は以下の手順による：

- (1) 情緒主の発話文にタグを付与する場合：
 - (1-1) 発話の意図を解析し、意図タグを付与する。
 - (1-2) (1-1) の意図に起因して、非情緒的な心的状態を推定し、心的状態タグを付与する。
 - (1-3) これまでの心的状態より、補完的な心的状態を推定し、心的状態タグを付与する。
 - (1-4) 心的状態に対して情緒生起原因を見い出して、情緒を推定し、情緒タグを付与する。
- (2) 情緒主ではない人物の発話文にタグを付与する場合：
 - (2-1) 発話の意図を解析し、意図タグを付与する。
 - (2-2) (2-1)の意図に起因して、情緒主に生じる非情緒的な心的状態を予測し、心的状態タグを付与する。
 - (2-3) これまでの心的状態より、補完的な心的状態を予測し、心的状態タグを付与する。
 - (2-4) 心的状態に対して情緒生起原因を見い出して、情緒を予測し、情緒タグを付与する。
- (3) (1)または(2)に戻る。

3.2.1.2 タグ付与

「イソップ物語」や「日本昔話」などの対話から、「勧誘」の場면을抽出して、タグ付与を行った。その中から「アリとキリギリス」におけるタグ付与の様子を図 3.2 に示す。ここでは、7ターン目から10ターン目までのやりとりを記載している。キリギリスを情緒主として、情緒推定を行っている。

7Gr : 「アリさん、僕、食べるものがなくて困っているんだ。」 〔言語, 生理の伝達〕 〔生理, 空腹〕 〔目標, 空腹を満たす〕 《悲しみ, 内的な不快》
8Ant : 「冬になるまで食べ物を集めなかったの？」 〔言語, キリギリスの行動に対する非難〕 〔記憶, アリがキリギリスに食べ物を集めることを提案した〕 〔記憶, キリギリスが食べ物を集めなかった〕 《悲しみ, 非難》
9Gr : 「そうなんだ。ああ、寒くて死にそうだ。」 〔言語, 生理の伝達〕 〔生理, 寒い〕 〔目標, 暖まる〕 《悲しみ, 外的な不快》 《恐れ, 生理異常》
10Ant : 「お家の中に入るかい？」 〔言語, プランの要求〕 〔プラン, キリギリスがアリの家に入る〕 〔予測, 部屋は暖かい〕 〔評価, 非常によい〕 《喜び, 立案》 《期待, 成算》 《喜び, 保護》 《喜び, 協力》 《好ましい, 保護》 《好ましい, 協力》

※ Gr はキリギリス, Ant はアリであり, 数字はターン数である

図 3.2 心的状態と情緒タグの付与の例

7Gr によると、キリギリスの生理状態が悪いと推定される。これにより、「内的な不快」により《悲しみ》が生起していることが推定される。8Ant によると、アリがキリギリスに、過去の行動について確認し、「非難」をしていると読み取ることができるので、キリギリスは「非難」されたことにより《悲しみ》が生起すると推定される。9Gr によるとキリギリスが寒さを感じていることがわかるので、「外的な不快」により《悲しみ》が、また、その程度が甚だしいことから「生理異常」により《恐れ》が生起していると推定される。10Ant によるとキリギリスの行動を促す発話であるのでキリギリスにプランが成立する。さらに、キリギリスは「アリの部屋が暖かく、良いプランを得た」と信じていると推定される。このことは〔目標、暖まる〕も考慮すると「立案」による《喜び》や、「目標の実現に成算があること」による《期待》の推定に結び付く。また、アリに対して「保護」や「協力」の得られたことによる《喜び》や《好ましい》という情緒の生起が推定される。

上述の方法により、1名の熟練者が付与したコーパスの規模を表 3.2 にまとめる。勧誘の対話場面を 29 対話抜き出してタグ付与を行った。242 ターンで構成されており、そのうち、78 箇所は、情緒的な状態を直接的に表していた。これ以外の箇所についてタグ付与を行ったところ、377 個の情緒タグを付与することができた。その内訳を表 3.3 に示す。377 個のうち 133 個 (35%) は《喜び》であった。最も少なかったのが《恐れ》であった。「勧誘」という場面であったので、《恐れ》や《怒り》などが少なかったと思われる。また、情緒を見分ける基準が、125 種類用意されていたが、そのうち 52% を使うことで対話における情緒が説明された。情緒タグは、比較的広範囲の事例を扱うことができるという見通しを得た。

表 3.2 1人目のタグ付与結果

項目	規模
対話の場面数	29 対話
ターン数	242 ターン
情緒タグの付与数	377 個

表 3.3 情緒タグの付与数と使用率

情緒タグ	付与タグ数(付与率)	使用率 (付与された種類数 / 定義された種類数)
喜び	133 (35%)	78% (14 / 18)
好ましい	88 (23%)	69% (11 / 16)
期待	53 (14%)	29% (4 / 14)
悲しみ	38 (10%)	61% (11 / 18)
嫌だ	35 (9%)	61% (14 / 23)
怒り	11 (3%)	22% (4 / 18)
驚き	10 (3%)	40% (2 / 5)
恐れ	9 (2%)	38% (5 / 13)
合計	377 (100%)	52% (65 / 125)

※ 定義された種類数については、第 5.1 節および付録 2 を参照

3.2.1.3 共通性の評価

2人の付与者が同一の対話に対して独立に情緒タグを付与した際のタグの共通性を評価する。評価の尺度は、カッパ値、一致率、および、同意率を用いる。カッパ値および一致率は両者の付与したタグ同士を比較する。原因と情緒名が完全に一致する場合、および、情緒名が一致する場合

の2通りについてそれぞれ集計する。一方、同意率は、1人目の付与者が2人目の付与者によるタグに対して同意した割合として集計する。

2人目の付与者は、情緒生起原因を熟知しているので、対話を読みながら非情緒状態を想定して情緒タグを付与する。したがって、情緒生起原因の違いのある箇所は、情緒推定の過程が2者の間で異なっていることになる。なお、情緒タグを付与しなかった部分は《なし》として集計する。

2者が付与したタグを比較した結果を表 3.4 にまとめる。付与した対話は 17 対話 (179 ターン) である。1人目が付与したタグは 339 個、2人目が付与したタグは 349 個であった。一方、原因ラベルを無視して情緒名だけをみると、1人目が付与したタグは 276 個、2人目が付与したタグは 286 個であった。たとえば、図 3.2 の最後に付与した情緒タグに「《好ましい, 保護》, 《好ましい, 協力》」がある。「保護」と「協力」は原因を表す。原因を無視すると《好ましい》であり、重複する数は数えないこととした。なお、対話の1ターンにつき複数のタグが付与されるため κ 値は参考値である。 κ 単独とは、1ターンに1つのタグが付与された箇所のみを集計した結果であり、 κ 複合とは複数のタグが付与された箇所は複数のタグが全て過不足なく一致することを一致と見なして集計した結果である。この表から次のことがいえる:

- 2名の付与者が付与したタグの一致率は、原因ラベルによる違いを考慮すると 39%, 原因ラベルによる違いを無視すると 58%であった。
- κ 値は、原因ラベルによる違いを考慮すると 0.24444 ~ 0.22037, 原因ラベルによる違いを無視すると 0.37580 ~ 0.34862 であった。
- 1人目の付与者が2人目の付与者のタグに対して同意できるかどうかを検査したところ、原因ラベル付き情緒タグに対する同意率は 93%であった。また、情緒名だけをみた同意率は 94%であった。

表 3.4 付与された情緒タグの比較結果

項目	1人目の 付与数	2人目の 付与数	一致率	κ 複合	κ 単独	同意率
原因ラベル付き情緒タグの付与数	339	349	39%	0.22037	0.24444	93%
原因ラベルを無視した情緒タグの付与数	276	286	58%	0.34862	0.37580	94%

※ 対話テキストの規模 17 対話, 179 ターン

原因ラベル付き情緒タグの一致率が低いことから、両者の間で、情緒を推定する過程に大きな違いがあることがわかる。しかし、原因ラベル付き情緒タグの同意率が 93%と高いことから、2人目の情緒推定過程に不自然さがあるわけではないことがいえる。すなわち、1人目の付与者が思いつかない観点で、2人目の付与者は情緒を推定していたことになる。

情緒推定過程に相違が生じた箇所を分析したところ、次のことが主たる原因であることが分かった:

- (a) 発話意図の解釈に関する付与者間の相違
- (b) 推論の深さに関する付与者間の相違

(a)について、発話意図の解釈には元々曖昧性のあることが知られている。たとえば、図 3.2 の 8Ant の「冬になるまで食べ物を集めなかったの?」という発話文に対して、「キリギリスを非難する」という発話意図ととらえる場合と、「キリギリスの状況を確認する」という発話意図ととらえる場合がある。前者の場合、キリギリスはアリの発話を受けて《悲しみ, 非難》や《嫌だ, 非難》という情緒タグが

付与された。一方、後者の場合、キリギリスは、アリが状況を把握してくれるので、アリがキリギリスに同情することや気遣って助けてくれること期待するため、《期待、同情》や《期待、保護》というタグが付与された。

(b)は、具体的には非情緒的な心的状態を推論する深さである。たとえば、図 3.2 の 7Gr の「アリさん、僕、食べるものがなくて困っているんだ。」という発話文からは、キリギリスの心的状態として、〔生理、空腹〕というタグが付与される。このことより《悲しみ、内的な不快》というタグが付与された。しかし、キリギリスの心的状態をさらに深く読み進めると、キリギリスはアリを頼ってきていることに気がつくため、キリギリスの情緒タグとして、《期待、協力》や《期待、成算》が付与された。

(a)や(b)の相違が生じる理由は他にも幾つか考えられる。基本的なものとしては、対話者の表情や音声などの非言語的情報が欠けていることが挙げられる。タグ付与者にとって、発話文から推定される情緒と表情などから推定される情緒が矛盾しない場合は適切な深さで推論が収束すると思われる。また、対話者の性格をタグ付与者が知っている場合にも性格と発言の関係から、同様に推論が収束すると思われる。たとえば、上述のキリギリスの性格を悪くとらえると、アリの親切心につけ込むような心的状態にまで推論が進む。

3.2.2 表情の参照可能なテキスト対話への表情・情緒タグの付与

3.2.2.1 設定

言語表現だけでなく、表情も参照すると情緒推定はどれだけ推定者間で共通するかについて調査する。推定者は、言語分析の6名の経験者(大学生、大学院生)である。

対象とする対話は、漫画「ちびまる子ちゃん」(第1巻～第10巻)である[63]。漫画を用いる利点は、漫画にはコマ単位で表情と台詞の対応がとられており、分析の単位が明確であること、および、漫画の登場人物の性格が推定者にわかりやすいことである。情緒推定において推定結果が一致しない理由の一つとして情緒主の性格を推定者が十分に知らないことが挙げられたが、漫画ではその問題が解消される。また、表情の識別について、現実の表情に対してはばらつきが生じるが、漫画の表情に対してはわかりやすく誇張して表情が描かれているのでばらつきが抑えられると予想できる。なお、漫画に描かれている対話は漫画家の創作による対話であり、現実のものではない。しかし、対話内容を見ると、交渉の仕方などは常識的な展開であるので、情緒推定を試みる上で不都合はないと判断した。

3.2.2.2 タグ付与

表情を参照した情緒推定とその過程の記述方法を説明する。推定者は、次の手順でタグを付与する：

2人の推定者は独立に以下の(1)と(2)を行う。

(1) 推定者は、漫画のコマの単位で登場人物に注目しながら、登場人物の表情を参照し、表情タグを付与する。

(2) 推定者は、漫画の台詞や文脈、および、登場人物の性格を考慮し、ならびに、表情タグも参照しながら、コマ単位に情緒タグを付与する。

ここで、(1)と(2)で付与したタグを一時的なタグとする。

(3) 2人の付与したタグを比較して、同一であるならばそれを正解タグとする。異なるならば、2人の協議により正解タグを決定する。

表 3.5 に作成したタグ付きコーパスの一部を示す。#1 と#2 は、同一のコマに「まる子」と「お姉ち

ゃん」が描かれていることを表している。X 者のタグと Y 者のタグは違いがあり、Y 者のタグが正解として採用されている。#4 は、まる子は描かれているが台詞がないことを表している。Y 者のタグ付与漏れがわかる。

構築したコーパスの規模を表 3.6 にまとめる。タグの規模として、正解タグの付与箇所数、表情タグ数、情緒タグ数を示す。「タグ付与箇所数」とは、コマ毎の登場人物のうち表情が平静でなかった者の数である。本コーパスではそのような者にタグが付与される。たとえば、表 3.5 でタグが付与された話者数はのべ 5 人である。

表 3.5 表情と情緒のタグ付きコーパスの一部(台詞は[63]より引用)

#	頁	コマ	話者	台詞	正解タグ		一時タグ			
					表情	情緒	X 者		Y 者	
							表情	情緒	表情	情緒
1	22	3	まる子	うちのもみの木は小さいね。	〈悲しみ〉	《悲しみ》	〈嫌悪〉	《嫌だ》	〈悲しみ〉	《悲しみ》
2			お姉ちゃん	仕方ないじゃん。	〈幸福〉	《悲しみ》	〈幸福〉	《悲しみ》	〈幸福〉	《悲しみ》
3		4	お姉ちゃん	ぎゃあっ！	〈驚き〉 〈恐れ〉 (青)	《驚き》 《恐れ》	〈驚き〉 (青)	《驚き》	〈驚き〉 〈恐れ〉 (青)	《驚き》 《恐れ》
4			まる子		〈驚き〉	《驚き》	〈驚き〉	《驚き》		
5	23	1	お姉ちゃん	まる子、あんたもみの木の鉢に金魚の死骸埋めたでしょ。						
6		2	まる子	そうだよ。						
7				だって肥料になると思っ。						
8			お姉ちゃん	やめてよ。	〈嫌悪〉	《嫌だ》	〈嫌悪〉	《嫌だ》	〈嫌悪〉	《嫌だ》
9				気持ち悪い。	(汗)		(汗)		(汗)	

※ 通番#は、本稿の説明用である。実際は巻番号・話番号等が別途存在する。

3.2.2.3 共通性の評価

1) 表情を参照する場合

まず、表情参照をする際の一時的な情緒タグの一致率を求める。まず 9 分類系の情緒タグについての結果を表 3.7 に示す。たとえば、漫画の第 1 巻と第 2 巻は、A 者と B 者が情緒推定を行ったところ一致率が 74.0%であった。本タグは、1カ所に複数の付与を認めているので、 κ 値は参考値として求めたものである。複数のタグの組を1つの種類と見なして集計したものが $\kappa_{\text{複合}}$ であり、単独のタグの付与された箇所のみを集計したものが $\kappa_{\text{単独}}$ である。この表より次のことが言える：

- 総合の一致率は 65.2%であった。総合とは、A 者、C 者、E 者側と B 者、D 者、F 者側のタグをそれぞれ束ねて比較することである。
- 推定者対ごとに一致率をみると、52.7%～74.0%であった。
- E-D 者間と E-F 者間の一致率が相対的に低い。

総合の一致率をみると、第 3.2.1 節で作成した対話コーパスの一致率は 58%であったことに比べて、本結果は高いことが分かる。しかし、E-D 者間の一致率は 58%を下回る結果となった。漫画「ちびまる子ちゃん」は易しく理解できる漫画であるため、7巻から10巻の内容が難しいというよりは、E 者による情緒の判断に問題があった可能性がある。

表 3.6 表情と情緒のタグ付きコーパスの規模

項目	規模	
冊子, 話	第 1 巻～第 10 巻, 104 話	
コマ	10,213 (コマ)	
文, 文字	29,538 (文), 388,809 (文字)	
タグ付与箇所	12,345 (のべ人・登場人物)	
表情タグ	14,040 (個)	100.0%
(内訳) 〈幸福〉	6,018 (個)	42.9%
〈嫌悪〉	2,608 (個)	18.6%
〈驚き〉	1,787 (個)	12.7%
〈悲しみ〉	1,360 (個)	9.7%
〈怒り〉	1,200 (個)	8.5%
〈恐れ〉	870 (個)	6.2%
〈背後〉	197 (個)	1.4%
情緒タグ	16,635 (個)	100.0%
(内訳) 《喜び》	4,469 (個)	26.9%
《嫌だ》	2,990 (個)	18.0%
《期待》	2,237 (個)	13.4%
《驚き》	2,010 (個)	12.1%
《恐れ》	1,757 (個)	10.6%
《悲しみ》	1,428 (個)	8.6%
《怒り》	1,347 (個)	8.1%
《なし》	207 (個)	1.2%
《好ましい》	190 (個)	1.1%

※ (タグ数は正解タグの数)

表 3.7 一時タグの比較

巻	推定者	一致率	(一致数)	κ 複合	κ 単独
1～2	A 者-B 者	74.0 %	(3,667)	0.532	0.607
3～4	C 者-B 者	71.9 %	(3,148)	0.513	0.600
5～6	C 者-D 者	68.6 %	(2,904)	0.497	0.597
7～8	E 者-D 者	52.7 %	(2,625)	0.313	0.408
9～10	E 者-F 者	60.1 %	(2,808)	0.363	0.496
全て	総合	65.2 %	(15,152)	0.444	0.546

次に、表情参照する際の一時的な情緒タグと正解の情緒タグを比較する。その前に、正解タグの信頼性を調べるために、正解タグをランダムに抽出し、同意率を調べたところ、97% (414 個の同意 / 425 個のサンプル)であった。第 3.2.1 節で作成した対話コーパスの同意率は 94%であったが、本節の結果は良好な部類である。

それでは、一時タグと正解タグの比較結果を表 3.8 に示す。この表から以下のことが分かる：

- 総合の再現率と適合率はそれぞれ 0.788 と 0.838 である。
- 推定者ごとにみると、C 者と E 者の再現率・適合率が総合よりも低い。

なお、E 者については正解タグとの一致率が低い。E 者の担当した 7 巻から 10 巻までの正解タグの同意率が 97% (158/163)であった。正解タグの不備ではなく E 者のタグ付与に問題があったと考えられる。

表 3.8 一時タグと正解タグの比較

推定者	再現率	適合率	一致率	(一致数)
A 者	0.876	0.931	90.3 %	(4,578)
B 者	0.848	0.889	86.8 %	(8,432)
C 者	0.766	0.832	79.8 %	(7,068)
D 者	0.822	0.901	86.0 %	(8,049)
E 者	0.670	0.690	68.0 %	(6,862)
F 者	0.798	0.843	82.0 %	(3,958)
総合	0.788	0.838	81.3 %	(38,947)

2) 表情を参照しない場合

表情を参照することの効果を検査するために、表情を参照せずにタグを付与し、一致率を求めた。対象は、第 2 巻、第 6 巻の各第 1 話、第 2 話の合計 4 話 (1,217 文) とした。表 3.9 に、表情参照のある場合の一致率と表情参照のない場合の一致率を比較して示す。表情参照のある場合とは、前述の結果の該当話から抽出した値である。表 3.9 より以下のことが分かる。

- 一致率は、表情参照のない場合が 60.5% であり、表情参照のある場合は 67.7% であることから、表情参照のある方が共通性が高い。

表 3.9 表情参照の有無による情緒タグの一致率の違い

表情参照	一致率	κ 複合	κ 単独
あり	67.7%	0.485	0.575
なし	60.5%	0.382	0.472

情緒推定の一致率を高めることに表情参照が有効であることが確認できた。また、表情参照の無い場合でも、第 3.2.1 節の一致率よりも高い結果が得られたのは、相対的に見て漫画の対話が長く、状況がつかみやすいこと、ならびに、漫画の登場人物の性格が推定者の間で共有できていることが理由として挙げられる。

しかしながら、漫画「ちびまる子ちゃん」は情緒推定において条件の良い対象であると容易に予想されるのだが、コーパス構築を通じて確認できたことは、一時タグの間の一致率が最高でも 74.0% であったことである。これは情緒推定の難しさを表している。

3.2.3 日記文への情緒タグの付与

第 3.2.1 節と第 3.2.2 節では、対話テキストにおける情緒推定の共通性について評価した。本節では、日記テキストにおける情緒推定の共通性について評価する。対話テキストでは対話によって作られる状態が情緒推定の手がかりとなったが、日記文では、文に表される事態が手がかりとなる。すなわち、情緒原因や情緒状態を表す表現を的確にとらえることが重要である。そこで、情緒の過程の種類に着目して、情緒推定の共通性を調査する。

3.2.3.1 設定

ここでの情緒推定は、日記文を対象に日記の作者の情緒を手で推定する。使用する日記文は、図書「英語で日記を書いてみる」[64]に収録されている文章である。日英対訳になっているが、日本語の文章を使用する。この日記は、情感でテーマ分けされた章立てとなっており、情緒の種類全体をカバーすると予想されるので、実験対象として適している。日記の作者が固定された人物で

はなく、かつ、各日記の話題が関連していないので、情緒推定者は日記を常識的に解釈しながら情緒推定を行うことになる。

日記文は、複文や重文を含んでいる。これらは、表す事態が複雑なので、節に分解する。節に分解する基準として、連用中止節、接続詞相当表現で区切られた節に着目する。引用節や連体修飾節は必ずしも分割する訳ではない。引用節や連体修飾節の表す事態は、文脈の表す時点と異なる時間や場所の事態を表すことがあるためである。以下に節単位に分割した例を示す。

- 1) 今朝スーパーで買ったサバはとても新鮮で、
- 2) 油がのっていて、
- 3) すごくおいしかった。

1) において、「スーパーでサバを買う」という事象があるが、1)から3)は、買い物をした時点とは異なるため、情緒の推定者は、どちらの時点を基準におくのか迷ってしまう。また、連体節は係り先の名詞の説明であるため、分割すると係り先の名詞の解釈が不明確になってしまう。したがって1)では連体節を分割していない。こうした配慮をしながら、日記文を節単位に分割したところ、1,642 節に分けられた。

3.2.3.2 タグ付与

情緒推定を行う者は、情緒分析の経験者 6 名と、情緒分析の初心者 1 名である。それぞれ、a 者～ f 者と g 者と呼ぶ。推定者は、日記を順に読み、節の単位で、日記の作者の情緒を読み取って9分類系の情緒タグを付与する。タグ付与は、推定者が独立に実施し、最後に集計して1つのコーパスにする。集計結果をわかりやすくするために、タグを表の形式に変換する。

具体例として、タグを集積した日記コーパスの一部を表 3.10 に示す。たとえば、#1 では、a, c, d, e, f 者の5名が《嫌だ》を付与し、g 者が《怒り》を、b 者が《なし》をそれぞれ付与している。#2 では、a 者が《悲しみ》および《嫌だ》を付与している。#3 のように、全員が同一のタグを付与することもあれば、#7 のように2種類に意見が分かれたりすることもある。

表 3.10 日記コーパスの一部(日記文は[64]より引用)

#	節	3分類系	Positive			Negative			なし		
		5分類系	P			N		A		S	なし
		9分類系	喜び	好ましい	期待	悲しみ	恐れ	嫌だ	怒り	驚き	なし
1	お父さんに床の張り替えを手伝うように言われ						acdef	g		b	
2	えーっ、そんなことでせっかくの日曜日を台無しにしたくないよーと思ったけど、				af	e	abcdeg				
3	仕方なく起きて						abcdefg				
4	手伝った。						cef	g		abd	
5	最初は面倒だと思ったけど、					g	abcdf				
6	コツがつかめてきたら		ef	g						abcd	
7	結構楽しいと思えた。	abe	cdfg								
8	今日は台所の床が終わり		a	e						bcdfg	
9	来週の日曜日は廊下をする予定。			aceg			d			bf	

情緒タグ付与の結果の概略を表 3.11 にまとめる。この表には、情緒推定の経験者である a 者～ f 者の6名についてのタグ付与結果を集計している。g 者は初心者であるため、参考値として集

計する。この結果より次のことが言える：

- 情緒の総タグ数は、6名で 11,461 (個)であった。1名・1節あたり 1.16 (個)付与したことになる。
- 情緒《なし》が 20.4%であった。情緒のあるものが8割存在するので、8つの情緒が均等に出現すると1割ずつである。そこで比較してみると、《嫌だ》は多く、《怒り》は少ないことが分かった。

表 3.11 情緒タグ付き日記コーパスの規模

項目	規模			
話の数	286 話			
文の数	1,015 文			
節の数	1,642 節			
文字数	27,357 文字			
情緒タグの数	経験者6名		(参考)初心者1名	
	11,461 (個)	100.0%	1,700 (個)	100.0%
(内訳) 《なし》	2,338 (個)	20.4%	613 (個)	36.1%
《嫌だ》	1,835 (個)	16.0%	176 (個)	10.4%
《好ましい》	1,554 (個)	13.6%	176 (個)	10.4%
《喜び》	1,220 (個)	10.6%	103 (個)	6.1%
《悲しみ》	1,188 (個)	10.4%	230 (個)	13.5%
《期待》	1,108 (個)	9.7%	129 (個)	7.6%
《恐れ》	874 (個)	7.6%	103 (個)	6.1%
《驚き》	851 (個)	7.4%	91 (個)	5.4%
《怒り》	493 (個)	4.3%	79 (個)	4.6%

3.2.3.3 共通性の評価

日記コーパスにおける情緒タグの共通性を評価する。日記コーパスには、各節に対して6名の熟練者がタグを付与したので、6名の間での一致率を計測することができる。しかし、前述の2つのコーパスと同様の比較をするために、6名から2名の組を全通り作り、2名の一致率を共通性の評価とする。

表 3.12 に、2名組の一致率を掲載する。たとえば、a 者と b 者の一致率は、56.7%である。こうして個別の組に対して一致率が得られるのであるが、全体を通じて考察を深めるために、表 3.13 に示すとおり、2名組の一致率の最小値と最大値、および、2名組の総合値をまとめる。2名組の総合値とは、a-b 者、a-c 者、a-d 者、a-e 者、a-f 者というように、ある1名と残りの5名による全ての組でタグの比較をして求めることである。

これらの表より、次のことが言える：

- 全員の総合の一致率は 52.3%である。
- 各者の総合の一致率で見ると、e 者が最も高く 54.6%、d 者が最も低い 50.0%である。
- しかし、 κ によると e 者の κ 単独は c 者や f 者に劣り、e 者の κ 複合は最低である。表 3.14 によると e 者は最も多くタグを付与している。他者と共通するタグが多いのは、曖昧な付与が多かったためだと思われる。

表 3.12 2名組のそれぞれの一致率

	a者	b者	c者	d者	e者
b者	56.7%	—	—	—	—
c者	51.9%	49.5%	—	—	—
d者	48.4%	46.0%	53.0%	—	—
e者	57.0%	51.2%	58.9%	50.9%	—
f者	48.7%	47.4%	56.2%	51.5%	54.3%

表 3.13 2名組の一致率と κ 値

推定者	一致率			κ 複合			κ 単独		
	総合	最小値	最大値	総合	最小値	最大値	総合	最小値	最大値
a者	52.6%	48.4%	57.0%	0.312	0.266	0.388	0.402	0.335	0.492
b者	50.2%	46.0%	56.7%	0.325	0.273	0.388	0.407	0.342	0.492
c者	54.0%	49.5%	58.9%	0.314	0.282	0.362	0.433	0.373	0.496
d者	50.0%	46.0%	53.0%	0.312	0.266	0.387	0.391	0.335	0.445
e者	54.6%	50.9%	58.9%	0.299	0.273	0.315	0.431	0.402	0.468
f者	51.7%	47.4%	56.2%	0.351	0.313	0.387	0.434	0.387	0.496
全員	52.3%	46.0%	58.9%	0.318	0.266	0.388	0.416	0.335	0.496

表 3.14 推定者ごとのタグ付与数

推定者	a者	b者	c者	d者	e者	f者
タグ数	1,955	1,730	2,126	1,781	2,203	1,666

次に、正解の情緒との比較を行う。日記コーパスにおける正解の情緒を多数決で決定する。多数決で決定する根拠は、情緒推定が人々に読み取られる情緒を推定することを目指していることにある。多数決による正解の情緒の決定方法は、過半数(6名中4名)の一致のある情緒タグとする。そのような情緒タグの得られない箇所は、情緒が固定的に推定できない箇所と見なして、評価の対象外とする。

表 3.15 に正解タグの規模を示す。節全体のうち、正解の情緒を決定することができた割合は、71%(1,167節/1,642節)であった。正解タグは、1つの節に対して平均1.03個であった。

表 3.15 日記コーパスにおける正解タグの規模

項目	規模
総節数	1,642 (節)
正解タグのある節数	1,167 (節)
正解タグの数	1,207 (個)

各推定者のタグと正解のタグを比較した結果を表 3.16 に示す。この表より次のことが言える：

- 総合の再現率は 0.812，適合率は 0.702，一致率は 75.3%であった。
- a者からe者は一致率がほぼ同一である。再現率と適合率はトレードオフの関係になっており、c者とe者は付与数が多く、相対的にみて再現率は高いが適合率が低い。
- f者の一致率は突出して高い。

表 3.16 日記コーパスにおける情緒タグの正解率

推定者	再現率	適合率	一致率	(一致数)
a 者	0.795	0.664	72.4%	(960)
b 者	0.754	0.733	74.3%	(910)
c 者	0.869	0.661	75.1%	(1,049)
d 者	0.770	0.721	74.5%	(929)
e 者	0.883	0.658	75.4%	(1,066)
f 者	0.798	0.811	80.4%	(963)
総合	0.812	0.702	75.3%	(5,877)

3.2.3.4 情緒推定の手がかりごとの共通性

「嬉しい」や「頭に来た」のように直接的に情緒を表す表現から情緒を推定する場合は、他の表現から推定する場合よりも、情緒推定の共通性が高いと予想される。しかし、直接的な表現がテキストにあまり出現しないのであれば、情緒推定の全体的なパフォーマンスは良くなる。そこで、日記文において、情緒推定の手がかりとなる表現の種類を分類し、各種類の表現に対する人間の情緒推定の共通性を調査する。

情緒推定で手がかりとされる表現は、「情緒生起原因を表す表現」、「情緒状態を表す表現」、「情緒反応を表す表現」である。日記コーパスにおいてこれらが出現する頻度と共通性について分析する。

日記コーパスの各節を、上記の3種類の表現あるいはその他の表現に分類し、出現頻度を求めたところ表 3.17 のとおりとなった。1つの節に、原因と反応の両方が見られる場合は両方ともに分類したため、合計値は、総節数の 1,642 を超える。この結果より、情緒生起原因を表す節が多く存在することが確認できた。なお、その他には「未然の事態」を表す表現が多く分類された。未然の事態とは、「これが夢だったらいいのに。」や「彼女は大変だと思うかもしれないが、」のように想像上のことを述べる節である。

表 3.17 日記コーパスにおける表現の種類の内訳

表現の種類	割合	(該当する節の数)
情緒生起原因を表す節	68.3%	(1,122)
情緒状態を表す節	15.7%	(258)
情緒反応を表す節	5.1%	(84)
その他	11.4%	(187)

※ 総節数 1,642 に対する割合

次に、共通性を調べる。表 3.18 に、各表現における一致率、一致率の最大値と最小値、 κ 複合、および、 κ 単独をまとめる。この表より次のことが言える：

- 情緒状態を表す節における情緒推定は共通性が高い。基本的な情緒を表す節は、単独の情緒タグが付与されやすく、かつ、そのタグが一致しやすいと思われる。 κ 単独が高いのはそのためであろう。
- 残りの節(原因, 反応, その他)における一致率はほぼ同水準である。

表 3.18 日記コーパスにおける表現の種類ごとの2名組の一致率と κ 値

表現の種類	一致率	(最大/最小)	κ 複合	κ 単独
情緒生起原因を表す節	49.7%	55.7% / 41.9%	0.296	0.372
情緒状態を表す節	63.1%	71.9% / 53.6%	0.371	0.575
情緒反応を表す節	51.8%	61.6% / 43.1%	0.271	0.388
その他	51.6%	58.0% / 44.8%	0.292	0.400

表 3.19 に多数決 (6名中4名以上の得票タグ) で正解タグを決定した場合の正解の割合を示す。正解存在率は、多数決で正解タグの決定できた節の割合を表す。この表より以下のことが言える:

- 情緒推定の正解率に著しい開きは見られない。
- 正解存在率は、情緒状態を表す場合において 90%と高く、残りの3つの場合において 70%前後である。

正解率に開きがないのは、情緒推定の経験者による多数決で正解タグが決定されたためである。多数決では、6名中4名以上のタグの一致という条件があり、一致しやすい節が抽出されたと考えることができる。しかし、注目すべきは、正解存在率である。推定者の間で共通に読み取られる情緒が存在する節の割合を、正解存在率が表している。情緒生起原因を表す節は、情緒状態を表す節には及ばないものの、他と同等の割合で存在している。

表 3.19 日記コーパスにおける表現の種類ごとの正解率

表現の種類	再現率	適合率	一致率	(一致数)	正解存在率
情緒生起原因を表す節	0.799	0.703	74.8%	(3,692)	70% (748 / 1,122)
情緒状態を表す節	0.848	0.707	77.1%	(1,241)	90% (232 / 258)
情緒反応を表す節	0.842	0.660	74.0%	(293)	68% (57 / 84)
その他	0.804	0.697	74.7%	(685)	73% (137 / 187)

以上の結果より次のことが言える。

- 情緒状態を表す節は、正解存在率が 90%であることより共通的に情緒の推定できる可能性が高いといえる。しかし、日記コーパスにおいて存在した割合が 15.7%であったことより、この種類の表現を手がかりとする情緒推定の方法は、実際にドキュメントを解析する際、網羅性の点で不安がある。
- 情緒生起原因を表す節は、正解存在率が 70%であることより共通的に情緒の推定できる可能性があり、かつ、日記コーパスにおいて存在した割合が 68.3%であったことより、この種類の表現を手がかりとする情緒推定の方法は、網羅性を確保するために必要である。

3.3 まとめ

人間による情緒推定の共通性を、3つのコーパスの構築を通じて調査した。コーパス1では、対話テキストに対して情緒推定過程を明示するための情緒タグを付与した。情緒生起原因を表すタグを使うことが情緒推定過程の分析の助けになることが確認された。対話での意図解析の違いや、情緒主の性格の読み取り方の違いが、情緒推定の不一致につながることを確認された。コーパス2では、表情の参照可能な対話テキストとして、漫画を題材として構築した。表情タグおよび情緒タグの付与を行った。表情の参照の有無による情緒推定の共通性の違いが定量的に明らかになった。

コーパス3では、日記を題材として、情緒タグを付与した。特に、情緒的過程に基づき文を分類し、推定される情緒の共通性を調査した。テキスト対話は日記文よりも情緒推定の共通性が高い傾向が見られた(表 3.4, 表 3.7, 表 3.13)。

本研究では、情緒生起原因に注目するので、日記文において、「情緒生起原因」、「情緒状態」、「情緒反応」、および、「その他」の4種類に節を分類し、出現頻度とともに情緒推定の共通性を調べた(表 3.17, 表 3.18)。情緒状態を表す節における一致率は 63.1%と高いものの、日記文においては出現割合が 15.7%であった。一方、情緒生起原因を表す節における一致率は 49.7%と相対的に低いものの、出現割合が 68.3%と高かった。テキスト全体から情緒を推定する際、推定できる機会を逃さないために、「情緒生起原因」を表す表現を対象にする必要性が確認された。

第4章 パターン言語処理の基礎と実現

本章では、情緒生起原因に基づくタイプ1の情緒推定を行う方式の基礎を示すとともに、その実現に不可欠な言語知識ベースを設計する。本方式に対する要請は、言語表現の表層レベルの情報から情緒が推定できることである。そこで、まず、三浦の言語過程説を参照して、言語の表現と意味についての捉え方を確認する。次に、意味処理を、「非線形言語モデル」に基づいて実現する方法「パターン言語処理」を示す。その方法で使用する言語知識ベースは、「文型パターン」の形式である。本章の後半では、文型パターンの記述と運用を効果的に行うことについて示す。

4.1 言語表現の生成と理解の過程

言語知識ベースは、言語表現を意味処理するために用いる知識ベースである。言語表現と意味処理についてここで確認しておく。

言語表現とは、単語、句、節、文、文章という文字列や音素の列による表現である。こうした表現に対して、「意味がある」や「意味が分かる」ということがある。「意味」とは何かについては様々な議論がなされているが、三浦の言語過程説[65]においては、次のように述べられている：

「音声や文字には、その背後に存在した対象から認識への複雑な過程的構造が関係づけられているわけで、このようにして音声や文字の種類にむすびつき固定された客観的な関係を、言語の「意味」と呼んでいるのです。」

この説明において、「対象」、「認識」、「(言語)表現」、および、「話し手」という4つの関係が重要である。この4つについて注意しながら、言語表現の生成過程と理解過程を概観しよう。

まず、言語表現の生成過程を概観する。言語表現は話し手により生成される。話し手は、ある対象を視覚や聴覚などの知覚器を通じて認識する。このとき話し手は知覚的特徴をとらえて、話し手の持つ認識の空間に展開する。話し手は、対象に対する特徴をよく表す言語表現を選択して言語表現を生成する。たとえとして、対象が話し手の外界の物事である場合で説明したが、対象が話し手の心の中(想像)にある場合も同様である。

ここで、重要な点は、「言語表現の対応する特徴は、社会的な約束で決まっている」という点である。三浦の言語過程説でいうところの「固定された客観的な」というのは「社会的な約束」である。本研究では、「社会的な約束として言語表現の対応する特徴」のことを「言語表現の概念」と呼ぶことにする。

たとえば、図 4.1 において、話し手は、レーシングカーのタイヤ、シャーシ、エアロパーツなど知覚的特徴をとらえて認識している。こうした特徴を総合して言い表す単語が「レーシングカー」であり、言語表現に使われている。

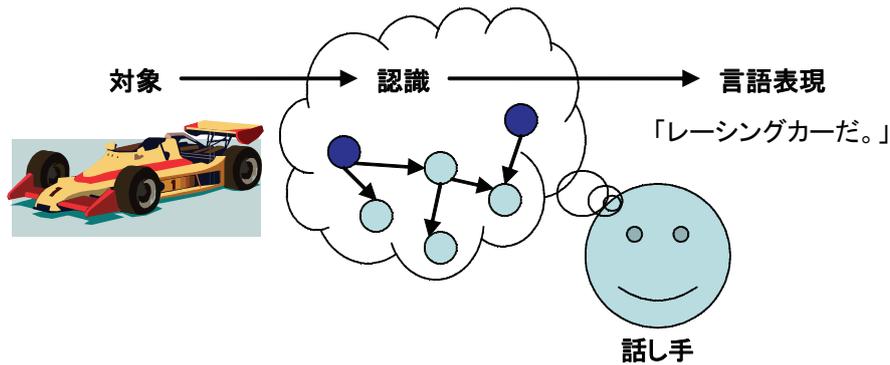


図 4.1 言語表現の生成過程

次に、言語表現の理解過程を概観する。話し手の言語表現を聞き手が受け取ると、聞き手は言語表現の概念を聞き手の認識の空間に展開する(図 4.2)。言語表現を文字通りに解釈することに相当する。言語過程説では、言語表現の文字通りの解釈までが議論の中心である。しかし、通常、言語表現の文字通りの解釈だけで聞き手の認識がとまることはなく、文脈に応じた解釈を試みたり、聞き手自身の既知の事柄と関連性を推論したり、あるいは、具体的なイメージをたてる。このように「追加的な推論」を行うことがある。これは語用論では推移と呼ばれる[49]。

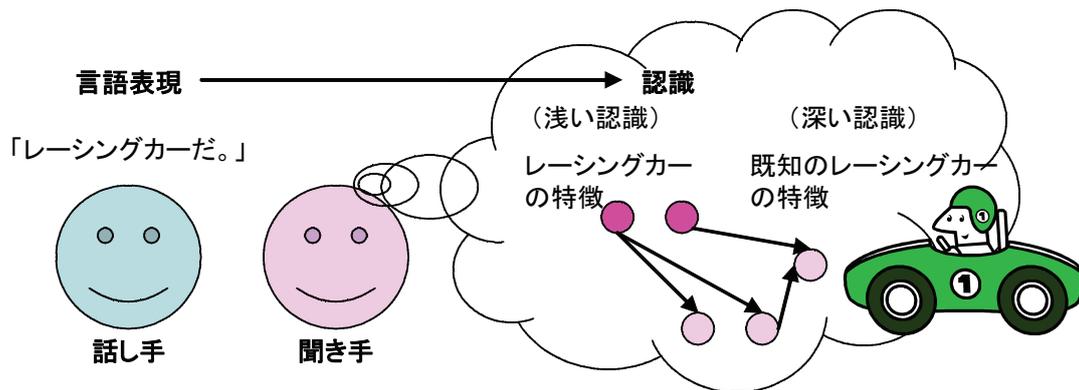


図 4.2 言語表現の理解過程

以上より、本章では用語を次のように定義して用いる。

- **話し手**：言語表現を生成する人
- **聞き手**：言語表現を理解する人
- **対象**：物理的・心理的に存在する物事・属性や関係
- **認識**：特徴や特徴間の関係・結びつき
- **言語表現**：単語、句、節、文、文章を形成する文字列や音素列
- **言語表現の概念**：社会的な約束として言語表現に対応する特徴
- **「対象を認識する」**：対象の特徴を抽出し認識に取り込むこと
- **「言語表現を生成する」**：言語表現を作ること。ただし、言語表現の概念が対象の特徴を話し手が伝えたいと思う程度によく表さなければならない。
- **「言語表現を理解する」**：言語表現の概念を、認識に取り込むこと

ここで、「特徴」という用語に定義が依存している。

- **特徴:** ある対象を、他の対象と区別するための根拠

特徴の定義は抽象的である。それは、定義が対象に依存しているためである。たとえば、対象が物であるならば物理的に形や大きさなどが区別するための根拠として挙げられるし、事であるならば時間変化などが挙げられる。

なお、図 4.1 および図 4.2 のポンチ絵では、理解を助けるために、認識をネットワークに模して描いている。しかし、特徴や特徴の関係がネットワークという形式であるということを、本研究で主張するものではない。生成や理解の過程において、様々な特徴が使用可能な状態にすることを「認識に取り込む」と呼ぶことにする。

4.2 意味処理への接近

4.2.1 意味解析と意味理解

池原は計算機による言語の理解過程を「意味処理」と呼び、「意味解析」と「意味理解」の2つに分類した[66]。意味解析は、前述の理解過程における浅い認識までの処理である。言語表現は、1つの表現でも複数の語義が対応することから、周辺の表現との関係から話し手がどの語義で表現を使用したのかを解析する必要がある。その他、意味理解との境界が不明確ではあるが、構文解析・格解析も関わっている。一方、意味理解は、前述の理解過程における深い認識の処理を指し、照応・省略解析、談話構造解析を含む。

4.2.2 言語表現の概念の分類

意味解析は、社会的約束に従って、言語表現の概念を捉える処理である。池原によると、言語表現の概念は、言語表現の複雑さに基づき、次の2つに分類される[10]。

- **単一概念:** 「単語」が表す概念
- **複合概念:** 複数の単語による表現が表す概念

1つの単語が(多義はあるにせよ)概念を表すことは改めて主張すべきことではない。しかし、池原の捉え方の特徴は、句、節、あるいは、文が1つの概念を表すと捉える点である。この捉え方は、言語表現を構成する単語を1つでも変更すると、その表現の概念は別ものになるという。たとえば、「父はタバコを吸う。」という文の概念と、「パパはタバコを吸う。」という文の概念は、完全には一致しない。たとえば「話し手の幼さ」という点が異なることが伺える。

文の意味解析を行う上で、文は複数の単語で構成されるので、「文の概念」は「複合概念」である。単語の概念は、数十万のオーダーで単語辞書に蓄積可能であっても、文は単語の組み合わせで作られるため、辞書として蓄積することは、現状では可能性が極めて低い。

そこで、考案されたのが「表現構造の概念」である[10]。表現構造とは言語表現の組み合わせ方を指す。たとえば、「猫はネズミより強い。」という表現には、「～は～より～」という表現構造がある。表現構造の概念とは、表現構造に対する人々の間での共通の解釈である。すなわち、この例の場合、「2者の比較」という解釈が可能である。そこで次の種類の概念を導入する。

- **抽象的概念:** 表現構造が表す概念

文を構成する単語を変更すると複合概念が変化するように、表現構造を構成する部分表現を変更すると、やはり、抽象的概念は変化する。たとえば、「猫もネズミより強い。」という文において、「～も～より～」という表現構造が存在する。ここで、「猫」と「ネズミ」の「強さ」を比較するので「2者の比

較]をしているが、「～も」という表現からは、「ネズミ」と「強さ」を比較した対象が「猫」以外にも文脈上存在していたことが伺える。したがって、「～も～より～」と「～は～より～」とは抽象的概念が異なる。

4.2.3 非線形言語モデル

言語表現の意味解析を行う上で、複合概念をありのまま扱うことを避けるために、抽象的概念を扱うことが1つの解決策である。そこで、提案されたのが「非線形言語モデル」である[10]。非線形言語モデルは、「言語表現が非線形要素と線形要素で構成される」ととらえる言語モデルである。ここで、非線形要素と線形要素の定義は次のとおりである：

- **非線形要素**：表現構造を構成する部分表現
- **線形要素**：非線形要素の表す抽象的概念が変化しない範囲で、書き換え可能な部分表現

※ 部分表現とは対象とする言語表現を構成する部分的な言語表現

たとえば、図 4.3 をみてみよう。「昨日買った鯖はいつもの鯖よりおいしい。」という文を非線形言語モデルに基づき解析する際、抽象的概念として「2者の比較」とすると、非線形要素として「～は～より～」という部分表現と、線形要素として「昨日買った鯖」、「いつもの鯖」、「おいしい」という3つの部分表現とに分けられる。

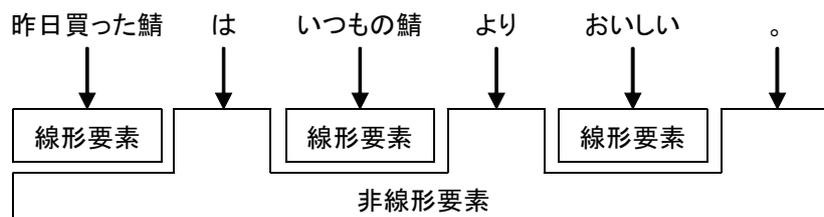


図 4.3 非線形言語モデルに基づく言語表現の区別

非線形言語モデルの性質を以下に説明する。

非線形要素は、表現構造を構成するために、文字列の上で離れた単語の間でも連携している。さらに、非線形要素の周辺の線形要素に対して制約を与えている。その制約は、抽象的概念が変化しないようにするために存在する。たとえば、「川は山より流れる。」は、「～は～より～」という形式が内部に存在するが「2者の比較」ではない。このことから分かるように、線形要素に対する統語的な制約が必要である。他にも「燃費は自動車より優れている。」^{*13} という表現においては、比較の観点が最初の線形要素に示されることから「2者の比較」ではない。この場合、線形要素に対する意味的な制約が必要である。

線形要素は、「おいしい」というように単語が該当することがあれば「いつもの鯖」というように修飾句を伴う表現や、「昨日買った鯖」というように連体節を伴う表現もある。線形要素は、元の言語表現全体よりは単純になるため、より単純な複合概念、または、単一概念を表す。

線形要素として取り出した部分表現は、別の表現構造を見出すことにより、さらに、線形要素と

*13 たとえば、文脈としてバイクの話をしている場面を想像しよう。

非線形要素の区別がなされる。たとえば、「昨日買った鯖」という表現は、「～買った～」という非線形要素と「購入物」という抽象的概念で解析すると、「昨日」と「鯖」が線形要素として抽出される。なお、「昨日」と「鯖」は単語であるので、これ以上分解することはできないので、「非線形要素」として扱われる。このように、非線形言語モデルは、線形要素に対して再帰的に適用することができる。

4.2.4 意味解析

4.2.4.1 非線形言語モデルに基づく意味解析

言語表現を非線形言語モデルで解析するという操作は、次の2つのことを行うことである。

- (1) 構成要素の分類：言語表現を構成する部分表現に対して、線形要素であるか非線形要素であるかの分類を行うこと
- (2) 複合概念の分離：複合概念から、抽象的概念を抽出し、より小さい複合概念、または、単一概念を抽出すること

意味解析のために、処理対象となる言語表現が与えられると(1)の操作を実行する。その結果、非線形要素として表現構造が判明し、残りの部分から線形要素が得られる。次に、(2)の操作として、表現構造に対応する抽象的概念が抽出され、一方で線形要素の概念が抽出される。

これら両操作は、理想的にいえば、表裏一体である。構成要素の分類ができることと、複合概念の分離ができることは、同時に成立する。しかし、現実的には、非線形要素が線形要素に対して、完全に正しく制約をかけることができる保証がないため、一見、構成要素の分類ができていても、複合概念の分離が不正確である場合があったり、得られた概念の曖昧性が解消されない場合があったりする。

4.2.4.2 文型パターン

上述の「(1)構成要素の分類」および「(2)複合概念の分離」という操作を行うための知識ベースの要件として次の項目が挙げられる：

- 表現構造の特徴を抽出すること：表現構造の特徴とは、字面の並び、および、線形要素との位置関係である。たとえば、「2者の比較」を表す「～は～より～」という表現構造において、「は」と「より」は重要なキーワードである。
- 線形要素へ制約条件を課すこと：線形要素に対する制約には、品詞や活用形という統語的な制約、および、意味属性などによる意味的な制約がある。たとえば、「2者の比較」において、述語となる線形要素は、形容詞あるいは形容動詞でなければならないという統語的制約がある。また、「～は」と「～より」が互いに比較される名詞でなければならない、そのような名詞の範囲は、名詞の意味属性で制限することが考えられる。

字面、統語的制約、意味的制約という表現上の特徴を記述する形式として、文型パターンがある。文型パターンは、従来の言語処理研究において使用されていた。たとえば、日本語語彙大系では、日本語の基本的な用言 6,000 語に対して日英翻訳における語義解析のために、約 14,800 パターンを掲載している。このパターンは上記の項目に関して、変数、字面、記号、および、意味属性制約を用いて記述している。変数は名詞を表す N が使用されている。字面は、助詞(相当表現を含む)および用言(終止形)を表す。記号は、助詞の選択を表す。意味属性制約は、変数に適合可能な意味的範囲を表す。このように、文型パターンは、上述の2つの要件を満たす。

ここで、従来の文型パターンの記述言語仕様は単純なものであった。そこで後の第 4.3 節では、効果的に文型パターンを運用するために、記述言語の仕様を改良し、その効果を確認する。

4.2.4.3 文型パターンの操作

解析対象の言語表現に対する文型パターンの操作について説明する。

- (1) 照合：文型パターンと言語表現とを「**照合**」するという操作がある。照合は、文型パターンに指定された記述要素が順序通りに言語表現の構成要素に対応することを確認する操作である。照合により、対応関係が確認できたことを「**適合**」と言い、確認できなかったことを「**不適合**」と言う。

文型パターンは、1つのパターンに1つの抽象的概念が対応している。言語表現の種類は様々であるので、大量の文型パターンが作られる。そこで、関連情報を伴う文型パターンを大量に収録した言語知識ベースを「**文型パターン辞書**」と呼ぶ。

- (2) 検索：文型パターン辞書から、解析対象の言語表現に適合する文型パターンを全て抽出し、変数などに対応する情報を全て列挙することを「**検索**」と言う。検索の結果得られる文型パターンのことを「**適合パターン**」という。そして、パターンや変数などの対応関係の情報1件ぶんのことを「**適合ユニット**」と呼ぶ。

4.2.4.4 抽象的概念の取り扱い

言語表現に文型パターンが適合することにより、適合パターンの持つ抽象的概念がその言語表現に存在することが解析される。抽象的概念を、たとえば、「2者の比較」というラベルで出力することは、抽象的概念を扱う一つの手段である。しかし、抽象的概念のラベル化は一般には容易なことではない。パターンの適合をラベルで表すのではなく、適合したことを、抽象的概念を使用した処理を起動するトリガーにするほうが適切な意味処理を行い易い。たとえば、日本語語彙大系の日英パターン対は、日本語文型パターンの適合に対して、英語文型パターンが定まる構造である。

意味解析という処理は、言語表現の直接的な解釈を行うことを指すのだが、パターン言語処理においては、適合パターンの決まることの意味解析に相当する。パターンの適合がトリガーとなり新しい情報の解析が行われると、意味解析の段階を越えて、意味理解の段階に入ると言える。

4.2.5 意味理解

4.2.5.1 追加的推論の戦略

文型パターンを用いて意味解析ができた後、意味理解を行う。意味理解では、言語表現の表す概念に関して追加的な推論を行う。その追加的な推論とは、言語表現の意味解析から得られた情報(新情報)と、既存の情報(たとえば、文脈情報や長期記憶的な情報)を組み合わせで行う推論である。ここで、推論には、大きく2つの戦略がある：

- 前向き連鎖：
新情報から連鎖的に推論を進めて、推論のゴールが成立することを示す(図 4.4(a))。
- 後ろ向き連鎖：
新情報から推論のゴールに対する仮説をたてて、その仮説を証明する。仮説の証明には、仮説の成立に必要な事柄がサブゴールとなり、それらの証明を行う(図 4.4(b))。

前向き連鎖では、言語表現からの意味解析の結果(図 4.4(a)の新情報)を出発点として推論を進める。すると、ゴール以外の事柄が多く含まれた、いわゆる、発散した状態になる。一方、後ろ向き連鎖では、意味解析結果(図 4.4(b)の新情報)から仮説が立てられ、仮説が成り立つために必要な部分を文脈情報から確認するという方法が成り立つ。ただし、この確認が直ぐにできない場合、すなわち、さらに多くの仮説の証明が要求される場合には、両戦略に違いが無い[67]。

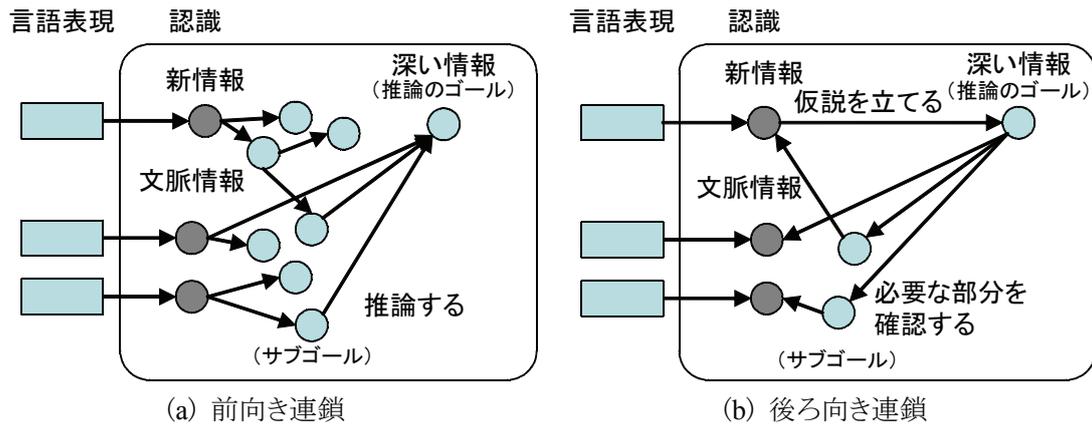


図 4.4 意味理解における追加的推論の戦略

4.2.5.2 意味理解としての情緒推定

言語表現が情緒状態を明示しない際の情緒推定は全て意味理解としての情緒推定である。ここでは、情緒推定方法のタイプ(第 2.2.3 節参照)ごとに戦略を考えてみる。

タイプ1とタイプ2においては、言語表現の表す事態が情緒の過程と言えるかどうか問われる。この場合、後ろ向き連鎖が実現しやすいと考えられる。すなわち、文型パターンの表す概念に対して、情緒の過程が対応付けられていれば、文型パターンが入力文と適合することによって、情緒の仮説を立てることができる。線形要素との組み合わせに依存して、その仮説が棄却されることが必要であるため、後ろ向き連鎖によりその判断をすればよい。ここで、後ろ向き連鎖の際に立てられるサブゴールは、「浅い認識の事柄」、「文脈情報の事柄」、「常識的な事柄」の3つのうちのいずれかであることが予想される。「浅い認識の事柄」、「文脈情報の事柄」として扱うことができない場合に「常識的な事柄」として対処することが考えられ、また、常識的な事柄については、Liu らのように統計的な裏付けを用いることで対処することが可能であるだろう。

タイプ3～5においては、両方の戦略が必要である。Rosis らの方法は、談話プランに情緒の情報を対応付けておき、プランの前提条件に仮説を裏付ける条件を記述した[54]。この方法は後ろ向き連鎖の方法である。一方、前向き連鎖の方法の例として、第 3.2.1 節で示したコーパスの構築が挙げられる。対話文から得られる情報に、推定者が推論を加えて、非情緒的な心的状態のタグを付与し、情緒を判定する条件の成立をもって、情緒タグを付与した。

本研究は、原因に基づくタイプ1の情緒推定を実現するものであるため、後ろ向き連鎖による情緒推定を採用することにする。その過程を以下にまとめる：

- (1) 処理対象となる言語表現の表層レベルの解析を行う。
- (2) 知識ベース(情緒推定用の文型パターン辞書)から、適合する文型パターンを検索する。
- (3) 適合パターンに対応付けられている情緒の情報を参照する。ここで、情緒の種類(喜び、悲しみ、怒りなど)、情緒主、情緒対象などが、仮説として立てられる。
- (4) 情緒主と情緒対象の関係を検査して、仮説が成立するかどうかを判定する。

ここで、(4)については、Liu らが示したように、格要素の関係に対する統計値により判定できるので、基本的に解決可能と思われる部分である。問題なのは(2)の知識ベースである。たとえば、「N1 が N2 を貰う」というパターンは、所有的移動という概念を表すが、情緒主が N1、情緒対象が N2、そして、情緒名は「喜び」という情緒の情報をこのパターンに対応付けておかなければなら

ない。このようなパターンの網羅的な作成が問題となる。

4.2.6 意味処理のための枠組み

以上の考察により、意味処理のための知識ベースは次の2つの情報が必須である：

- (a) 言語表現の構成要素を分類するための文型パターン
- (b) 応用のための情報

(a)は、言語表現を意味解析するために必須である。(b)は、意味理解において、前向き連鎖で推論するための直接的な連鎖を示すこと、あるいは、後ろ向き連鎖のための仮説とその補足条件を示すことのために必須である。(b)については、第5章において具体的に示すことにするので、本章では(a)について詳しく述べる。

(a)の文型パターンに対して次の3つの要件がある：

- (i) 文型パターン辞書から、解析対象の言語表現に適合する文型パターンを検索できること
- (ii) 文型パターンが、解析対象の言語表現に柔軟に適合すること
- (iii) 文型パターンが、意味的に正しく適合すること

(i) について、第 4.3 節で方法を示し、第 4.4 節でその実装が妥当な動作特性であることを示す。(ii) について、第 4.3.1 節において、柔軟な適合ができるように文型パターンの記述子を設計し、第 4.5 節でその効果を実験的に評価する。(iii) について、第 4.3.1 節において、意味制約付き変数を導入し、第 4.6 節において、意味属性制約の効果について評価する。

4.3 文型パターンの記述子と照合

4.3.1 記述子

本節では、文型パターンの記述子を示す。

既に述べた通り、文型パターンは、言語表現の構成要素において、非線形要素と線形要素を分類するための知識の形式である。表現構造を構成するために「字面」という記述子を使用され、線形要素の該当位置と統語的制約を示すために「変数」という記述子を使用される。しかし、字面と変数だけで記述された文型パターンは、柔軟性に欠けるため、多様な言語表現に対処できない。また、2つの単純な記述子だけでは統語的な制約が記述し難い。さらに、文型パターンの適合結果を有効に活用するために機能的な記述子が必要である。そこで、記述子の拡張を行う。それではまず要請を確認しよう。

言語表現にパターンが柔軟に適合するために、言語表現の次の点に着目する。

- 表記性：漢字、平仮名、片仮名、音便、活用、異体同義の表現についての柔軟性
- 語順性：格要素の順序、副詞の位置、主題の位置についての柔軟性
- 任意性：省略、挿入に対する柔軟性

言語表現にパターンが厳密に適合するために、言語表現の次の点に着目する。

- 語義：単語や句に対する意味的な条件
- 語形：活用形に対する条件
- 文節境界：単語と単語の間が文節の境界か否か

文型パターンの機能を高めるために、次の点に着目する。

- まとめ機能：複数の記述子をまとめる機能
- 記憶機能：パターンの記述子に対して、言語表現が対応した結果を記憶する機能

以上の要請に対して、文型パターンの記述子を4種類に大別し、それぞれに、具体的な記述子を設ける。

- (a) 字面：出現形字面，基本形字面
- (b) 変数：意味自由変数，意味制約付き変数
- (c) 関数：ドット関数，ハット関数
- (d) 記号：選択記号，順序任意記号，移動可能記号，任意記号，補完記号，離散記号，文節境界記号，非文節境界記号，まとめ記号，記憶記号

「出現形字面」は、適合する字面を厳密に指定する。たとえば、パターンと言語表現の間で、漢字と平仮名の表記が異なると、不適合となる。一方、「基本形字面」は、単語の基本形をシングルコートで囲みパターンを記述するのだが、言語表現の基本形に基づき適合を判定する。たとえば、出現形字面によるパターン「もなか」は言語表現「もなか」と適合するが「モナカ」には適合しない。しかし、基本形字面によるパターン「モナカ」は言語表現「モナカ」にももちろん適合するが、「もなか」や「最中」とも適合する^{*14}。基本形に基づく判定なので、パターン「走る」に対して言語表現「走り」や「走ら」も適合する^{*15}。こうして、字面における表記の柔軟性が確保される。

「意味自由変数」は、「変数名」と「変数番号」の組で記述する。変数名により適合する品詞(名詞句や連体節など単語以外の統語的構成要素を含む)を指定し、変数番号により適合する言語表現を記憶する。一方、「意味制約付き変数」は、意味自由変数に「意味属性制約」を追加した記述形式とする。変数に適合した言語表現が意味属性制約を充足することで、意味属性制約付き変数の適合が決まる。

「ドット関数」は、先行するパターンの記述要素に関連して、後続する表現を指定する。たとえば、用言に後続する表現を指定する。具体的には、パターン「V1.kako」は、変数「V1」とドット関数「.kako」で構成されており、表現「走った」や「飲んだ」の「た」や「だ」が「.kako」と適合する。こうして、助詞や助動詞に見られる音便の表記の揺れを吸収することができる。また、「本だ」，「本です」，「本である」のように判定詞の違いが、たとえばドット関数「.da」で吸収することができる。

「ハット関数」は、先行するパターンの記述要素に対する表現上の制約および統語的制約を指定する。たとえば、パターン「V1^rentai」は、変数「V1」が連体形であることを制約している。

「選択記号」は、「|」を用いて、言語表現と対応すべきパターン記述要素列を選言的に指定する。たとえば、異体同義の表現を効率的に吸収する。具体的には、パターン「N1(から|より)」は表現「山から」と「山より」の両方とも適合する。他に、パターン「山は(AJ1|AJV1)」は、述語が形容詞でも形容動詞でもよいことを示している。

「順序任意記号」は、複数のパターン記述要素が順序を問わず言語表現と対応することを指定する。記述形式は、中括弧内に複数の記述要素列をコンマ区切りで列挙するという形式である。たとえば、パターン「{N1は，N2に，N3を}渡す」は、表現「彼は彼女に手紙を渡す」と適合することはもちろんのこと、表現「手紙を彼女に彼は渡す」も適合する。こうして、格要素の順序に柔軟性を持たせることができる。

「移動可能記号」は、指定するパターン記述要素が、パターン上で離れた幾つかの位置のい

*14 ただし、「もなか」と「最中」の形態素解析結果に、基本形として「モナカ」が出力されること。

*15 ただし、「走り」、「走ら」の形態素解析結果が、基本形として「走る」が出力されること。

れかで、ちょうど1回対応することを指定する。記述形式について、「\$数[^]{パターン記述}」により移動可能な要素を定義し、「\$数」により可能な移動先を定義する。たとえば、パターン「\$1N1 は \$1[^]{ADV2}N3 を\$1V4.kako」は、表現「突然太郎はドアを開けた」、「太郎は突然ドアを開けた」、「太郎はドアを突然開けた」のいずれも適合する。こうして、記述子の位置の柔軟性が高まる。

「任意記号」と「補完記号」は、パターン記述要素のうち、言語表現と対応することが必須ではないことを指定する。記述形式は、大括弧あるいはカギ括弧「< >」で任意となる要素を囲む形式である。たとえば、パターン「<N1 は>V2」は表現「彼は行く」、「行く」の両方とも適合する。また、パターン「V1 [.kako]」のように使うことで、表現「行く」、「行った」の両方とも適合させることができる。このように、補完主語や時制の任意化が可能である。

「離散記号」は、言語表現のうち、文型パターンで解釈しない部分を指定する。記述形式は、「/ 離散記号名」である。〈離散記号名〉には「半角英語小文字列」を使用し、解釈しない言語表現の統語的種類を指定する。たとえば、パターン「/tN1」は、連体節の離散記号「/t」を含んでいるので、表現「昨日スーパーで買った鯖」は、下線部が離散記号に対応しながらパターンと適合する。なお、離散記号は挿入の可能性を指定するものであるので、たとえば、表現「鯖」のように、連体節が含まれていなくてもパターン「/tN1」と適合する。こうして、文型パターンの中に挿入可能な表現が指定できる。

「文節境界記号」は、単語と単語の間に文節境界のあることを指定する。記述形式は「!」である。同様に、「非文節境界記号」は、単語と単語の間に文節境界の無いことを指定する。記述形式は「+」である。たとえば、パターン「N1+は!V2」は、表現「彼は帰る」と適合する。「N1」と「彼」が対応しているが、「彼」と「は」の間に文節境界が無いので、「+」が対応する。「は」と「帰る」の間に文節境界があるので、「!」が対応する。文型パターンと言語表現を照合する際、言語表現の形態素情報を利用するため、連続する変数、節の境界、離散記号の定義などにおいて、統語的に正しく照合するためにこれらの記号が使用される。

「まとめ記号」は、小括弧を用いて、1個以上のパターンの記述子をまとめて1つの記述要素とみなすことを指定する。たとえば、パターン「N1 は|が V2」は、選択される記述要素列が「N1 は」と「が V2」になってしまう。助詞の選択とするならば、「N1 (は|が) V2」としなければならない。また、パターン「N1 は(ND2 をする|V3)」は、表現「彼は勉強をする」、「彼は勉強する」の両方とも適合する。パターン上で述部が明示できる。

「記憶記号」は、対応づけられた文型パターン記述子を記憶する。記述形式は、「#記憶番号」を記憶したいパターン記述要素の直前に配置する。たとえば、パターン「N1#1 (は|が) V2#2 [.kako|.genzai|.mirai]」は、表現「彼は帰った。」と適合するが、「#1=[は], #2=[.kako]」という結果が得られることから、格の対応、時制の対応が判定できる。

4.3.2 照合

4.3.2.1 照合の基本方針

非線形言語モデルに基づく意味解析では、言語表現の構成要素に対して線形要素と非線形要素の分類を行うために、文型パターンを使用する。1つのパターンが1つの表現構造を表すので、日本語の表現構造を網羅するためには、かなりの数のパターンを用意しなければならない。

意味解析においては、パターン辞書の中から解析対象となる言語表現に適合する全ての文型

パターンを検索する。さらに、検索結果として、線形要素と非線形要素の分類結果などの適合ユニットを全て出力する。

ただし、線形要素の内部構造について、解析する必要はなく、対応する表現の範囲が同一であるならば内部構造の相違は区別しない。たとえば、表現「赤い家の扉」は名詞句であり、内部構造は「赤い」の係り先が「家」である場合と「扉」である場合とに分けられるのだが、名詞変数「NP」との照合において区別しない。

以上より、照合の基本方針をまとめる：

- 原則として、1つの言語表現に対してパターン辞書内の全パターンと照合する。
- 適合した全パターンを出力する。
- 全ての適合の仕方を出力する。
- パターン記述子に対応した言語表現の内部構造についての曖昧性は不問とする。

4.3.2.2 照合の方法

本研究では、文型パターンごとにネットワークを作成して、幅優先 ATN により適合する文型パターンを検索するという方法をとる。

ATN (Augmented Transition Network) は、トップダウン構文解析に用いられる手法の一つである[68]。句構造規則に対応する状態遷移ネットワークを使用し、解析対処となる言語表現の構成要素を入力として、ネットワーク上の状態遷移により解析を進める。非終端記号に対して別のネットワークを対応させるので、ネットワークから別のネットワークへと状態遷移の場を移すことがある。

たとえば、文型パターン「N1 は N2 より (AJ3 | AJV3)。」は図 4.5 のようにネットワーク化される。start の指しているノードから状態遷移を開始する。N は名詞変数、AJ は形容詞変数、AJV は形容動詞変数である。AJ と AJV に選択記号が係っているため、ネットワークでは平行のアーキが次ノードに接続している。pop のアーキに辿り着くと照合が終了となる。

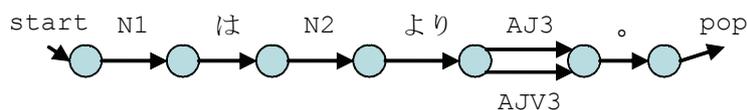


図 4.5 ネットワーク化された文型パターン

名詞変数 N は、実際には接辞が付随することが認められていたり、形容詞に接尾辞が加わり名詞化する場合も認めていたりするため、複雑な定義となる。図 4.6 は名詞変数 N の定義例である。接頭辞 PF，純然たる名詞 NE，接尾辞 SF，形容詞 AJ，名詞化する接尾辞 SFN により名詞が定義されている。文型パターンにおいて、変数 N の照合をするために、この名詞変数のネットワークに状態が遷移する。

ATN の基本的な動作であるが、図 4.6 のアーキのうち始点と終点が同一のノードであるアーキがある。これは、0回以上の繰り返しを表す。また、pop アーキを通過すると、図 4.5 のネットワークに戻るのだが、N1 の照合のために N のネットワークに遷移していた場合は、N1 のアーキの終点に遷移し、N2 のために N のネットワークに遷移していた場合は、N2 のアーキの終点に遷移する。

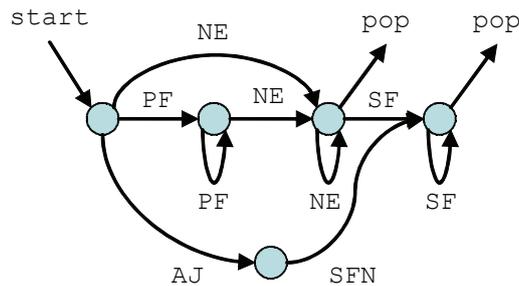


図 4.6 名詞変数 N のネットワーク

1) ATN を用いる理由

ATN のメリットは、ネットワークの状態遷移の条件に、処理を定義することができる点である。文型パターンの記述子には、既に述べた通り、記憶記号のように、解析とは異なる処理が含まれている。また、今後の文型パターンの機能拡張に対処しやすいと予想される。

一方、デメリットは計算量である。単純な ATN は無駄な状態遷移が生じるため、チャート法に劣る。しかし、履歴の利用による ATN は $O(n^3)$ であると言われており^{*16}、チャート法と同水準の計算量に持ち込める[69]。

2) 文型パターン毎のネットワーク化の理由

ATN のネットワークを作成する際、原則として、1つの文型パターンに対して1つのネットワークを作成する^{*17}。複数のネットワークをトライ構造でまとめることは可能であるが、次の理由からネットワークをまとめることはしない：

- 文型パターンはバリエーションが広いいため、トライ構造化で共通化できる部分は多くはない。
- 文型パターンには字面の記述子があり、ネットワークをまとめない方が、次の方法で効率化が期待できる。任意記号や選択記号の係らない字面は、パターンが解析対象の言語表現と適合するための必須の条件であるので、解析対象の言語表現にその字面が存在するかどうかを判定すれば、ATN を用いる以前に文の不適合を検出できる。

3) 幅優先の理由

ATN において状態遷移を行う際、曖昧性が生じる。この対策として、幅優先の方法と深さ優先の方法がある。幅優先の方法では、曖昧性が生じたとき、複数の状態を同時に保持し、同時に状態遷移を進める。深さ優先の方法では、状態の分岐を記憶しておき、状態遷移を1つ進め、残りの状態は後から進める。次の理由から幅優先の方法を選択する：

- 文型パターンのネットワークは一直線になりやすい。分岐は、選択記号、順序任意記号、離散記号により一旦は生じるものの、再び同じ共通のノードにアークが結び付く。ゆえに、複数の状態遷移が、共通のノードを通過することが多いため、共通の状態遷移が生じ易い。

*16 解析対象となる言語表現の単語数を n とする。

*17 ただし、移動可能記号については原則から外れて、位置を指定する箇所ごとに文型パターンを展開するため、複数のネットワークが対応する。

- 構文構造を解析することが目的ではないので、共通の状態遷移に対して結果を共有できる。
- 深さ優先の方法ならば通過ノードへのマーキングが必要であり、管理が冗長になるが、幅優先の方法では、共通の状態遷移の発生が検出しやすい。

4.3.2.3 ATN を用いた検索

ATN を幅優先で実行するために複数の状態遷移を管理する。本研究では状態遷移を「エージェント」という喩えをすることで簡単に説明する。ゆえに、本研究ではこの ATN を AB-ATN (Agent based Breadth first-ATN) と呼ぶことにする。

エージェントは、遷移状態を表す情報やパターン記述子を働かせるための情報を保持する。具体的には次の情報がある：

- **ATN のための基本情報：**
ネットワーク上の位置、スタック (pop 先などの保持)、言語表現の参照先
- **機能維持のための情報：**
統語的な制約先、意味的な制約先、記憶記号・変数の記憶、順序任意記号の管理、移動可能記号の管理
- **エージェントのための情報：**
親エージェント、エージェントが目覚める時刻

エージェントは、状態遷移の過程で得た変数の値を保持する。制約先は、ハット関数や意味属性制約に示される制約に係る先の形態素を指すのだが、変数の照合の際に係り先を保持する。これらは一部であるが、エージェントは、上述の情報を更新しつつ状態遷移を進める。なお、複数のエージェントの同期は、解析対象の言語表現の形態素の参照位置に基づく。

エージェントの動作は次の7種類である：

- **move 動作：** エージェントは、ネットワーク上に存在し、アークに付与された遷移条件と参照している言語表現の構成要素を比較する。もし、条件が成立すれば、エージェントは、アークを辿り、アークの先のノードへ移動する。この動作を move 動作と呼ぶ。ATN での push と pop もこの動作に位置づける。
- **kill 動作：** エージェントが遷移条件を検査して、成立しなければ、エージェントは消滅する。
- **memorize 動作：** 変数や記憶記号に対応して、適合した形態素や記述子エージェントが記憶する。
- **sleep 動作：** move 動作の際、記述子によっては2つ以上の形態素を参照することがある。形態素の参照位置で、幅優先探索の同期をとるため、形態素の先読みをしたエージェントは、その差の時間だけ休眠する。
- **awake 動作：** 形態素の参照位置により同期がとれると、エージェントは目覚める。
- **breed 動作：** 1つのノードから複数のアークが出ている際、エージェントは子供を産み、各アークに配置する。このとき親エージェントの情報を、子供エージェントが引き継ぐ。
- **join 動作：** ネットワーク上の同一ノードでエージェントが重なるとき、エージェントの情報を適切に引き継ぎ、1つのエージェントにまとめる。

これら7種類の動作を組織的に実行することで、AB-ATN が実現できる。

4.3.2.4 記述子のネットワーク化

move 動作において、アークに付与された遷移条件を検査する。文型パターンの記述子は、その検査ができるようにネットワークに組み込まなければならない。

1) 字面

「出現形字面」および「基本形字面」は、アークに字面の条件として組み込まれる。パターンの作成時に、字面の形態素数は不明であるため、複数の形態素を、1つの字面型条件とする。照合の際、適合した形態素の数が判明するので、その数に応じて sleep 動作が働く。

2) 変数

「意味自由変数」は、通常の ATN での扱いと同一である。すなわち、定義の複雑な変数は、別のネットワークにより検査し、単純な変数はその場で形態素と比較する。

一方、「意味制約付き変数」は、まず「意味自由変数」と同じ手続きにより照合する。そして、パターン照合の後処理の段階(ATNによる照合の外の段階)で、変数に対応した表現に対して意味制約をかける。意味制約は、ATNの段階でかけることも可能であるが、実践的には格交替や省略格への対処を行うために、意味制約はATNの外の段階とするほうがよい。

3) 関数

「ドット関数」は、意味自由変数とほぼ同一の方法でネットワーク化する。異なる点は、対応する言語表現を記憶しない点である。

一方、「ハット関数」は、係る記述子のアークの直ぐ次のアークに配置する。たとえば、パターン「V1[^]rentaiN2」では、動詞変数 V1 にハット関数[^]rentai が係るので、図 4.7(a) のように、V1 の存在するアークの終点のノードを始点とするアークに、[^]rentai を配置する。ここで、ハット関数の係り先は V1 であるが、V1 が複雑な定義である変数ならば、その定義の中にハット関数の係り先が指定されていなければならない。図 4.7(b) では、動詞変数 V の定義例であるが、ハット関数の係り先は、純然たる動詞の変数 VE とするか、それとも、「用言性名詞 ND を 'する'」のうち 'する' の部分とするかである。これを指定するために、VE と 'する' に係り先を明示する記号「[^]」が付与されている。

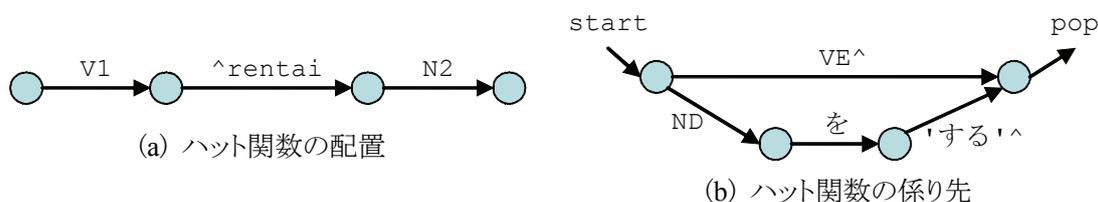


図 4.7 ハット関数のネットワーク化と補助記号付与の例

4) 記号

「選択記号」は、通常の ATN で指定されるとおり、ノードから出るアークを選択肢の数だけ列挙する。

「順序任意記号」は、構成要素に対してカウンタを付与し、それを並列に並べたものを、構成要素の数だけ繰り返す。たとえば、パターン「{N1 は、N2 に、N3 を}V4」は、図 4.8 のようにネットワーク化する。カウンタのアーク cnt は、通過すると加算され、同一カウンタは1回しか通過できない。

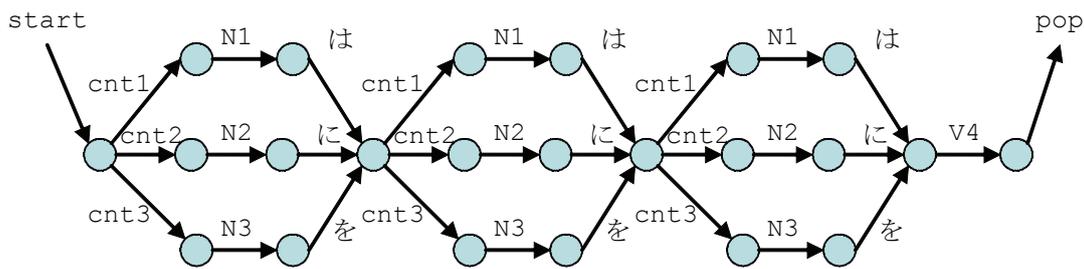


図 4.8 順序任意記号のネットワーク化の例

「移動可能記号」は、複数のパターンに展開する。たとえば、パターン「 $\$1^{\{N1\}}VP2$ ので $\$1VP3$ 」は「 $N1$ は $VP2$ ので $VP3$ 」と「 $VP2$ ので $N1$ は $VP3$ 」の2つに展開する。

「任意記号」と「補完記号」は、通常の ATN で指定されるとおり、jump アークを用いて実現する。

「離散記号」は、ATN の処理としては、ドット関数と同一である。離散記号の性質として、指定する統語構造が任意で挿入可能という点があるが、これは離散記号の定義において実現する。たとえば、「 $/k$ 」は連体修飾要素、すなわち、形容詞連体形、形容動詞連体形、連体詞、「名詞+の」であるが、これらが0回以上繰り返す構造を、 $/k$ に定義する。

「文節境界記号」と「非文節境界記号」は、特殊記号のアークとしてネットワークに配置する。

「まとめ記号」は、文型パターンをネットワーク化する際に構造を形成する働きをする。

「記憶記号」は、アークに補足情報として配置する。すると、エージェントがそのアークを通過するとき、アークの記述子をエージェントが記憶するようになる。たとえば、パターン「 $V1\#1(.kako|.genzai|.mirai)$ 」の場合、図 4.9 のようにネットワークを作る。記憶記号#1 の係るアークに全て、すなわち、まとめ記号の係る範囲に、#1 を配置する。エージェントがアークを成功裏に通過すると、エージェントの保持する情報(第 4.3.2.3 節)としての#1 にそのアークの記述子を記憶する。

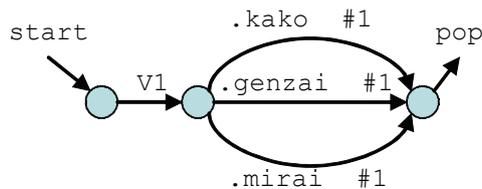


図 4.9 記憶記号のネットワーク化の例

4.3.2.5 高速化

AB-ATN には3つの点で高速化の余地がある。

- 絞り込み前処理
- join 動作
- サブネットワーク照合の履歴

1) 絞り込み前処理

文型パターンのネットワークには、必ず通過しなければならないアークが明らかな場合があり、さらに、そのアークに出現形字面の条件が付与されていることがある。ここで、このような場合の字面のことを必須字面と呼ぶことにする。

文型パターンと言語表現を照合する際、文型パターンの必須字面がその言語表現に存在しないならば、その文型パターンは絶対に適合しない。したがって、必須字面が言語表現に存在するかどうかを調べるだけで、文型パターンと言語表現の不適合が判断できる。

必須字面の判定を、語順を問わない判定とするならば、「文字のビットマップ」を使うことで数バイトの検査で判定できる。文字のビットマップは、必須字面に含まれる文字に対応するビットを立てたものである。文型パターンの必須字面の文字ビットマップと、解析対象の言語表現の文字ビットマップの **and** 演算をして、必須字面の文字ビットマップが残ることで判定できる。

あ	あ	い	い	…	亜	唾	…	凜	瑤
0	1	0	1			0	0		0

図 4.10 文字ビットマップの例

2) join 動作

文型パターンのネットワーク上で、複数のエージェントが同じノード上に存在し、`sleep` しておらず、機能維持の情報が整合するならば、その後の状態遷移は同一である。ゆえに、それらのエージェントは合体して1つにする。なお、サブネットワーク上では、`join` 動作はしない。

3) サブネットワーク照合の履歴

変数、ドット関数、離散記号は、サブネットワークを用いて照合する。サブネットワークの照合はコストが高い。その照合結果を保存しておき、再利用することで、重複する照合を削減することが考えられる。この処理を履歴処理という。

履歴処理の使用には次の全ての条件が満たされなければならない：

- 解析対象の言語表現が同一
- サブネットワークが、履歴と同一(異なる変数番号, 異なる文型パターンでも良い)
- サブネットワークと照合し始める言語表現の位置が、履歴と同一

履歴処理が提供する情報は、サブネットワークの照合により得られる状態遷移である。照合結果が曖昧な場合、複数の状態遷移となる。状態遷移をエージェントが管理しているので、履歴処理はエージェントを産み出す処理に他ならない。

たとえば、言語表現「₁ 彼₂ の₃ 車₄ は」を、2つのパターン「NP1 に」と「NP1 は」と照合する場面で説明する。まず、「NP1 に」と照合する途中段階で、「彼」を参照すると、「NP1=彼」と適合するので、履歴には、参照番号「1」と変数「NP」の組をキーとして、「NP=彼」という記録が登録される。次に「車」までを参照すると、参照番号「1」と変数「NP」の組をキーとして、「NP=彼の車」という記録が登録される。

次に、「NP1 は」というパターンを照合する段階で、「彼」を参照する際、すなわち、参照番号が「1」の際、「NP1」の照合を行おうとすると、履歴には、「NP=彼」および「NP=彼の車」が登録されていることから、2つの適合結果が即座に得られる。その結果、状態遷移が2通りとなるため、エージェントが2つ作られる(図 4.11)。

履歴処理はここまでであるが、理解を深めるためにその続きの説明をする。エージェント2は、「NP=彼」とした場合であるので、言語表現の参照先は2である。エージェント3は「NP=彼の車」とした場合であるので、言語表現の参照先は4である。したがって、エージェント2はすぐに同期がとられて、「の」と「は」の照合が始まる。エージェント2の照合は失敗し `kill` 動作が働く。一方、エー

エージェント3は同期をとるため sleep 動作となる。時間が進み同期がとられると、「は」の照合に成功し、エージェントが4の状態になる。

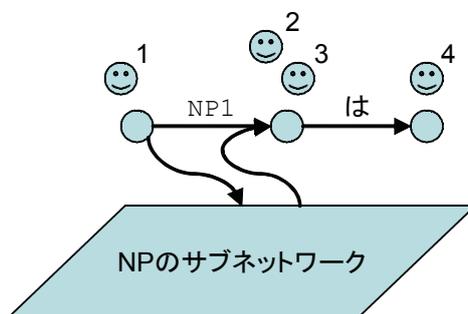


図 4.11 履歴処理の様子

4.3.3 構造照合型文型パターン検索

本節では、文型パターン辞書を用いて入力される言語表現を解析するシステムを構成する。このシステムは、入力表現と文型パターンの照合を行うことで、辞書から適合するパターンを検索する。そこで、このシステムを「構造照合型文型パターン検索システム」と呼ぶ。

4.3.3.1 システムの構成

構造照合型文型パターン検索システムの構成を図 4.12 に示す。「形態素解析」、「意味属性付与」、「絞り込み前処理」、「AB-ATN 照合処理」、「意味制約後処理」、「文型パターン辞書」、および、「変数・関数・記号定義」で構成される。言語理解の表層レベルの解析に該当するのは、「形態素解析」、「意味属性付与」である。本検索の中心的な処理部を「SPM (Structural Pattern Matching and search)」と呼ぶ。SPM は、「AB-ATN」、「絞り込み前処理」、「変数・関数・記号定義」で構成される。形態素情報と文型パターン辞書が、SPM への入力となる。

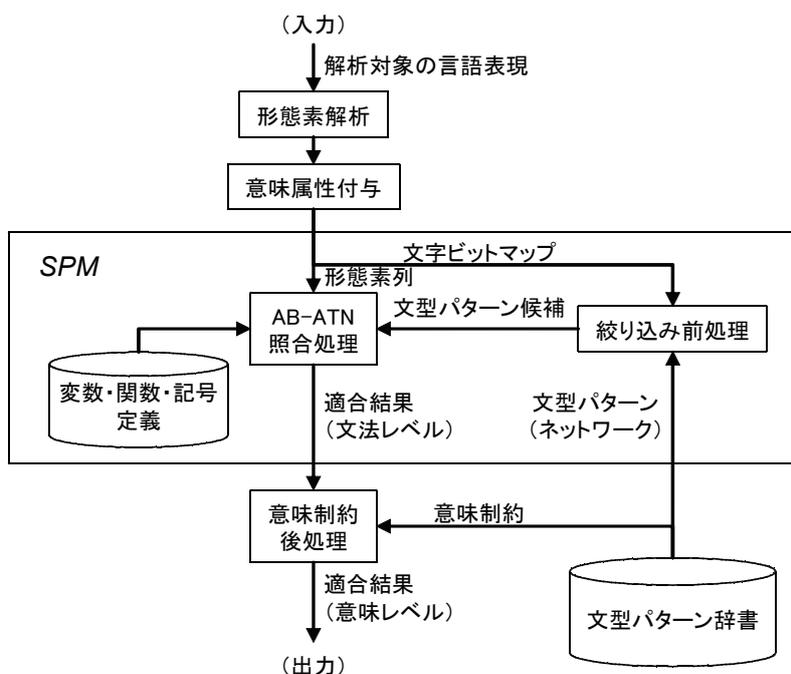


図 4.12 構造照合型文型パターン検索 SPM のシステム構成

1) 表層レベルの解析

「形態素解析」は、既存の形態素解析ツール^{*18}に、形態素調整の処理を付与したものである。形態素調整は、見出し語を入力文での活用形に変更する。たとえば、入力文「はしります」に対して「走ります」と変更する。意味属性付与は、日本語語彙大系[13]の体言・用言の意味属性、講談社類語大辞典[38]の意味コード、日本語表現意味辞書[70]の体言・用言の意味属性を付与する。

2) SPM

「絞り込み前処理」は、「文型パターン辞書」より、入力文と適合する可能性のある文型パターンを抽出し、「AB-ATN 照合処理」に引き渡す。

「AB-ATN 照合処理」は、入力文の形態素列と文型パターンの照合を行う。「変数・関数・記号定義」には、サブネットワークを含んでいる。変数、関数、記号(離散記号)の定義を自由に追加・修正することができる。定義に依存した文型パターンの性能については、第 4.5 節で述べる。

AB-ATN 照合処理の結果は、適合結果として一旦ファイルに出力される。そのファイルには、入力文に適合した文型パターン、適合したパターン記述子、変数や記憶記号等への代入値が収録されている。この段階の適合結果は、文法レベルの照合に成功したものである。

3) 後処理

「意味制約後処理」は、文型パターン辞書より意味制約の情報を抽出し、適合結果の中から意味的制約を充足するものを抽出する。この適合結果は意味レベルの照合に成功したものである。

4.4 文型パターン辞書検索の基本性能の評価

4.4.1 実験条件

構造照合型文型パターン検索システムの基本性能を評価する。基本性能の評価は、形態素解析が終了してから計測を開始し、文法レベルの適合結果を得るまでの時間を対象とする。意味制約後処理については、応用時に使い分けを行うので、基本性能の評価対象外とする。

評価実験に使用する計算機は、CPU が AMD Athlon64 (2.4GHz)、メモリが 2GB、OS が SUSE-Linux (64bit) である。

ここで実験に使用する文型パターン辞書は、日英対訳コーパスから作成した単語レベル日本語文型パターン辞書である[71]。日英対訳と日英単語レベルパターン対の例を示す：

- AC000008-00
 日文：あの建物はああ見えても新しい。
 英文：That building is still new despite appearance.
 日パ： /y\$1^{/cfGEN1/kN2 (IM:12100, IM:13660) は}/tcfk ああ/f (見え|みえ) ても\$1/yfAJ3 (IY:A5B0)。
 英パ： AJ1 N2 be still AJ3 despite appearance.
- AC000024-00
 日文：資本主義と社会主義は相入れない概念だ。
 英文：Capitalism is incompatible with socialism.
 日パ： /ytcfkN1 (IM:14120, IM:14A10) と/tcfkN2 (IM:14120, IM:14A10) は /cf 相入れない!概念だ。

*18 NTT との共同研究のプロジェクトにおいて ALT-JAWS を使用した。その他のテーマに関して、奈良先端大学院大学・松本研究室の chasen および cabocha を使用した。cabocha は、文節境界を得るために使用し、構文情報は使用していない。

- 英パ: N1 be incompatible with N2.
- AC000051-00
 - 日文: これは人々に愛唱されている古い民謡の一つです。
 - 英文: This is one of the old folk songs people love to sing.
 - 日パ: /y#1{/tcfkN1(IM:10000,IM:15000)は,/tcfkN2(IM:11100,IM:16770)に}/cf(V3.rareru.teiru^rentai|ND3をされている)/fAJ4(IY:A5B0)^rentai/fN5(IM:14230,IM:15243)の/kN6(IM:16720,IM:16730).#da。
 - 英パ: N1 be N6 of AJ4 N5 N2 V(V3|ND3).

この辞書には、単語レベルパターンが 122,718 パターンが収録されている。変数の種類は 12 種類、離散記号の種類は 17 種類であり、定義のための規則は 255 個である。

実験は leave-one-out cross validation 法とする。すなわち、入力文は、このコーパスにおける日本語文であるが、その入力文から作られたパターンは照合に使用しない。

4.4.2 実験の実施と結果

4.4.2.1 実験 4-1 : 文の長さ と 検索時間の関係

単語レベル文型パターン辞書から文に適合するパターンを全て検索する時間を計測する。コーパスから文の長さ(単語数)ごとに 200 文ずつ抽出した 5,527 文を用いる。長さが 29 以上の文は 200 文に満たなかったため、得られた文全てを用いた。図 4.13 に実験結果を示す。このグラフには、パターンが検索できた場合(ヒット)とできなかった場合(ミス)についての平均検索時間、および、それらの総合の平均検索時間を載せている。次のことがわかる:

- ヒットの場合とミスの場合では、ミスの場合が僅かに時間が短い、大差がない。
- 200 文まで抽出できなかった長さの文(長さ 29 以上)では、2つの場合とも検索時間のばらつきが目立つ。
- 長さ 28 以下について総合の場合のプロットを近似式で表すと、 $y = 0.0004x^2 + 0.004x - 0.007$ となった。文の長さ m に対して、実効的には、 $O(m) \sim O(m^2)$ である。

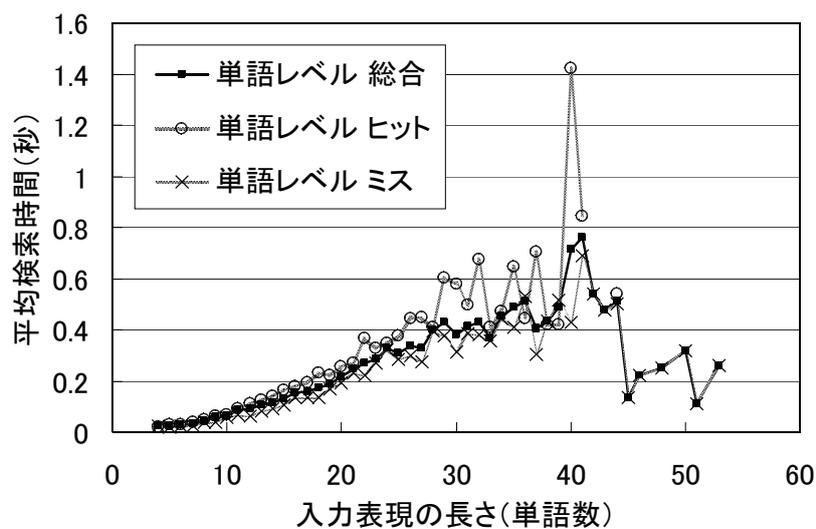


図 4.13 文の長さ と 検索時間の関係

4.4.2.2 実験 4-2 : 辞書の規模と検索時間の関係

単語レベルパターン辞書を 10 分割して、徐々に用いるパターンを増やしながら、辞書の規模ごとの平均検索時間を調べる。コーパスからランダムに取り出した 6,181 文を用いる。結果を図 4.14 に示す。横軸は辞書を使用した割合で、全て使用した場合を 10 とする。また、プロットした点の近似式を求めた。次のことが読み取れる：

- パターン辞書の規模 p (パターン数) に対して平均検索時間は、 $O(p)$ である。

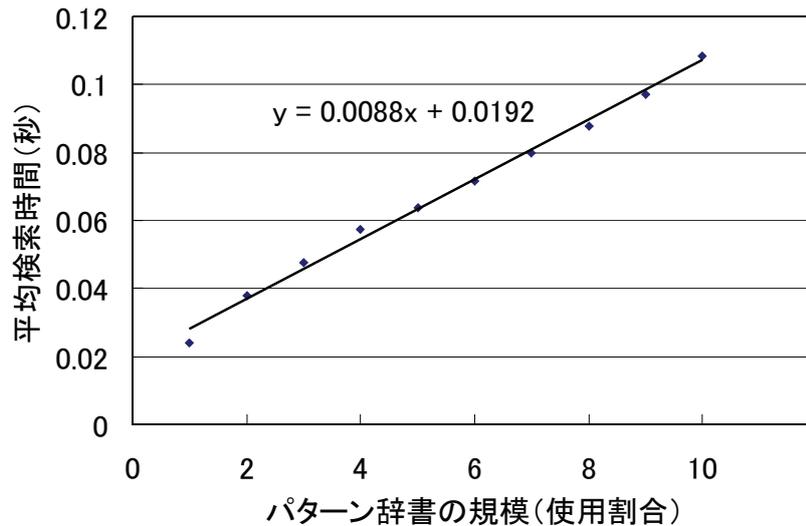


図 4.14 辞書の規模と検索時間の関係

4.4.2.3 実験 4-3 : 文の長さメモリ使用量の関係

メモリ使用量は、遷移状態(エージェント数)、履歴(キー、値)、および、変数等への代入値が主たる要因となって決まる。実験 4-1 と同じ文集合を用いて検索実験を行い、文の長さに対するメモリ使用量を計測する。

表 4.1 に、各要因について、実験で使用された個数の最大値、1 個あたりに使用するメモリ (byte 単位)、実験で使用された最大値・平均値 (Mbyte 単位)を示す。図 4.15 は、これらの合計値を文の長さごとに集計した平均値を示す。表 4.1 と図 4.15 より次のことが言える：

- 遷移状態は1個あたりのメモリ使用量は大きいですが、全体で見るとメモリ使用量に占める割合は大きくはない。
- 代入値のために使用したメモリ使用量が最も大きい。1つのパターンとの照合の途中で失敗した場合でも、入力文が変わらない限り、使用した代入値は履歴に残すために破棄しなかったためである。
- 使用メモリ量の平均値(図 4.15)と最大値(表 4.1)を比較すると、平均使用メモリ量の主要因は「代入」である。

表 4.1 各要因のメモリ使用量

要因	個数		使用量	
	最大値(個)	1個当り(byte)	最大(MB)	平均(MB)
遷移状態	576	648	0.36	0.008
履歴:キー	1,337	8	0.01	0.003
履歴:値	3,492	24	0.08	0.020
代入	3,318,437	12	37.98	2.888

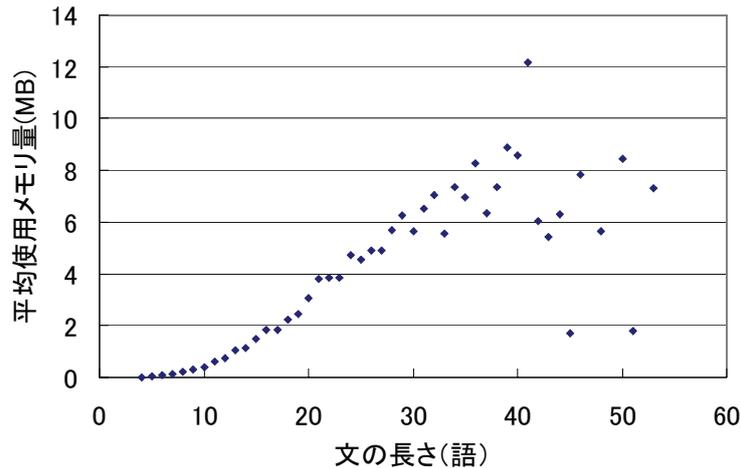


図 4.15 文の長さ和使用メモリの関係

4.4.2.4 実験 4-4 : 絞り込み前処理の効果

絞り込みの条件は、文型パターンの中で適合することが必須の字面である。そこで、まず、単語レベル文型パターン辞書における必須字面を集計した。必須字面の種類数は、93,792 種類であり、使用頻度が 1 であるものは 88,303 種類、2 であるものは 3,335 種類であった。逆に使用頻度の高いものは表 4.2 のとおりであるが、約 12 万の単語レベルパターンにおいて、「。はを」の字面を必須字面とする文型パターンは 1%に満たない。こうした、必須字面の特殊性により絞り込みの効果が期待できる。

表 4.2 単語レベル文型パターン辞書において使用頻度の高い必須字面(上位10件)

順位	必須字面	頻度
1	。はを	1,127
2	。は	909
3	。を	708
4	。には	559
5	。とは	467
6	。にあを	459
7	。のはを	457
8	。が	450
9	。がは	365
10	。にを	321

次に、絞り込み前処理を無効にして文型パターンを検索する。文の長さごとの平均検索時間を求めるため、実験 4-1 と同じ文集合を用いて実験する。実験1と比較した結果を図 4.16 に示す。この結果より次のことが言える：

- 検索時間に大きな開きが見られる。
- 文の長さが 16 語を超えたあたりから、絞り込み無しの場合の平均検索時間の増加が鈍くなっている。

ここで、後者の理由を調査すると、処理の途中でアボートが生じていることが分かった。長さ 16 語以上の入力において、アボートが見られ始め、長さ 30 語の文集合においては 37%(99/156)の文がアボートした。アボートの理由はメモリ不足である。不適合でも本システムでは代入のためのメモリを使用するためである。

また、アボートしなかった文の実行結果を調べると、平均検索時間は、絞り込み無しの場合、1.449 (秒/文)、絞り込み有りの場合、0.225 (秒/文)であった。絞り込みにより検索時間は約 6 分の 1 に短縮されていたことが分かった。

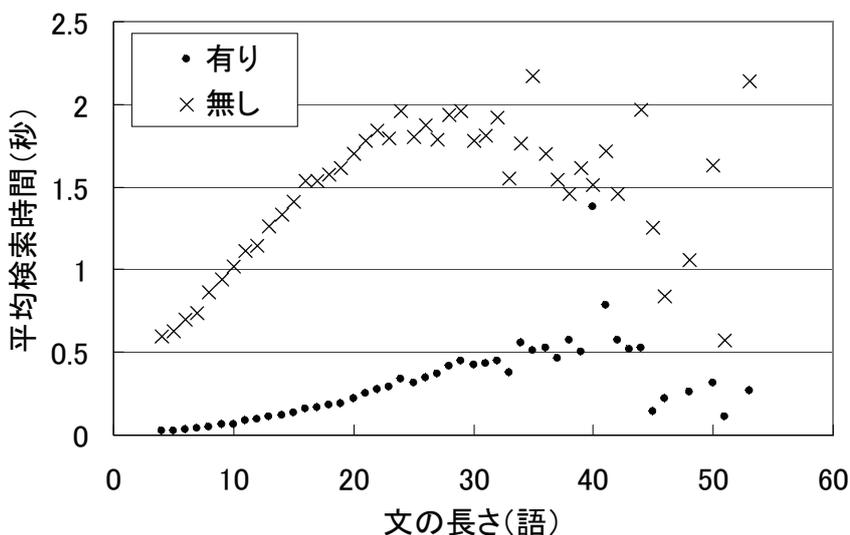


図 4.16 絞り込み前処理の有無による平均検索時間の比較

4.4.2.5 実験 4-5 : join 動作の効果

join 動作の効果を調べるために、join 動作の機能を停止させ検索時間を計測する。実験 4-1 と同じ文集合を用いて実験したところ、join 動作をさせない場合に多くアボートが発生した。アボートせずに動作した 99 文について動作状況を調べる。

その結果(表 4.3)から次のことが言える：

- エージェント数(遷移状態の並列数)は、最大値で 20 分の 1 (8,356/163,901)に、平均値で 5 分の 1 (1,833.2/8,688.6)に、それぞれ抑えることができた。
- join 動作により、検索時間は 9 分の 1 に短縮されていたことが分かった。
- join 動作により、アボート数が大幅に抑えられていたことが確認された。

エージェント数が爆発的に増えるのは、幅優先の方法であるためである。以上の結果より join 動

作が幅優先の方法に必要な不可欠であることが確認された。

なお、本実験において、サブネットワークを定義する際、ネットワークの改善を行っていなかった。ネットワークの改善とは、非決定的なネットワークを作成することである。パターン照合において、サブネットワーク上の通過経路の違いは問わないので、サブネットワーク内からさらに下位のサブネットワークを呼び出す場合に、それをあらかじめ展開しておくことは許されている。さらに、展開後に非決定的なネットワークに変更すると、重複する条件判断を削減することができる。

表 4.3 join 動作の有無による処理性能の比較

検査項目	join 動作あり	join 動作なし
join 回数	788,127 回	-
1 ネットワーク当たり*1 エージェント数*2	1.23 回	-
最大値	8,356 個	163,901 個
最大値の合計	181,489 個	860,171 個
1 文当たり平均	1,833 個	8,689 個
検索時間	10.9 秒	93.3 秒

*1 照合処理をしたネットワークの数で join 回数を割る

*2 瞬間的に並列的に存在した数

4.4.2.6 実験 4-6 : 履歴機能の効果

履歴機能の効果を調べるために、履歴機能を使用する場合と使用しない場合について検索時間を比較する。文集合は、実験 4-1 と同じものを使用する。

検索処理がアボートした回数を求めた(表 4.4)。長さが 16 以上になるとアボートの発生が始まり、長さが 30 になると 57 回(156 回中)となった。アボートの理由はメモリ不足である。また、アボートせずに検索できた場合の検索時間を比較したところ、履歴機能により検索時間が 38 分の 1 に短縮されていたことが分かった。

表 4.4 履歴の有無による処理性能の比較

処理項目	履歴機能あり	履歴機能なし
文数		5,527 文
アボート率	0.036%	78.1%
アボート数	2 文	4,319 文
文数		1,208 文
検索時間	49.74 秒	1904.93 秒

4.4.3 実験のまとめ

以上の結果を以下にまとめる:

- 基本性能に関して
 - 処理時間は、文の長さ(単語数 m)に対して $O(m) \sim O(m^2)$, 辞書のサイズ(パターン数 p)に対して $O(p)$ である。
 - 使用メモリは、文の長さ(単語数 m)に対して $O(m^2)$ である。実験値として、最大で約 38MB の場合があったが、平均的には 2.8MB であった。

- 高速化に関して
 - 「絞り込み前処理」は、検索時間を約 6 分の 1 に短縮する効果がある。
 - 「join 動作」は、検索時間を約 9 分の 1 に短縮する効果がある。
 - 「履歴機能」は、検索時間を約 38 分の 1 に短縮する効果がある。
 - 3つの高速化は、1つでも使用しないとメモリ不足により処理がアボートする。このことから、3つの高速化は、システムの安定化にも効果がある。

以上より、現在の計算機的能力から言えば、本システムは現実的な処理能力であるといえる。

4.5 文型パターンの記述子に関する評価

本節では、文型パターンが、解析対象の言語表現に柔軟に適合することを目指して設計された離散記号、任意記号、ドット関数(時制、相・様相関数)について、導入した効果を定量的に明らかにする。なお、評価実験の対象として、規模の大きな文型パターン辞書が必要となるので、日英機械翻訳用に構築された文型パターン辞書を使用する(第 4.4 節を参照)^{*19}。

4.5.1 柔軟性を高める記述子

1) 離散記号

文型パターンの記述子は、適合の条件が厳しい。たとえば、パターン「N1 は N2 より AJ3。」は、「猫はネズミより強い。」と適合するが、「黒い猫はネズミより強い。」と適合しない。なぜならば、変数 N1 に形容詞が適合しないからである。

このパターンの表す「2者の比較」という抽象的概念に、N1 に先行する形容詞の存在は任意である。そこで、抽象的概念に対して任意な表現が文型パターンに挿入できる位置を示すために、離散記号を使用する。たとえば、パターン「/kN1 は/kN2 より AJ3。」には、離散記号/k が導入されているので、形容詞などの連体修飾語句の挿入された表現に対して柔軟に適合する。

その他の離散記号を含めて5種類がある。表 4.5 にその一覧を示す。離散記号は既存の文型パターンに追加して使用できる。その挿入位置の目安もこの表に示す。

こうして、上述の例にさらに離散記号を挿入すると、パターンは「/ytcfkN1 は/cfkN2 より/cfAJ3。」となる。なお、連続する離散記号を1つにまとめて記述することができる。

表 4.5 離散記号の種類と使用位置

記述子	適合する表現	使用位置
/y	副詞的な節 (連用中止節, 従属節, 仮定節)	パターンの先頭, 副詞的な節の直後
/t	連体節	連体節の無い名詞句の直前
/c	格要素	格要素の前後
/f	副詞的な句・単語 (副詞, 形容詞連用形, 形容動詞連用形)	格要素の前, 用言の直前
/k	形容詞的な句・単語 (形容詞連体形, 形容動詞連体形, 「名詞+の」, 連体詞)	名詞句の直前

*19 ただし、文型パターン辞書の開発と平行して本研究は進められたため、パターン数などの細かい値に異なるところがある。

2) 任意記号

文型パターンの記述子は、意味解析において決定力が強い。言い換えると、意味解析の決定力が弱い記述子は、文型パターンと言語表現の適合する条件を弱くすべきである。照合だけを考えて、決定力の弱い記述子は、離散記号で記述すれば簡単である。しかし、応用への処理を考慮すると、離散記号は、変数のように表現を記憶しないため、必ずしも有効ではない。そこで、任意記号により、決定力の弱い記述子は、適合条件として必須でなくする。

たとえば、「スクーターは自動車より燃費が良い。」という表現は、パターン「/ytcfkN1 は /cfkN2 より /cfAJ3。」に適合するが、「燃費が」が「/c」に適合するため、「燃費が」の解析には別途処理が必要となる。一方、任意記号付きのパターン「/ytcfkN1 は /cfkN2 より /cf[/kN3 が] /fAJ4。」を用いると、「N3=燃費」となるので、直ぐに情報が抽出できる。

評価実験で使用する辞書には、表 4.6 に示すように任意記号が使用されている。この辞書の文型パターン数は 122,619 パターンであり、そのうち 38,489 パターン(31%)が任意記号を使用している。

表 4.6 任意記号の使用例(使用頻度の高い上位 10 件)

順位	任意記号の使用例	意味	使用頻度
1	[REN]	連体詞の任意化	9,480
2	[N の]	「名詞+の」の任意化	7,022
3	[ADV]	副詞の任意化	6,306
4	[AJ]	形容詞の任意化	1,512
5	[AJV]	形容動詞の任意化	1,409
6	[TIME]	時詞の任意化	687
7	[その]	「その」の任意化	538
8	[TIME の]	「時詞+の」の任意化	371
9	[TIME は]	「時詞+は」の任意化	347
10	[自分の]	「自分の」の任意化	277

※ 使用頻度は、記述子を含むパターン数

3) 時制・相・様相関数

日本語の助詞・助動詞、および、相当表現は音便による表記の揺らぎや同義異体の表現が見られるため、助詞・助動詞に対するドット関数が利用できる。たとえば、関数「.kako」は、「た／だ」に対応し、関数「.dantei」は「だ／です／である」に対応する。

時制・相・様相については、応用の処理の段階ではルールに基づく自動調整が考えられる。そのため、時制・相・様相について文型パターンが介入する必要の無い場合が考えられる。たとえば、表現「今朝の鯖はいつもの鯖よりおいしかった。」に対して、「2者の比較」という概念を解析する目的の文型パターンは、「過去」と「現在」の区別を重視しないので、「/ytcfkN1 は /cfkN2 より /cf[/kN3 が] /fAJ4[.kako]。」というパターンの記述ができる。

評価実験で使用する辞書において、時制の関数について、32,384 パターンが関数.kakoを含み、54 パターンが関数.darouを含んだ。89,451 パターンが関数.kakoも.darouも含まなかった。一方、相・様相関数については、表 4.7 に示す。

表 4.7 相・様相のドット関数(使用頻度の高い上位 10 件)

順位	記述子	適合する表現	使用頻度
1	.teiru	ている／でいる	9,565
2	.rareru	れる／られる	7,378
3	.dantei	だ／です／である	7,166
4	.hitei	ない／ず	6,731
5	.desu/.masu	です／ます	4,302
6	.meireigo	て下さい／で下さい／給え	2,804
7	.you	よう	2,734
8	.suitei	ようだ	1,779
9	.sase	せる／させる	1,199
10	.tekureru	てくれる／でくれる	1,071

※ 適合する表現は、終止形で記載したが、いずれの活用形でも適合する。

※ 使用頻度は、記述子を含むパターン数

4.5.2 評価パラメータ

離散記号、任意記号、時制・相・様相のドット関数を、日英文型パターン辞書に導入することで、この辞書の柔軟性が高まった程度を評価するためのパラメータを考案する。

評価パラメータは、これらの記述子の導入の有無により、文型パターンと文との適合の度合いの違いを表すものとした。1文と1パターンの照合では記述子導入の効果が分からないので、大量の文とパターンを照合することにより、パターンの適合した文の割合を求めて、その割合を基本的な評価パラメータとする。ここでこの割合を文型パターン辞書の「被覆率特性」と呼ぶことにする。

記述子の導入時の被覆率と、導入しない時の被覆率の差を評価パラメータとして使用することが考えられる。ところが、ある程度まで被覆率が高められた状態からさらに被覆率を高める際の労力と、ほとんど被覆率が高められていない状態からそれを高める際の労力とでは、大きな違いがあるため、被覆率の単純な差だけで説明することは誤解が生じやすい。そこで、「文型パターン辞書の拡大率」を新たに提案する。

4.5.2.1 文型パターン辞書の被覆率特性

大量の文とパターンの照合により、適合した文の数や、パターンの数を以て評価するパラメータを導入する：

- $R1$: 照合成功率
入力文集合に対して、パターンの適合した文の割合
- N : 平均適合パターン数
入力文集合に対して、適合したパターンの数

文型パターン辞書が、様々な種類の文に適合することが望ましい。その度合いを直接的に表す評価パラメータが $R1$ である。一方、適合したパターンが正しく意味解析を導くとは限らないが、適合したパターンの多い方が、意味的な正解を含有する可能性が高い。そこで、入力文1文につき平均何件のパターンが適合したのかを表すパラメータが N である。

$R1$ と N は、文集合を構造照合型文型パターン検索システムに入力し、その照合結果から求める。そのために次の計算式を用いる：

$$R1 = M / I \times 100 \quad (\%) \dots\dots\dots(4.1)$$

$$N = P / I \quad (\text{patterns})\dots\dots\dots(4.2)$$

ここで、 I は入力文集合の文数である。 M はパターンが適合した文の数である。 P は適合したパターンの数である。

図 4.17 を例に説明する。約 12 万の日本語文を収録したコーパスが入力文集合として、構造照合型文型パターン検索に与えられる。文型パターン辞書には、単語レベル文型パターンが約 12 万件収録されている。この辞書のパターンを leave-one-out cross validation 法で検索・照合を行う^{*20}。照合の結果を見ると、入力文に1つ以上パターンが適合する場合と全く適合しない場合がある。パターンの適合した文を「適合文」と呼び、そのパターンを「適合パターン」と呼ぶことにする。また、パターンの適合しなかった文を「不適合文」と呼ぶことにする。

この図の拡大部分においては、入力文数 $I = 3$ に対して、適合文の数 $M = 2$ 、不適合文の数は 1 である。また、適合パターン数 $P = 5$ である。したがって、 $R1 = 67 \%$ 、 $N = 1.67$ (patterns) となる。

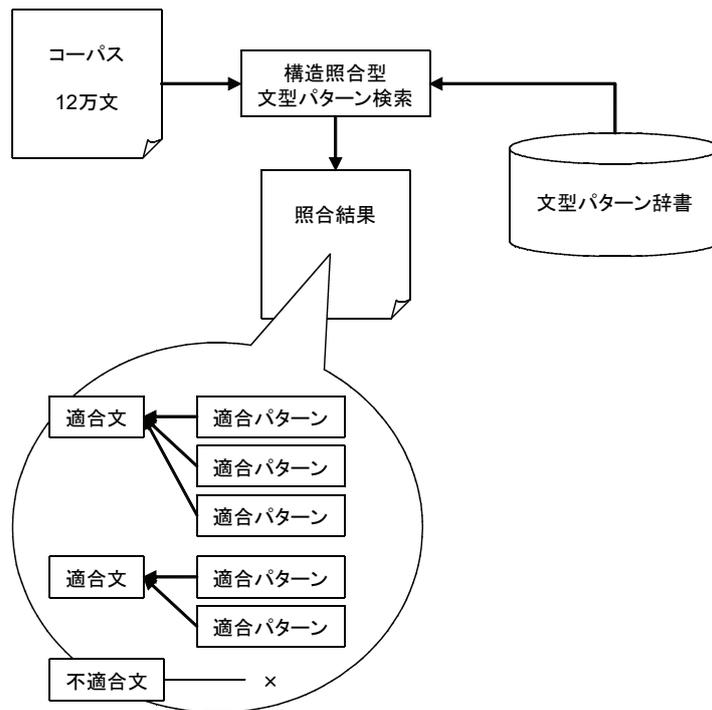


図 4.17 文型パターン辞書の照合実験の構成

ここで、 $R1$ と N の特性について調べる。 $R1$ と N は、文型パターン辞書の規模に対して単調増加する。その特性については、辞書を分割し、サイズの異なる辞書を用いて照合実験を行うことで得られる。

*20 このコーパスから文型パターン辞書を作成した。しかし、照合実験において、入力した文から作られたパターンは照合に使用しない。

図 4.18 がその結果である. (a)と(b)の両方とも, 横軸に文型パターン辞書のサイズ(パターン数 p とする)を設定し, 縦軸に R_1 と N の値を設定している. 各プロットが実測値, 実線が近似曲線である. 具体的に, R_1 については, 各サイズについて 5 回のランダム抽出で辞書を作成した.

辞書の規模が 12 万件のとき, $R_1 = 48.2(\%)$, $N = 13.4$ (patterns)であったところ, 辞書の規模が 4 万件のときは, R_1 が $32.1(\%)$ から $36.2(\%)$ までの揺れがあった.

プロットした点に対して以下の式をベースに回帰分析し, 近似式を求めた.

$$R_1 = (1 - \exp(-\lambda_1 p^{\lambda_2})) \times 100 \quad (\%) \dots\dots\dots(4.3)$$

$$N = \lambda_3 p \quad (\text{patterns}) \dots\dots\dots(4.4)$$

その結果各係数は次のようになった:

$$\lambda_1 = 0.005038$$

$$\lambda_2 = 0.47171$$

$$\lambda_3 = 0.0001093$$

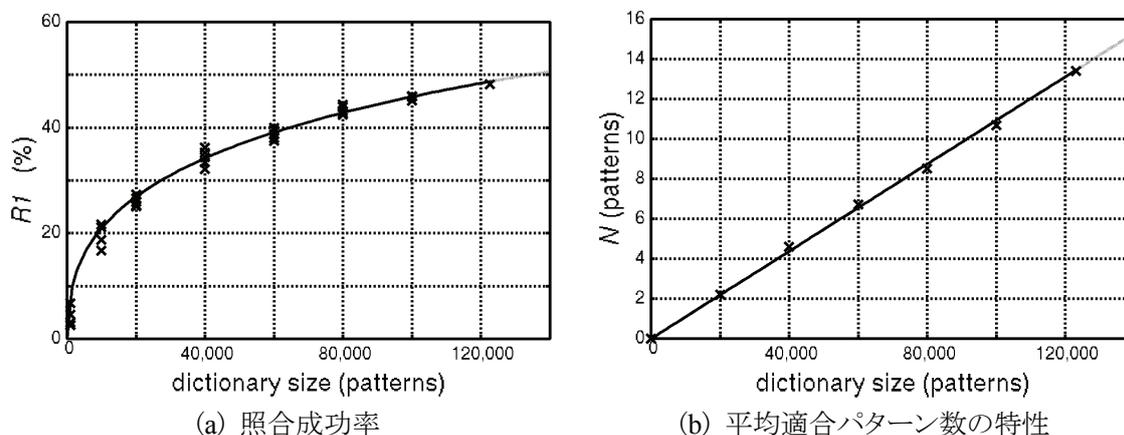


図 4.18 単語レベル文型パターン辞書の被覆率特性

4.5.2.2 文型パターン辞書の拡大率

図 4.18(a) によると, パターン辞書の規模(パターン数) p に応じて R_1 の増加量は異なる. たとえば, R_1 を 10 向上させる難しさを, 元々の R_1 が 20 である場合と 40 である場合とで比べてみると, 前者の場合がパターン数を 2 万程度追加すれば実現できる一方で, 後者の場合が約 6 万パターン追加する必要がある. そこで, 被覆率の差を比較するのではなく, パターン辞書の規模に換算して比較するための評価パラメータを提案する.

- η : 文型パターン辞書拡大率

評価対象の文型パターン辞書(対象辞書と呼ぶことにする)のパターン数を, 基準文型パターン辞書(基準辞書と呼ぶことにする)の文型パターン数に換算した際の倍率

η は, 次の式で定義する.

$$\eta = X / B \dots\dots\dots(4.5)$$

ここで、 B は基準辞書のパターン数、 X は、対象辞書と同じ性能を出すために必要な基準辞書のパターン数である。

基準辞書の被覆率特性は、図 4.18 のように実験的に求めることができ、パターン数 p の関数となっている。この関数を $c(p)$ とする。実験的に求めた対象辞書の被覆率を C_t とし、関数 $c(p)$ の逆関数 c^{-1} を用いると、対象辞書と同じ性能を出すために必要な基準辞書のパターン数 X は、次式により得られる。

$$X = c^{-1}(C_t) \dots\dots\dots(4.6)$$

一般に関数 c は実験的に求められる近似曲線であるので、誤差を打ち消すためには、基準パターン辞書の規模 B も換算値に置き換える。基準辞書の被覆率を C_b としたとき、 B は次式により得られる。

$$B = c^{-1}(C_b) \dots\dots\dots(4.7)$$

ゆえに、式(4.5)に代入して、以下の式により文型パターン辞書拡大率を求める。なお、換算のために使用する被覆率特性の関数を明示するために添え字を付けることがある。

$$\eta_c = \frac{c^{-1}(C_t)}{c^{-1}(C_b)} \dots\dots\dots(4.8)$$

たとえば、図 4.18(a) の被覆率特性が、基準辞書のものであったとする。ある対象辞書が、 $R1 = 40$ (%)であったとき、対象辞書の規模は、基準辞書のパターン数に換算すると、約6万件となる。基準辞書の規模が約12万件であるので、 η_{R1} は約 0.5 となる。

基準辞書と対象辞書の性能を比較して、対象辞書の方が優れているならば、 $\eta > 1$ となり、劣っているならば、 $\eta < 1$ となる。

4.5.3 柔軟性の評価

4.5.3.1 方法

柔軟性の評価実験に使用する日英文型パターン辞書には、既に、離散記号と任意記号が導入済みである。これらの記号の効果を調べるためには、これらの記号を削除した辞書を新たに作成し、これを対象辞書として使用する。削除前の辞書を基準辞書として拡大率を計測し、 $\eta < 1$ ならば対象辞書すなわち削除した辞書の劣位が明らかになり、削除された記号の効果があつたことが判定できる。一方、時制・相・様相のドット関数を任意化する場合の効果については、基準辞書において実際に任意化することで対象辞書を作成して、拡大率を計測する。

4.5.3.2 対象辞書の作成例

基準辞書のパターン、および、対象辞書のパターンの作成例を示す。

1) 離散記号と任意記号の削除による対象辞書の作成

離散記号と任意記号の削除により作成される対象辞書のパターンの例である：

(原文)

こんなに客が少なくては商売上がったりだ。

(基準辞書のパターン)

/y#1[こんなに]/tk 客が/cfAJ2(て|で)は/tcfkN3 上がったりだ。

(/t 削除の対象辞書のパターン)

/y#1 [こんなに]/k 客が/cfAJ2 (て|で)は/tcfkN3 上がったりだ。

(任意記号削除の対象辞書のパターン)

/y#1 こんなに/tk 客が/cfAJ2 (て|で)は/tcfkN3 上がったりだ。

こうして作成された対象辞書は、機能が劣るはずである。後述の実験結果においては、この対象辞書の拡大率 η が 1 より小さくなることを期待する。

2) 時制のドット関数の任意化

以下は、時制のドット関数を任意化する例である：

(基準辞書のパターン1)

/yN1 は/cf うなずきながら/tcfkN2 の/cfkN3 を/cfV4.kako。

(時制を任意化した対象辞書のパターン1)

/yN1 は/cf うなずきながら/tcfkN2 の/cfkN3 を/cfV4[.kako|.darou]。

(基準辞書のパターン2)

/yN1 は/cf 新しく/cf 買った/cfkN2 を/cfV3.hitei。

(時制を任意化した対象辞書のパターン2)

/yN1 は/cf 新しく/cf 買った/cfkN2 を/cfV3.hitei[.kako|.darou]。

時制の任意化とは、時制を表す部分を「[.kako|.darou]」とすることである。これにより、過去の表現、未来(だろう)の表現、現在の表現(過去も未来も無い場合)のいずれにも適合する。

ただし、過去形でしか成り立たない表現も一律に任意化するため、意味的には正しくないパターンになることがある。しかし、こうした任意化は無意味なことではない。時制の任意化に対する正誤判定は、人手により1件ずつ判定する必要があるコストが高い。もし、今回の時制任意化に効果が無いならば、正確な任意化作業に着手する意味が無いことが分かる。規模の大きな知識ベースの開発過程において、作業方針を立てる上で、今回の評価実験は意味がある。

3) 相・様相のドット関数の任意化

相・様相のドット関数は、基準辞書には、37種類が使われているが、本評価実験では、表 4.7 に示した 10種類について任意化の対象とする。以下に例を示す：

(任意化前) N1 が V2.teiru.hitei。

(任意化後) N1 が V2.teiru[.hitei]。

否定を表す関数「.hitei」が任意化された。たとえば、「雪が降っている。」は任意化前のパターンは適合しないが、任意化後のパターンは適合する。

4.5.3.3 結果

作成した対象辞書は 20 個になる。その内訳は、離散記号を全て削除した対象辞書(1個)、離散記号の各種類について削除した対象辞書(5個)、任意記号を削除した対象辞書(1個)、時制のみ任意化した対象辞書(1個)、相・様相の各種類について任意化した対象辞書(10 個)、相・様相をまとめて任意化した対象辞書(1個)、時制・相・様相をまとめて任意化した対象辞書(1個)である。

第 4.5.2 節で使用したコーパスを使用し、また、同じ照合方法で対象辞書の照合実験を行う。それぞれの対象辞書について拡大率を求めることができる。被覆率として $R1$ と N があるので、それぞれの近似曲線を使用して、拡大率 η_{R1} と η_N を計測する。

計測した結果を表 4.8 に示す. この表から次のことが言える:

- (1)の拡大率の低さから, 離散記号を全て使用することが柔軟性を高める上で効果のあることが分かった.
- (1-3)の拡大率の低さから, /c が柔軟性を高める上で重要であることが分かった. 単語レベルの文型パターンでは格要素を吸収する変数が他に無いためだと思われる.
- (2)は, 任意記号の無い辞書は柔軟性が低いことを表している. 連体詞の変数 REN が必須であるパターンの柔軟性が低いことは, 調べるまでもないが, 事例(原文)から文型パターンを作成する上では, REN などの修飾要素的な表現を任意化しない場合は, 被覆率特性からみて, 辞書の規模が2分の1から3分の1程度になることに相当することが確認された.
- (3)によると, 時制関数の任意化と, 相・様相関数の任意化は同程度の柔軟性があることが分かる.
- (1)と(3)を比べると, 離散記号のほうが, 時制・相・様相関数を任意化するよりも, 文型パターンの柔軟性を高める能力が高いことが分かった.

以上により, 離散記号を導入することが, 文型パターン辞書の柔軟性を高める上で有効であることが確認できた.

表 4.8 記述子加工の効果

対象辞書のタイプ	η_{RI}	$(1/\eta_{RI})$	η_N	$(1/\eta_N)$
(1) 全ての離散記号の無い辞書	0.05	(20.0)	0.14	(7.14)
(1-1) /y の無い辞書	0.65	(1.54)	0.81	(1.23)
(1-2) /t の無い辞書	0.96	(1.04)	0.98	(1.02)
(1-3) /c の無い辞書	0.51	(1.96)	0.50	(2.00)
(1-4) /f の無い辞書	0.77	(1.30)	0.89	(1.12)
(1-5) /k の無い辞書	0.83	(1.20)	0.83	(1.20)
(2) 任意記号の無い辞書	0.39	(2.56)	0.72	(1.39)
(3) 時制・相・様相関数を任意化した辞書	2.15	-	2.29	-
(3a) 時制関数を任意化した辞書	1.53	-	1.36	-
(3b) 相・様相関数を任意化した辞書	1.49	-	1.40	-
(3b-1) .teiru を任意化した辞書	1.19	-	1.10	-
(3b-2) .rareru を任意化した辞書	1.06	-	1.08	-
(3b-3) .dantei を任意化した辞書	1.02	-	1.02	-
(3b-4) .hitei を任意化した辞書	1.05	-	1.02	-
(3b-5) .teinei を任意化した辞書	1.03	-	1.03	-
(3b-6) .meireigo を任意化した辞書	1.05	-	1.02	-
(3b-7) .you を任意化した辞書	1.02	-	1.02	-
(3b-8) .suitei を任意化した辞書	1.07	-	1.01	-
(3b-9) .sase を任意化した辞書	1.01	-	1.05	-
(3b-10) .tekureru を任意化した辞書	1.01	-	1.02	-

4.5.3.4 η_{RI} と η_N はどちらがよいパラメータであるか?

表 4.8 によると, η_{RI} と η_N に大きな違いが見られない. ここで, 対象辞書の規模を変えながら拡大率を計測してみる. 具体的には, まず, 基準辞書から 100 パターンを抽出した小さい基準辞書を作成する. 次に, それに対して時制の任意化を行い小さい規模の対象辞書を作成する. こうして, 規模の異なる基準辞書と時制の任意化をした対象辞書の組を作ることができる. 各組において, 拡

大率を求めると表 4.9 の結果を得た. 辞書の規模が 100 ~ 500 パターンの間は, 不安定である. しかし, 1,000 パターンになると η_N が安定し始めた. η_{RI} が安定し始めるのは, 40,000 パターン以上であることから, η_N の方が小さい規模での実験で求めやすい.

時制の任意化を正しく行う上では, 手作業による分析の行程がある. 表 4.9 の結果は, 1,000 件程度の辞書を試作することで, η_N を参照することが可能となり, より大規模に手作業を行う価値があるかどうか推定可能であることを示している.

表 4.9 対象辞書の規模と拡大率の関係

対象辞書の規模	η_{RI}	η_N
100	0.01	0.15
500	0.13	0.15
1,000	4.07	1.12
5,000	1.64	1.30
10,000	1.83	1.40
20,000	2.05	1.34
40,000	1.72	1.47
80,000	1.69	1.39
122,619	1.53	1.36

4.6 意味属性制約に関する評価

「文型パターンが, 意味的に正しく適合すること」が, 非線形言語モデルの意味解析の成立に重要である. 第 4.2.4.2 節で示したとおり, 統語的な制約を文型パターンに与えること, および, 変数に意味属性制約を与えることにより, 「意味的に正しい適合」を実現しようとしている. そこで, 本節では, 変数に意味属性制約を与えることの効果について評価を行う.

4.6.1 意味属性制約の効果の評価方法

「意味的に正しい適合」ということを判断するために, パターン言語処理による日英機械翻訳の実験を行う. 適切な英訳が可能であるならば, 意味的に正しい適合をしていると判断する.

具体的には, 日本語パターンが「 A の B 」という形式の場合に, 英語表現の形式を特定するタスクである. たとえば, 「テレビの音声」という日本語表現を入力すると, 「 B on A 」という英語表現の形式(英語パターン)を出力するタスクである. ここで, 使われる知識は次のような対訳パターン対である:

$$A(962 \text{ 機械}) \text{ の } B(1087 \text{ 音声}) \Leftrightarrow B \text{ on } A$$

変数 A, B は名詞変数であり, 適合可能な名詞を表す意味属性制約が付与されている. 意味属性制約と英語パターンの組を大量に作成しておき, 英語の表現形式を決定する翻訳タスクの実験を行う.

4.6.2 意味属性制約付き「 A の B 」型名詞句パターンの作成

4.6.2.1 作成手順

まず, 英語辞典等の対訳例文 33 万件より, 日英名詞句対応の標本を半自動的に収集したところ, 約 3 万件の標本が得られた. ただし, 形式名詞, 複合名詞, 接辞付き名詞, および, 比喩表現は対象外とした.

次に、英語表現に着目して標本を分類すると、49種類の英語パターンを得た。たとえば、「空の星」の対訳として「star in the sky」があるとき、「B in A」型に分類した。分類した英語パターンの中には標本全体から見て出現度数が極めて低いものも存在する。学習データとして十分な標本が確保できないため、出現度数が10件以上である28種類を対象とする。これらは、標本全体の99%に相当する。

最後に、意味属性制約を日本語パターンに付与する。意味属性は、日本語語彙大系[13]で分類されている「一般名詞意味属性」を使用する。名詞AとBの付与された意味属性の共起関係が、パターンの意味的に正しい適合に寄与すると期待される。付与は次の2つのステップによる。

ステップ1. 学習データの作成：一般に名詞には多義があるため、1つの名詞に対して複数の意味属性が対応する。意味属性の共起を利用するために、名詞句における各名詞の意味属性を文意に合わせて絞り込む。標本のうち5,000件を学習用データとして使用する。手作業により意味属性を選択した。

ステップ2. 制約の汎化：ステップ1の標本を用い、制約を自動生成する。その方法は、一般名詞意味属性体系の階層構造を利用して、パターン間での意味属性制約の範囲が重複しないように、上位の意味属性制約に置き換えるというものである[72]。

たとえば、「店のカレー」という表現において、「店」には「374 企業」と「434 商店」という2つの意味属性があり、「カレー」には「848 飯」と「852 料理(その他)」という意味属性がある。ステップ1において、この表現に対して「434 商店」と「848 飯」に絞り込む。次にステップ2において、汎化により「430 店舗等」と「847 料理」を得る。

4.6.2.2 作成結果

作成した意味属性制約の一部を表4.10に示す。また、英語パターンごとに意味属性制約の種類数を表4.11に示す。汎化の際、兄弟関係にある意味属性制約が採用され、共通する親の意味属性制約が不採用になる場合には、1つの学習用標本から複数の制約が作成される。そのため、表4.11の件数は、5,000件を大きく上回り、合計で、440,545件となった。

表 4.10 意味属性制約と英語パターンの共起(一部)

Aの意味属性制約	Bの意味属性制約	英語パターン
408 図書館	902 出版物	B in A
1147 情報	1031 概要	A in B
388 場所	878 壁	B of A
919 帳	2583 枠	B of A

表 4.11 英語パターンごとにみた意味属性制約の種類数(一部)

英語パターン	件数	英語パターン	件数
A+B	46,315 件	A of B	6,107 件
B in A	23,894 件	A on B	1,010 件
B from A	6,946 件	B as A	597 件

※ 合計 440,545 件

4.6.3 「AのB」型名詞句の翻訳における意味属性制約の評価実験

収集した標本のうち、学習に使用しなかった標本をランダムに1,000件抽出して、翻訳実験を行

う。翻訳の評価は、「◎:標本と訳出が一致」、「○:標本と一致しないが正解」、「△:使用する文によっては正解」、「×:誤った英語形式」とする。これらの評価は、英語パターンが正しく選択できたかどうかのみに注目する。「○」の例を以下に示す:

○の例: 患者の体温

標本: the patient's temperature (英語パターン: *A's B*)

訳出: temperature of a patient (英語パターン: *B of A*)

実験結果を表 4.12 に示す。評価が◎と○の場合に明らかな正解と言えるので、60.1%において判断が正しく行われたといえる。意味属性を使用しない場合は、収集した 3 万件の中で最も割合の高かった「*A's B*」パターン(26.61%)あるいは「*B of A*」パターン(21.54%)を使用することになる。デフォルト規則として前者を使用すると、正解率が 26%程度と予想されるため、これをベースラインと考えると、意味属性制約による選択が有効であったと言える。

表 4.12 日英名詞句翻訳における意味属性制約の効果

評価	◎	○	△	×
割合	42.1%	18.0%	19.3%	20.6%

4.7 まとめ

表層レベルの言語情報から情緒推定を行うために、パターン言語処理の基礎的な考え方および実現方法について述べた。

表層レベルの言語情報からの意味処理を行うために、非線形言語モデルを導入した。非線形言語モデルは、言語表現が「表現構造を表す非線形要素」と「表現構造で結び付けられる線形要素」で構成されるとする言語モデルである。このモデルに基づく意味解析は、非線形要素と線形要素を分類する知識を使用するものである。この知識の形式として文型パターンを採用した。

パターン言語処理では、文型パターンが適合することにより、あらかじめ文型パターンに対応付けられていた情報が、意味処理結果として出力されるというものである。したがって、パターン言語処理による情緒推定を実現するためには、文型パターンの記述子の設計、表層レベルの言語情報とパターンを照合する方法、および、文型パターンに情緒の情報を対応付けることという3つの課題があることになる。本章では、文型パターンの記述子の設計、および、照合方法について説明をした。

大量の文型パターンによる言語知識ベース(文型パターン辞書)から、入力文に適合する文型パターンを検索するツールとして、ATNに基づく検索ツール *SPM* を作成した。高速化のために、(1) 絞り込み前処理、(2) join 動作、(3) サブネットワーク照合の履歴を導入した。その結果、実効的には $O(m)$ から $O(m^2)$ での照合処理が可能となった。

文とパターンを柔軟に照合するためには、文型パターンの記述子に工夫が必要である。文型パターンの柔軟性を得るために、離散記号、任意記号、時制・相・様相の関数を文型パターンの記述に導入することを提案し、その効果を実験的に評価した。ここで、文型パターンの柔軟性を定量的に評価する必要がある。そこで、文型パターン辞書の規模の拡大率 η に換算して評価する方法を提案するとともに、離散記号、任意記号、時制・相・様相の関数について、記述子を導入した効果を拡大率に基づき評価した。離散記号・任意記号が導入済みの辞書を基本とするとき、離散記号

を除去すると $\eta_{RI}=0.05$ すなわち辞書の規模が 20 分の 1 になることから、離散記号が柔軟性を高めるために非常に効果があることが定量的に確認できた。

文とパターンを意味的に正しく適合させるためには、変数に対する意味属性制約も重要である。日本語語彙大系における一般名詞意味属性を、「A の B」型の日本語名詞句パターンに制約条件として課し、対訳となる英語パターンを選択する問題について実験したところ、良好な精度であったことから、意味属性制約の効果が確認された。この実験において半自動的に意味属性制約を与えることができた点は、今後の応用に役立つと言える。

以上の処理は、規模の大きい文型パターン辞書において良好に動作することが確認された。こうして、言語表現の表層レベルの情報から意味解析を行うための、基本ツール、および、その知識ベースの設計ができた。

第5章 情緒推定用の言語知識ベースの構築

本章では、情緒生起の原因をとらえるための言語知識ベースとして「情緒属性付き文型パターン辞書」の構築を行う。そのためには、文型パターンに情緒属性を対応付けなければならないのだが、文型パターンが情緒生起の原因を表すかどうかを人手による判断で解決することにする。文型パターンの表す事態が情緒原因になるか否かについては、主観的な判断になりがちであるため、本章では、まず、その判断基準を設けることにする。次に、その結果を用いて、文型パターンに対して情緒属性の付与を行う。ここで、文型パターンの網羅性を確保するために、日本語語彙大系に収録されている全ての文型パターンを対象とする。

5.1 情緒生起の概念分析

既に、第 1.2.2 章で述べたとおり、本研究では Plutchik の理論を参考にして、次の8つの情緒を基本的な情緒として概念分析を行う：

“喜び、悲しみ、好ましい、嫌だ、驚き、期待、恐れ、怒り”

情緒の生起の原因には、ある程度の一般性が存在すると思われる。たとえば、人が怒りや喜びを持つことに対して理由を説明することができ、その理由に同意できることから、一般性が存在しないとは言い切れない。

情緒の生起にはある程度一般性が見られるということは、情緒に関して人々の間に共通の概念があるといえる。一般に、概念は日常生活の経験から形成されるので、本節では日常的なエピソードを分析に用いる。ただし、エピソードのみに頼ると重要な特徴を見落とす恐れがある。そこで、概念分析の方法は、次の2つの手順をとる：

1) ボトムアップ分析

情緒毎に関連のある事象をエピソードの中から拾い出し、当該情緒を生起させる特徴を分析する。すなわち、まず、拾い出した多くの事象を類似するものに分類し、次に、クラス毎に情緒を生起させたと考えられる特徴を抽出する。次に、個々の特徴を幾つかまとめて1つ上位の特徴を考え、さらに、それらをまとめてさらに上位の特徴を考える。これを繰り返して、注目している事象全体をカバーする最上位の特徴を求める。

2) トップダウン分析

心理学の文献などを参照し、ボトムアップ分析で欠落した特徴を補充する。

5.1.1 ボトムアップ分析

使用するエピソードは、イソップ物語を参考にして作られたシナリオから取り出す。シナリオは、日常生活に関連するエピソードが含まれている。たとえば、次のようなエピソードがある：

水汲みのエピソード

- (1) 朝起きたら、とても喉が渇いていた。
- (2) 水瓶の水を飲もうと思ったら、水瓶が空になっていた。
- (3) 住み処から小川に水を汲みに行った。
- (4) 小川で水を飲んだら、冷たくておいしかった。
- (5) 水瓶一杯に水を汲んだ。
- (6) 帰りに、もうすぐ住み処という辺りで転んだ。
- (7) 水瓶の水を半分ほどこぼしてしまった。
- (8) 悔しかったが、そのまま残りの水を持ち帰った。

これによると、(4)の事象で喜びが、(7)の事象で悲しみが各々生起している可能性がある。また、(7)に注目すると、必要な物を失い努力が無駄になったため、悔しい気持ちが生じている。このようにエピソードを観察すると情緒生起に結び付く事象やその原因となる特徴を知る手がかりが得られる。

こうして、エピソードから約 1,000 件の事象を抽出することができた。たとえば、次のような事象もある。

- 喉の渇きを潤すために、ぶどう棚のぶどうを手に入れた。
- 空腹を満たすために、川で魚を捕まえた。

これらの事象で構成されるクラスの場合、「目標を実現するために必要な物を獲得した」という特徴が見い出せる。そして、この特徴に対して簡単化のために「獲得」というラベルを与えてこれらの事象を分類することができる。また、(4)の事象を含むクラスに対しては「快」、(7)の事象を含むクラスに対しては「喪失」という特徴が見い出せ、ラベルを与えることができる。他のエピソードや事象についても同様に分類することで、「思惑どおり」、「発見」、「立案」、「完遂」、「有効」、「同意」、「優越」などの特徴が見い出せる。

得られた特徴を幾つか類似するものにまとめてみる。たとえば、「獲得」、「完遂」、「有効」は、目標を実現するための実行結果であるので「実行結果」という特徴が、「思惑どおり」、「発見」は、必要な情報を収集できたので、「情報収集」という特徴がそれぞれ見い出せる。そして、最後に「現状態が前状態よりも好都合である」という「最上位特徴」にたどり着く。

こうして、喜びの階層的な特徴がエピソードから得られる。また、他の基本情緒の特徴も同様に得られる。

5.1.2 トップダウン分析

ボトムアップ分析での事象は、児童を想定して集めたものであるため、重要な事象を見逃しているかもしれない。そこで、心理学などの知見を参考にして、トップダウン的に欠落した特徴を補う。

Murray は、情緒を特殊な種類の動機と考え、その生起には飢餓などの身体変化がともなうこと、および、学習による影響があることを述べた[73]。換言すると生理的動機と心理的動機があると言える。これを参考にして、最上位特徴を大きく「生理的」と「心理的」の2つの中位特徴に分ける。特に前者には、疲労などの内的なものと、かゆさ・暑さなどの外的なものとの顕著な区別があるので、新たに「内的な快」と「外的な快」を最下位特徴に導入する。

しかし、後者には、そのような顕著な区別が見あたらない。Pfeifer は、情緒生起の状況として目標実現の過程を重視している。また戸田によると、人間は集団を形成し、集団の統制は集団内の

順位付けによって実現される[74]. そこで, 中位の特徴として, 「目標実現」と「対人関係」を設ける. 前者は, 目標実現の過程に着目して, 「目標設定」, 「情報収集」, 「計画」, および, 「実行結果」に分ける. また, 後者は, 順位付けの有無により「優劣関係」と「仲間意識」に分ける.

以上のようにしてボトムアップ分析で得た階層的特徴を補正する.

5.1.3 特徴フレームの形成

8つの基本情緒それぞれに関し, 上述の方法で人手による分析を行った. その結果から得られた「喜び／悲しみ」における最下位特徴を表 5.1 に示す. また, 「喜び／悲しみ」と「怒り」の階層的特徴をフレームで表現して, それぞれ図 5.1, 図 5.2 に示す. なお, 図 5.1 の下線部は, トップダウン分析で補正された最下位特徴である.

得られた特徴フレームは, 125 個の最下位特徴, 51 個の中位特徴, および, 最上位特徴で構成される. 詳細は付録 2 に示す.

最上位特徴として, 「喜び／悲しみ」や「恐れ」は常識的である. しかし, 「怒り」の「規則や当然のことがらが守られていない」は, 若干考察を要した. 根源的には, 動物社会の縄張り荒らしや人間社会の掟破りに対して生起する情緒としてとらえている. 8種類のいずれの特徴フレームも, 大きく「生理的」と「心理的」に分かれ, 後者はさらに「目標実現」と「対人関係」に大別されている. これら以外のまとめ方もありえようが, 試行錯誤の結果, 最も整理された構造と判断したものである.

表 5.1 「喜び／悲しみ」の最下位特徴

最下位特徴名 (特徴ラベル)	説明
内的な快／不快	内的な生理状態が良く／悪くなった
外的な快／不快	外的な生理状態が良く／悪くなった
思惑通り／思惑違い	目標実現に有効な物事が予測した通り認識できた／できなかった
発見／見落とし	目標実現に有効で自分の知らなかった物事を認識した ／目標実現に有効な物事が行動途中に存在していたことを後で知った
判明／不明	目標実現に必要で有効な物事が, 努力の未分かった／努力の未分らない
立案／無計画	目標実現に有効な計画が生成できた／できなかった
完遂／断念	努力して計画を実行し, 目標を実現した／目標を実現できずに諦めた
獲得／喪失	目標実現に必要な物事を, 努力して手に入れた／なくした
有効／無駄	目標実現への行動結果は, 予測以上だった／効果がなかった
同意／反対	他者は自分の意見を, 受け入れた／受け入れなかった
同感／反感	他者はある事に対して自分に賛同的／反抗的な感情を持った
協力／非協力	他者は自分の目標実現のための行動を支援した／しなかった
仲直り／仲たがい	不仲になった者と再び仲良くなった／仲良だった者と不仲になった
優越／劣等	自分の能力を周囲の者より高く／低く自己評価した
賞賛／非難	自分の能力は他者に高く／低く評価された
服従／反抗	他者は要求の通りに行動した／しなかった
厚遇／冷遇	他者は自分の状態を配慮した／無視した行動をとった
保護／見放し	他者は自分を危険から守った ／他者は自分の危険を知りつつ何もしなかった

(喜び／悲しみ (現状態は前状態よりも好都合／不都合である (生理的 (内的な快／不快; 外的な快／不快); 心理的 (目標実現 (情報収集 (思惑通り／思惑違い; 発見／見落とし; 判明／不明); 計画 (立案／無計画); 実行結果 (完遂／断念; 獲得／喪失; 有効／無駄)); 対人関係 (仲間意識 (同意／反対; 同感／反感; 協力／非協力; 仲直り／仲たがい); 優劣関係 (優越／劣等; 賞賛／非難; 服従／反抗; 厚遇／冷遇; 保護／見放し))); その他)))

図 5.1 「喜び／悲しみ」の特徴フレーム

(怒り (規則や当然のことが守られていない (生理的 (内的な不快; 外的な不快); 心理的 (目標実現 (目標設定 (強制); 情報収集 (嘘); 実行結果 (失敗; 妨害; ずる)); 対人関係 (仲間意識 (反対; 反感; 仲間外れ; 脱退; 暴露; 責任逃れ; 責任転嫁); 優劣関係 (非難; 反抗; 無礼; 不平等)); その他)))

図 5.2 「怒り」の特徴フレーム

5.2 文型パターンへの情緒属性の付与

情緒生起原因に着目したタイプ1の情緒推定では、言語表現の表す事態が、情緒生起原因と一致するかどうかを判定する(第 2.2.3 節参照)。言語表現で事態を表すには、用言が中心であるので、用言についての言語知識ベースを構築する。

用言の表す事態については、日本語語彙大系における文型パターンでとらえることができる。この文型パターンは、基本的な日本語の用言 6,118 語を網羅している。そして、14,819 パターンで語義をカバーしている。

情緒推定用の言語知識ベースにするためには、パターンの1つ1つについて、情緒生起原因を表すかどうかを判定し、その結果を、「情緒属性」と呼ぶスロット・フィラー対で表し、新たに文型パターンに付与する。ここで、判定の基準として、第 5.1 節で求めた特徴フレームを使用する。付与者の判定コスト低減と付与作業の客観性向上に寄与すると期待される。

5.2.1 情緒属性

「情緒属性」のスロットとフィラーは、下記のとおりに設計する。

- 情緒名: 《喜び》, 《悲しみ》, 《好ましい》, 《嫌だ》, 《驚き》, 《期待》, 《恐れ》, 《怒り》
- 原因: 情緒生起の特徴ラベル
- 情緒主: 情緒を持つ人物. 文型パターン中の変数で記述
- 情緒対象: 情緒を向ける対象. 文型パターン中の変数で記述

「情緒名」のフィラーは、情緒タグで説明した9分類系より《なし》を除く8種類である。「原因」のフィラーは、特徴フレームで示したラベルである。「情緒主」、および、「情緒対象」のフィラーは、文型パターンにおける変数である。

具体例を図 5.3 に示す。この図には、日本語語彙大系に記載されている「用言意味属性」、「日本語文型パターン」、「その変数への意味属性制約」、および、「英語文型パターン」、それから、本研究で付与すべき「情緒属性」を示す。「貰う」には、下記の他に「(手紙を) 貰う」、「(延長を) 貰う」、「(妻を) 貰う」があるが、ここでは、対象物の制約が最も緩い概念の「貰う」のパターンである。

この文型パターンには、幾通りかの情緒生起原因が見い出せる。ここに3通りを例として示す。それぞれを「セット」と呼ぶことにする。セット(1)は、「情緒主 N1 にとって目標実現に必要な物 N2 を貰うこと」を根拠に「喜び」を推定するための知識である。セット(2)は、「贈り物を貰った」や「おしぼりを貰った」という気配りのあったことを根拠に「好ましい」を推定するための知識である。また、セット(3)は、用言意味属性を広く解釈することになるが、たとえば、貰った物が「病気」の場合である。

用言意味属性:	所有的移動								
日本語文型パターン:	N1 が N2 を N3 に/から/より 貰う								
意味属性制約:	N1 (3 主体) N2 (*) N3 (3 主体)								
英語文型パターン:	N1 get N2 from N3								
セット(1)	<table border="1"> <tr><td>情緒名:</td><td>喜び</td></tr> <tr><td>原因:</td><td>獲得</td></tr> <tr><td>情緒主:</td><td>N1</td></tr> <tr><td>情緒対象:</td><td>N2</td></tr> </table>	情緒名:	喜び	原因:	獲得	情緒主:	N1	情緒対象:	N2
情緒名:	喜び								
原因:	獲得								
情緒主:	N1								
情緒対象:	N2								
セット(2)	<table border="1"> <tr><td>情緒名:</td><td>好ましい</td></tr> <tr><td>原因:</td><td>厚遇</td></tr> <tr><td>情緒主:</td><td>N1</td></tr> <tr><td>情緒対象:</td><td>N3</td></tr> </table>	情緒名:	好ましい	原因:	厚遇	情緒主:	N1	情緒対象:	N3
情緒名:	好ましい								
原因:	厚遇								
情緒主:	N1								
情緒対象:	N3								
セット(3)	<table border="1"> <tr><td>情緒名:</td><td>悲しみ</td></tr> <tr><td>原因:</td><td>内的な不快</td></tr> <tr><td>情緒主:</td><td>N1</td></tr> <tr><td>情緒対象:</td><td>N2</td></tr> </table>	情緒名:	悲しみ	原因:	内的な不快	情緒主:	N1	情緒対象:	N2
情緒名:	悲しみ								
原因:	内的な不快								
情緒主:	N1								
情緒対象:	N2								

図 5.3 情緒属性が3セット付与された文型パターンの例

5.2.2 付与の方法

5.2.2.1 付与の条件

文型パターンに情緒属性を付与する際、以下の点に注意を払う：

- 1) **登場人物の視点：** 文型パターンにおいて人物を表す変数に対して情緒主を仮定し、その人物の立場に立つこと。人物を表す変数が複数あるときは、各変数に対して立場を考えること。変数に人物が想定できない場合は、付与を行わない。たとえば、「N1(388 場所, 2610 場)に雨が降る」は、人物が登場しないので、付与を行わない。
- 2) **事態の解釈：** パターンの表す事態の解釈は次の点に注意すること。
 - 2-1) 時制, 相, 様相が文型パターンに指定されない限り, 終止形で事態を解釈すること。
 - 2-2) 情緒主の視点で事態を解釈すること。
 - 2-3) 事態の解釈と情緒生起の特徴が一致するかどうかを判断すること。
- 3) **多義の考慮：** 情緒名, 原因, 情緒主, 情緒対象が複数通り認められるならば, それぞれについてセットを作ること。1つのセットに絞り込むことはしない。

情緒主の視点を意識しながら付与を行うことは、原因となる事態を見逃さないための付与作業者の重要な心がけである。たとえば、「N1 が N2 を N3 から略奪する」という結合価パターンにおいて「略奪する」という言葉のイメージに頼ると「略奪された N3 の情緒《怒り》」だけに気をとられやすい。しかし、人物の視点を変数毎に置きながら事態を想像することにより、「略奪に成功した N1 の獲得による《喜び》」に気がつく。

情緒生起の原因をあらかじめ定義しているため、文型パターンが情緒生起の原因を表すかどうかを客観的に判断できる。ここで、パターンの表す事態が、情緒生起の原因を表すと主観的に思われるのだが、最下位特徴のラベルのいずれとも合わないことがある。その場合、2つの対処法がある。1つは中位特徴や最上位特徴を対応付ける方法である。これは、用言の語義が抽象的である場合に用いる。もう1つは〈その他〉を対応付ける方法である。これは、何かしら具体性のある用言であるが最下位特徴には適さない場合に用いる。

5.2.2.2 付与の行程

付与作業は、人手で行う。人手の作業において、作業漏れは完全に排除できるものではない。そこで、可能な限り作業漏れを排除するために、3段階から成る行程を設ける。

第1版) 1人の作業者が日本語語彙大系の構文体系の先頭のパターンから順に、情緒属性の対応付けを行う。

第2版) 再び同じ作業者が、原因の区別に注意しながら、見直しを行う。

第3版) 上記の作業者とは異なる者(検査者)が検査して、対応付けに疑問のある箇所や追加案を作業者に示し、両者の協議により修正する。

ここで、作業者と検査者のトレーニングについて説明する。両者とも、特徴フレームの学習として、特徴ラベル、および、具体的な定義(付録2参照)を学習した。

作業者は、第1版の構築において、日本語語彙大系の先頭のパターンから順に分析を進める。なお、週に1回のペースで対応付けの結果からランダムサンプリングにより、正しく付与できているかどうかを情緒原因を熟知した者が検査した。特徴フレームに対する作業者の理解を確認しながら作業を進めた。第2版と第3版も同様に進捗確認を行いながら作業を進めた。

第3版の作成における検査者は、第 3.2.2 節で示した対話コーパスへの情緒タグ付与の経験者

であり、他者とのタグの一致率の高かった者である。検査者は、3名である。約 1.4 万パターンを3等分し、各検査者を割り当てた。1人あたり約 6,000 セットを検査するのだが、約 500 セットごとに作業者に検査結果を報告し、協議をした。協議内容は、情緒の原因と用言の語義が対応したと判断した理由である。

5.2.3 第2版から第3版への改訂の様子

第2版を改訂して第3版を作るという作業工程において、検査作業があった。ここで、検査者が指摘した箇所は、2,863 カ所であった。作業者が検査者の指摘を妥当であると認めた箇所は、2,494 箇所であった。修正に応じた箇所は作業者の見落としした箇所であったので、異論なく修正することができた。しかし、369 カ所は修正に応じなかった。

修正に応じなかった箇所のうち 307 カ所は、用言意味属性が「5 属性」であるパターン、すなわち、性状規定の表現であった。たとえば、「N1 は金遣いが荒い」というパターンがある。「『金遣いが荒い』という属性を N1 が有する」と話し手が信じていることが表される表現である。しかし、情緒主とする N1 が、この属性を認めているかどうかは、この表現だけでは判断できないため、N1 に情緒生起原因が成立していることは確認できない。したがって、原因を表す表現とは分類しなかった。ただし、N1 が話し手の場合、すなわち、N1 に1人称の代名詞や自身の名前が対応する場合は、情緒生起原因が成立する。その意味では検査者の指摘は正しい。今後、「5 属性」に属するパターンに対する特例を設けて、知識ベースを改良することが課題として挙げられる。

以上の検査によると、対応付け作業に対する信頼性を推察することができる。上述の検査作業によると、17,587 カ所が検査対象となっていたところ、2,494 カ所に修正が発生したので、第2版で問題のあった割合は、14.2% (2,494/17,587)である。第3版は指摘箇所が全て修正されているため、より信頼性が高いと推察される。

5.2.4 付与の結果

対応付けの結果を以下にまとめる：

- **日数：** 対応付け作業に要した時間(学習を除く実質的な付与作業の時間)は、1つの版に対して約3ヶ月であった。
- **第3版の内訳：** 詳細については次節で説明する。ここでは簡単に内訳を説明する。情緒属性のセットが対応付けられたものは、「情緒主が登場し、かつ、情緒生起の原因を表すもの」であり、対応付けられなかったものは、情緒主が登場しない(たとえば、「N1 が壊れる」)、または、情緒生起の原因を表さないパターンである。用言数、パターン数、および、セット数は次のようになる：

対応付けのあった用言の数と割合：	45.7% (2,798 個/6,118 個)
対応付けのあったパターンの数と割合：	52.2% (7,739 パターン/14,819 パターン)
対応付けたセットの数：	11,712 セット

5.3 構築した知識ベースの内訳

知識ベースの性質として、情緒生起原因を表す用言の多様性、および、用言から推定される情緒の曖昧性という観点から、本知識ベースの内訳を示す。

多様性とは、1つの生起特徴に対して表すことのできる用言の数である。たとえば、《喜び》の生起特徴として〈獲得〉があるが、〈獲得〉を表す用言は、「貰う」、「入手する」、「拾う」、「買う」などが数

多く存在する。日本語の言葉の豊かさを表す1つの指標となりえるもので、資料価値がある。一般のシソーラスにおいては、語義を区別する観点で分布が示されているが、本分析においては、情緒生起原因という観点から分析する。

一方、曖昧性とは、1つの用言の表す事態が原因となって結び付く情緒の数である。たとえば、「貰う」という用言は、前節において、3つの情緒属性のセットが対応付けられたので、曖昧性が3である。情緒推定においては、曖昧性解消が1つの課題であるため、用言の単位での曖昧性や、語義解析後の曖昧性など、曖昧性解消のための処理の深さを比較しながら曖昧性の程度を分析する価値がある。

5.3.1 情緒生起原因を表す用言の多様性

5.3.1.1 情緒名ごとの分布

14,819 パターンに対して、11,712 セットが対応付けられた。情緒名ごとにセットの数を数えると、表 5.2 の分布を得た。《好ましい》が最も多く、《驚き》が最も少なく付与された。もし、8種類の情緒が均等に付与されると、12.5%ずつとなるので、それを基準にすると、《恐れ》、《怒り》、《驚き》の原因を表す用言は種類が少ないことがわかる。その理由として次のことが考えられる。

- 《怒り》や《驚き》の数が少ないのは「規則や当然のこと」や「予測どおりでないこと／予測外のこと」ということを1つの用言で表しにくいことであると考えられる。
- 《恐れ》の最上位特徴を構成するキーワードとして「危害」が含まれているが、これは「不都合」の下位概念である。したがって、《嫌だ》や《悲しみ》よりは該当する数が少なくなることは当然のことである。
- 《好ましい》と《喜び》、それから、《嫌だ》と《悲しみ》は最上位特徴が類似しているが、《喜び》と《悲しみ》は、前状態と現状態の比較が考慮されるので、《好ましい》と《嫌だ》よりは該当する数が少なくなる。

表 5.2 文型パターンに対応付けられた情緒属性のセット数の情緒名ごとの分布

好ましい	喜び	嫌だ	期待	悲しみ	恐れ	怒り	驚き
2,642	2,295	1,907	1,818	1,355	1,105	506	84
22.6%	19.6%	16.3%	15.5%	11.6%	9.4%	4.3%	0.7%

※ 下段の比率は、総セット数 11,712 個に対する比

次に、情緒に対応する用言の例を表 5.3 に示す。格要素なしに語義がとらえにくい場合については括弧書きにより名詞の種類を補足する。用言に後続する括弧書きは相・時制・様相に関する制約である。格要素が固定的な場合は括弧書きなしに格要素を記載した。この表より次のことが言える：

- 《喜び》に関して、「惨敗する」は、「N1 が N2 に惨敗する」というパターンであり、N2 を情緒主とした場合である。逆に情緒主を N1 とした場合は《悲しみ》が付与される。
- 《驚き》に関しては、もともと副詞が含まれているパターン(例. ばったり遭う)や、偶然性を表す名詞が格要素に指定されているパターン(例. 不意を打つ)が目立つ。また、作業者の主観に依存して偶然性の決められたパターンがある。たとえば、「N1 (3 主体) が N2 (2364 雨) に遭う」は、作業者が「雨」に対して偶然性があると判断している。同様に「喧嘩を買

う」,「発掘する」,「目撃する」は事象に対する偶然性により判断された。

- 《期待》に関しては,「他者への依頼(例. 仰ぐ, 心当たりを当たる)」,「他者と連絡(例. のろしを上げる)」,「ある事象の起こる前に行く動作(例. 幕を開ける)」が見られる。これらの用言は,その結果に新たな事態が予測され,《期待》の最上位特徴と一致する。
- 《恐れ》に関しては,用言のみで身体的・社会的な危害を表すものと,格要素が危害を表すものが見られる。

表 5.3 情緒ごとに見た対応する用言の例

情緒名	用言の例
《喜び》	愛護する, (条件に) 合う, 軍配が上がる, (候補に) 挙げる, 会得する, 援護する, 知恵を借りる, 手を借りる, 胸に応える, ご馳走する, 授かる, 惨敗する, 飼育する, 証明する, 助言する, 尽力する, 着服する, 満腹に成る, …
《悲しみ》	盗難に遭う, 喘ぐ, 頭が上がる (ない), 誤る, 争う, 焼きを入れる, 恨みを受ける, 失う, 後ろ向きだ, 目を疑う, 訴える, 都を落ちる, 力を落とす, 思い切る, 夢にも思う (ない), 折り合う, 抗議する, 降参する, ぶつかる, 見送る, …
《好ましい》	会う, 呼吸が合う, 水を開ける, (部屋/家屋に) 上げる, (印象を) 与える, 集まる, シャワーを浴びる, 安住する, 安心する, 気を利かせる, 口を聞く, 協和する, (人/職業/労働で) 食う, 駆使する, 決定を下す, 訓示する, …
《嫌だ》	喘ぐ, 呆れる, (授業で) 当たる, 思案に余る, 怪しむ, 溝がある, 顔色をうかがう, 支配を受ける, 深手を負う, (身体障害/病気類を) 起こす, 気が置ける (ない), (敵・味方を) 抑える, (支払い/料金を) 溜める, 墮落する, …
《驚き》	(雨に) 遭う, ぱったり遭う, 手の内を明かす, (リストに) 上がる, 不意を打つ, (目/口を) 押さえる, 思いがけない, 喧嘩を買う, 裏を搔く, 不意を食う, 身につまされる, 発掘する, 舌を巻く, 見つかる, 目を見張る, 目撃する, …
《期待》	(資材を海外に) 仰ぐ, (薬品を) 仰ぐ, (立案を) 明かす, 幕を開ける, 例に挙げる, のろしを上げる, (心当たりを) 当たる, (関心/同情/話題が) 集まる, 依頼する, 電話を入れる, 打ち明ける, 夢に描く, 演説する, お参りする, …
《恐れ》	毒杯を仰ぐ, 脅威を与える, (人間活動に) 当たる, (強豪に) 当たる, 違反する, 慰留する, 隠蔽する, 迂回する, 疎い, うろたえる, 追い込む, 覚悟する, 疑念を掛ける, 偽証する, (力/知恵/背丈を) 比べる, (名称を) 汚す, …
《怒り》	煽る, 欺く, 言い返す, 喧嘩を売る, 横領する, べてんに掛ける, 強いる, しくじる, 詐欺をする, 作り話をする, だます, 挑戦する, 不幸に付け込む, ミスに付け込む, 手出しする, 反論する, 仕事を回す, どじを踏む, 笑う, …

5.3.1.2 生起特徴ごとの分布

生起特徴の階層を見ながら, 対応付けられた情緒属性セットの分布を調べた。その結果を表 5.4 に示す。最上位特徴を対応付けた数は「H」の行に, 中位特徴を対応付けた数は「M」の行に, 最下位特徴を対応付けた数は「L」の行にそれぞれ記した。斜線部は, 特徴フレームに該当する特徴ラベルが定義されていない部分である。内訳を以下に説明する:

- 11,712 セットがパターンに対応付けられ, 最下位特徴の対応付けは 10,929 セットであった。その割合は 93.3% である。逆に, 最上位特徴, 中位特徴, その他の特徴の対応付けは 6.7% である。最下位特徴に著しい不足の無いことが確認された。
- 最上位特徴に該当する用言は, 「危うい, 危険だ(恐れ)」, 「好都合だ(好ましい)」である。最上位特徴の定義をそのまま表している。しかし, たとえば, 「予想外だ」は, 情緒主がパターン上に登場しないため, 《驚き》に対応付けされていない。この場合は対応付けの対象

外であり、誤りではない。

- 中位特徴〈生理的〉に対応付けられた用言には、「全快する、治す、吹き返す、若返る」などがある。下位の特徴には〈内的な快〉、〈外的な快〉があるが、両方に共通する語義であるため、中位特徴に対応付けられた。これは中位特徴のラベルを正しく用いた例である。
- 中位特徴〈情報収集〉は、《期待》においては、「探りを入れる、(事件/問題に)メスを入れる、インタビューする、(流行/跡/変動を)追う」という、『情報を収集する行為を開始する時点の事象』を表す表現であった。これらは『調査』に関連する用言であるが、《期待》の最下位特徴には該当する特徴ラベルが無かった。「期待の特徴フレーム」に追加する必要がある。
- 中位特徴〈対人関係〉に対応付けられた用言には、「惚れる(期待)、熱愛する(期待)」という用言意味属性〔感情動作〕の用言が含まれていた。「惚れる」の直接表現は「好ましい」であるが、「惚れた後」には対人関係に関して様々な事象の発生が予測されることから《期待》と判断した。深く考えて対応付け作業が行われていたことが伺える。
- 中位特徴〈仲間意識〉は、《期待》においては、「気がある、縁が薄い、交際する、手を出す」という『仲間の関係を形成する前の時点の事態』を表す表現であった。これは中位特徴のラベルを正しく用いている。
- 〈その他〉に関して、564 セットが対応付けられた。幾つか特徴的なものを説明する。
 - 《好ましい》に関しては、「ドライブする、登山する、保管する、水上スキーをする、(たばこを)吸う」などがあり、個人の嗜好に依存している。これらは、逆に個人に依存して《嫌だ》と分類されても誤りではない。これらのように個人に依存するものは、一般性のある部分の知識ベース化の目的からは外れるものであり、別途識別のための属性値を付与しなければならない。
 - 《嫌だ》に関しては、「(授業で)当たる、言い訳が立つ、遅刻する、年を取る、血が滲む、舌が回る(ない)」などであり、これらは平均的な人々が嫌だと感じる原因として妥当であると考えられる。今後、事例をさらに収集し、最下位特徴を設計し追加する必要がある。
 - 《恐れ》に関しては、「隠れる、避ける、忍ばせる」などであり、隠れたり避けたりすることの生じた原因は《恐れ》であるため、これらの分類は直接的には「反応」の事象である。しかし、これらの事象は《恐れ》の解消にはなっておらず、「見つかること」や「避けきれないこと」のように2次的な《恐れ》の原因になることがある。ゆえに誤りではない。

5.3.2 用言から推定される情緒の曖昧性

5.3.2.1 用言意味属性ごとに見た情緒の曖昧性

情緒属性の全セットから、用言意味属性ごとに、対応した情緒の数を収集した結果を表 5.5 にまとめる。この表は、用言意味属性ごとに用言を集めて、その用言に対応付けられた情緒の数の比をまとめている。たとえば、「2 状態」の用言意味属性を有する用言は、25 個存在し、そのうちの 24% (6個)に《好ましい》が対応付けられていることを表している。1つの用言に複数の情緒属性が対応付けられることがあるので、割合の合計は 100%を超える。たとえば、「3 抽象的關係」を表す用言は1つであるが、《好ましい》と《嫌だ》が対応付けられているため、それぞれ 100%ずつとなっている。空欄は対応する情緒が全く無いことを、0%は四捨五入した確率値が 1%未満であることをそれぞれ示している。この表より次のことが言える：

表 5.4 生起特徴ごとに見た情緒属性セットの分布

		喜び	悲しみ	好ましい	嫌だ	驚き	期待	恐れ	怒り
最上位	H	0	0	1	0	0	0	2	0
生理的	M	10	5	42	37	0	10	7	0
	L	13	68	159	308	1	20	57	10
心理的	M	2	10	0	1	0	0	2	1
目標実現	M	0	0	0	0	0	0	0	1
目標設定	M			0	0				0
	L			454	561				95
情報収集	M	0	0	0	0	0	6	2	0
	L	158	57	294	67	59	276	283	13
計画	M	0	0				0	0	
	L	113	134				432	219	
実行結果	M	0	2	3	1	0		0	0
	L	1,149	488	642	250	13		4	126
対人関係	M	0	0	0	0		0	1	3
仲間意識	M	3	0	14	17		31	5	0
	L	212	122	548	349		471	299	156
優劣関係	M	0	0	0	0		0	0	0
	L	614	431	362	232		336	192	82
その他		21	38	123	84	11	236	32	19

- [6 所有], [19 所有的移動]の用言意味属性を持つ用言について
 - 《喜び》, 《好ましい》の対応付けられたものが多い.
 - [19 所有的移動]において《悲しみ》が対応付けられているのは、「失う者」の立場で解釈する場合である. 情緒主の区別が重要であると言える.
 - 《期待》が対応付けられているのは、「(資材を海外に)仰ぐ」のように未来において[19 所有的移動]の成立する用言に対してである. 用言の連鎖する事態に関する区別が重要であると言える.
- [27 消滅・破壊]の用言意味属性を持つ用言について
 - 《悲しみ》の対応付けられた物が多い.
 - 「M1 が壊れる」のように情緒主の登場しないパターンがあり, [27 消滅・破壊]の用言意味属性であっても〈喪失〉による《悲しみ》が対応付けられなかったものがある(これは対象外としているので誤りではない).
 - 《喜び》が対応付けられているのは、「撲滅する」のように、「努力して無くすること」という解釈の可能な用言である. 語彙分析の際, 用言の暗示する動作者の意図や努力に注意を払う必要がある.
- [20 属性変化]はとりたてて確率の高い情緒はない. [6 所有], [19 所有的移動], [27 消滅・破壊]は最下位特徴と関わるのが容易に想像できるが, [20 属性変化]は関わる生起特徴が想像しにくい. [20 属性変化]は該当する用言が多いので, 情緒推定に用言意味属性を使用するためには, 下位分類を設ける必要がある.
- [33 使役]は, 「促す, 恩を着せる, 強いる」などの用言であり, 《嫌だ》, 《期待》, 《怒り》の対応付けが多い. [34 可能], [35 開始], [36 終了]も対応付けられる情緒の確率に不自然さがない.

以上のことから、用言意味属性を用いて情緒を推定することは、判断の情報が不足していると言える。「情緒主」や「用言の連鎖や暗示する事態」などによる条件を平行して用いる必要のあることが明確になった。このことは用言意味属性の分解能が荒いことを示している。基本的な言語解析において用言意味属性だけではカバーできないことを示している。

表 5.5 用言意味属性から推定される情緒の曖昧性

用言意味属性	延べ用言数	喜び	悲しみ	好ましい	嫌だ	驚き	期待	恐れ	怒り	なし
1 事象	0									
2 状態	25	8%	4%	24%	8%					64%
3 抽象的關係	1			100%	100%					
4 存在	61	2%	2%	18%	7%					77%
5 属性	3,783	4%	4%	7%	8%	0%	2%	3%	1%	77%
6 所有	78	23%		45%	1%		9%	4%		37%
7 相對關係	363	31%	27%	21%	19%	0%	10%	8%	8%	23%
8 因果關係	48	6%	2%	10%	2%		4%	2%		83%
9 精神的關係	2			50%						50%
10 知覚状態	60		3%	22%	55%			2%		23%
11 感情状態	275	5%	25%	17%	20%	1%	17%	11%	3%	17%
12 思考状態	25	16%	8%	12%	16%		20%	16%	8%	24%
13 心的状態	9	22%	67%		11%	22%		33%		11%
14 身体状態	54		7%	17%	39%		4%	7%		31%
15 自然現象	155			1%	4%			3%	1%	95%
16 行動	159	18%	3%	16%	10%		30%	23%	6%	25%
17 物理的行動	1							100%		
18 物理的移動	511	8%	6%	19%	11%		12%	9%	1%	59%
19 所有的移動	409	63%	13%	32%	9%		19%	3%	6%	11%
20 属性変化	2,588	14%	9%	17%	12%	0%	6%	6%	3%	52%
21 身体変化	373	10%	13%	12%	30%		3%	6%	0%	36%
22 結果	269	29%	26%	9%	18%	6%	2%	15%	3%	23%
23 身体動作	1,811	12%	6%	17%	13%	0%	10%	8%	4%	50%
24 利用	121	18%	13%	20%	6%		17%	7%	8%	35%
25 結合動作	183	20%	1%	23%	5%		10%	2%	1%	51%
26 生成	276	24%	0%	13%	1%	1%	9%	3%		56%
27 消滅・破壊	96	22%	39%	2%	8%			6%	1%	41%
28 精神的行動	0									
29 精神的移動	1,023	14%	4%	25%	17%	1%	42%	8%	6%	22%
30 知覚動作	194	10%	2%	30%	20%	2%	16%	18%	1%	32%
31 感情動作	1,341	20%	18%	21%	18%	1%	17%	15%	8%	13%
32 思考動作	1,374	24%	8%	20%	7%	1%	25%	10%	2%	27%
33 使役	16	6%	6%	6%	44%		63%	13%	44%	6%
34 可能	11	27%		36%						36%
35 開始	99	19%	2%	3%	1%	1%	31%	12%	3%	48%
36 終了	100	24%	28%	1%	4%				1%	52%

5.3.2.2 用言ごとに見た情緒の曖昧性

用言ごとに対応付けられた情緒の種類数より、曖昧性を調べた結果を表 5.6 にまとめ、曖昧性の高い用言を表 5.7 に例示する。情緒《なし》を含むため、曖昧さは 9 となる。日本語語彙大系の全用言 6,118 個に対して、情緒《なし》も含めて曖昧性のない用言は、64.5%であり、曖昧性のある用

言は、35.5%であった。全体的に見れば曖昧性のある用言の割合は低い。しかし、曖昧性の高い用言をみてみると、和語の動詞であった(表 5.7)。和語動詞は使用頻度が高いため、実際の情緒推定においては曖昧性の問題が無視できない。

表 5.6 用言に対応付けられた情緒の曖昧性

情緒の曖昧さ (《なし》を含む情緒の種類数)	割合	該当する用言 (該当数 / 総用言数)
9	0.0%	(3 / 6,118)
8	0.3%	(16 / 6,118)
7	0.3%	(21 / 6,118)
6	0.5%	(32 / 6,118)
5	1.1%	(70 / 6,118)
4	3.0%	(184 / 6,118)
3	8.2%	(504 / 6,118)
2	22.0%	(1,343 / 6,118)
1	64.5%	(3,945 / 6,118)

表 5.7 曖昧性の高い用言の例

情緒の曖昧さ	用言の例
9	掛ける, 打つ, 付く
8	する, 起こす, 見る, 使う, 取る, 受ける, 出す, 出る, 上がる, 成る, 当たる, 踏む, 突く, 入れる, 付ける, 立つ
7	ある, 引く, 飲む, 下す, 開ける, 掛かる, 遣る, 言う, 守る, 上げる, 乗る, 折る, 揺く, 置く, 張る, 転じる, 当てる, 抜く, 変える, 落とす, 濡らす
6	逸らす, 引っ張る, 押す, 外す, 呼ぶ, 固める, 絞る, 行く, 差す, 止める, 示す, 収容する, 集める, 食う, 振る, 切る, 接する, 潰す, 吐く, 塗る, 買う, 払う, 変わる, 返す, 報いる, 無い, 明かす, 与える, 利かせる, 立てる, 流す, 裂く

5.3.2.3 文型パターンと情緒主の単位に見た情緒の曖昧性

語義の区別と情緒主の区別が、情緒の曖昧性解消にどの程度効果があるかを調べた結果を、表 5.8 にまとめる。表内の数値は、用言の数である。“語義のみ区別”の列は、1つの用言に複数のパターンが存在するとき、パターンごとに対応付けられた情緒の種類数の最大値を、その用言の「情緒の曖昧さ」とした。“情緒主のみ区別”の列も同様に、1つの用言において複数の情緒主が存在するとき、情緒主ごとの情緒数を用いた。和語と漢語により多義の度合いが異なるため、区別してまとめた。語義や情緒主の区別をしない場合(“区別なし”の列)は表 5.6 と同一である。表 5.8 より次のことが分かる:

- 和語は、語義の区別が情緒の曖昧性解消に効果があり、情緒主の区別はそれほどではない。
- 漢語は、情緒主の区別のほうが、語義の区別よりも効果が若干高い。
- 語義の区別と情緒主の区別は両方を行うとより効果がある。
- 両方の区別においても曖昧さが残された部分は、情緒主と動作対象の関係などにより曖昧性解消を行う必要がある。その割合は、20.6% (1,262 / 6,118) である。

表 5.8 語義の区別と情緒主の区別による曖昧性解消

情緒の 曖昧さ	区別なし		語義のみ区別		情緒主のみ区別		両方を区別	
	和語	漢語	和語	漢語	和語	漢語	和語	漢語
9	3	0	0	0	0	0	0	0
8	16	0	0	0	1	0	0	0
7	21	0	0	0	8	0	0	0
6	30	2	0	0	15	0	0	0
5	64	6	0	2	18	1	0	0
4	143	41	27	21	50	17	6	13
3	309	195	150	142	129	53	45	49
2	624	719	574	611	534	597	565	584
1	1,976	1,969	2,435	2,156	2,431	2,264	2,570	2,286

5.3.2.4 対極的な曖昧性

情緒推定の曖昧性が問題となるのは、推定された情緒を《 Positive 》, 《 Negative 》, 《なし》の3分類に大別したとき、対極となる《 Positive 》と《 Negative 》の両方となる場合である。そこで、3分類系に情緒属性を変換し、対極となる曖昧性がどの程度存在するのかを調べる。

表 5.9 に、語義の区別と情緒主の区別をしない場合とした場合について、対極的な曖昧性のある用言がどの程度存在するのかをまとめた。この表より次のことが言える：

- 語義・情緒主の区別により、対極的でない曖昧性が大きく解消された。しかし、対極的な曖昧性のある用言は、それだけでは解消されない。
- 語義・情緒主の区別した用言の 12.2%(748 / 6,118)が対極的な曖昧性を持つ。該当する用言の例を表 5.10 に示す。用言意味属性ごとにみて該当パターン数の多い順に例示する。このような用言は、特に、情緒主と情緒対象の関係に対する判定条件が必要である。

以上より、情緒推定において対極的な情緒の曖昧性について、語義、および、情緒主を区別しても解消できない用言は、日本語語彙大系の用言の中で、12.2% であることが分かった。

表 5.9 対極的な曖昧性のある用言の数

語義・情緒主の区別の有無	曖昧性がない	対極的でない曖昧性がある	対極的な曖昧性がある
なし	4,359	640	1,119
あり	5,341	29	748

表 5.10 対極的な情緒の曖昧性のある用言(一部)

パターン数	用言意味属性	用言の例
259	属性変化	魔法に掛ける, (組織へ) 乗り出す, 脱却する, 任命する, 鍛える, 転職する, 増減する, 火に油を注ぐ, 壊す, …
178	身体動作	出社する, 逃げる, (たばこ/薬品を) 吸う, 随行する, 歩く, 登山する, (怪我を) 浸す, 口にする, 伸ばす, 処刑する, …
166	思考動作	脳味噌を絞る, 思案する, 取り組む, 熟慮する, 選出する, 認める, 例に挙げる, 探る, 腹を決める, 自粛する, …
154	精神的移動	曝す, 物語る, 頭を突っ込む, (反対を) 叫ぶ, (命令が) 下る, 説く, インタビューする, 授業する, 伝染する, …
135	属性	休学する, 遠ざかる, 代表する, 目立つ, 相殺する, 一変する, 堅持する, 酷似する, (願望・失望と) 合う, …
108	感情動作	失恋する, 気が高ぶる, 追従する, 気に掛かる, 差別する, カンニングする, 引く, 過保護にする, 胸を膨らむ(せる), …
106	物理的移動	転じる, 出る(ていく), 吹き払う, (地域・陸地に) 渡る, 遠征する, 送還する, 肉薄する, 移住する, 移民する, …
97	相対関係	水を開ける, 顔を突き合わせる, (スポーツで) タイに成る, 対決する, 嗜好が違う, 統率する, 相当だ, 仕える, …

※ 用言に対して1個以上のパターン・意味属性が存在するため, 本表のパターン数の合計は, 748 を超える。

5.4 まとめ

本章では, パターン言語処理に基づく情緒推定に必要な言語知識ベースを構築した. この情緒推定は, 情緒生起原因を手がかりとしたタイプ1の方式である. 情緒推定を実現するための言語知識ベースを構築する問題に対して, (1) 情緒生起原因を列挙すること, (2) 列挙された情緒生起原因を判断基準として, 網羅的な文型パターンに対応付けることが課題であった.

(1)に対して, Plutchik を参考にした8つの基本情緒に対して情緒的なエピソードの分析により, 情緒生起原因の種類をボトムアップ的に列挙し, その後, 心理学などの知見に基づきトップダウン的に特徴の追加と整理を行った. こうして, 約 120 種類の原因を列挙することができた.

(2)に対して, 日本語語彙大系における用言パターンを参照し, 日本語における基本的な用言を網羅した文型パターン 14,819 件に対して, (1)で得た情緒生起原因を判断基準としながら情緒属性の対応付けを行った. 情緒属性として, 「情緒の種類」, 「情緒の原因」, 「情緒主」, 「情緒対象」を設けた. その結果, 14,819 件のパターンうち 52.2%は情緒原因の対応付けがあることが分かった. さらに, 情緒生起原因を表す用言の多様性と用言から推定される情緒の曖昧性について, 本知識ベースの内訳を示した.

第6章 インプリメンテーション

本章では、パターン言語処理に基づく情緒推定の方式を実装する。第4章では、パターン言語処理の方法とパターンの記述子を示した。第5章では、情緒推定のための文型パターン型知識ベースを構築した。本章では、両者を組み合わせて、本提案方式を実装する。ここで、第5章で示した文型パターンは、記述子が単純であり、柔軟性に欠ける。そこで、文型パターンの変形を行う。また、複数の文型パターンが適合するというパターン言語処理での曖昧性が問題になる。そこで、ヒューリスティクスによる曖昧性解消を行う。こうして実装した後、動作の確認を行う。

6.1 パターン言語処理に基づく情緒推定の構成

パターン言語処理に基づく情緒推定のシステム構成を図 6.1 に示す。全体的にみると、表層構造の解析、構造照合による文型パターンの検索、および、適合結果の選択という3段階がある。

表層構造の解析では、まず、入力された言語表現の形態素解析を行う。品詞コード、単語の基本形、および、文節境界を解析する。次に、意味属性コードを付与する。ただし、両処理において、品詞や意味属性の多義性を解消する必要はない。

構造照合による文型パターンの検索は、第4章で示した検索プログラム *SPM* を使用する。*SPM* は、文に適合した全てのパターンについて適合結果を出力し、かつ、1パターンにつき複数の適合の仕方があるときは、その全ての組み合わせについて出力する。

適合結果の選択は、2つの判断用の情報を追加した後に行う。1つめの追加とは、意味属性制約の充足状況の追加である。*SPM* は意味属性情報を検査しないので、この段階で制約が満たされているか否かのフラグを立てる。2つめの追加とは、情緒属性の追加である。適合した文型パターンに対応付けられている情緒属性を、情緒属性付き文型パターン辞書より検索する。こうして情緒の候補が作られる。次に、適合結果のソートにおいて、意味属性制約、情緒属性、および、適合の仕方から、特徴量を算出し、適合結果の順位付けを行う。

タイプ1の情緒推定では、文脈情報を参照しないので、適合結果の中で順位の最も良いものを推定結果として出力する。将来、タイプ3の情緒推定として文脈情報を参照する際は、文脈上得られる情報を特徴量に加えて、適合結果のソートの前に行う。

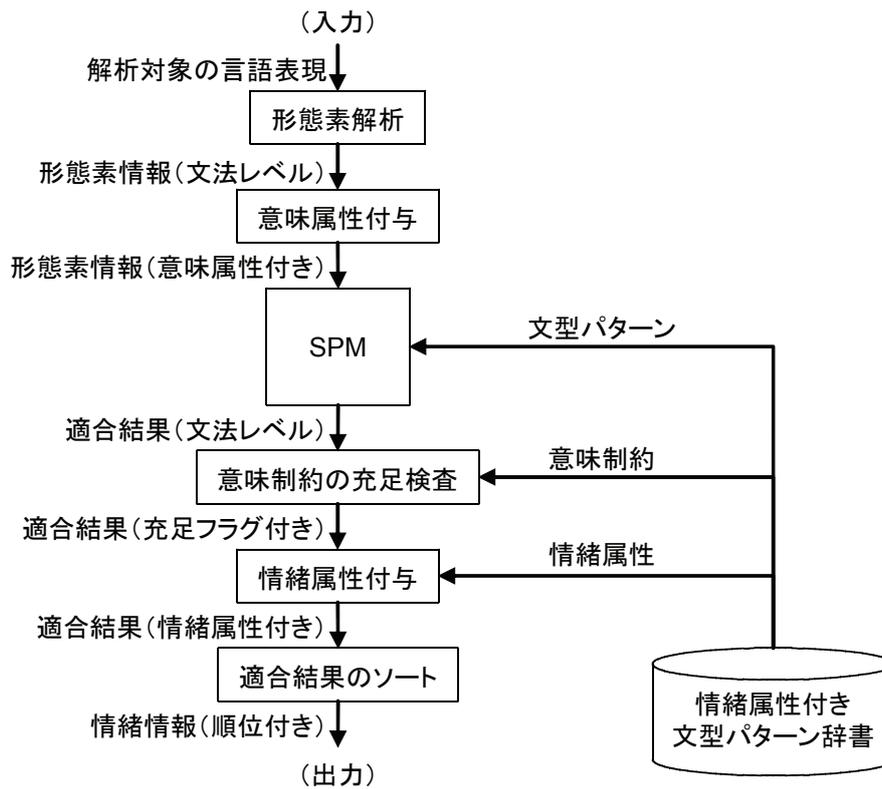


図 6.1 原因検査型タイプ1情緒推定システムの構成

6.2 文型パターンの変形

6.2.1 日本語語彙大系の文型パターンの構造

第5章では、日本語語彙大系の文型パターンに対して情緒属性を付与した。この文型パターンは、記述子として、変数、字面、および、選択記号が使用されている。その構造は次のようになっている:

```

〈日本語語彙大系文型パターン〉 ::= 〈見なし格要素列〉 〈述語〉 [ 〈後続表現〉 ]
〈見なし格要素列〉 ::= 〈見なし格要素〉 〈スペース〉 [ 〈見なし格要素列〉 ]
〈見なし格要素〉 ::= [ 〈連体修飾型ターム〉 ] 〈非連体修飾型ターム〉
〈連体修飾型ターム〉 ::= 〈名詞変数〉 〈連体修飾の格助詞および相当表現〉
| 〈連体修飾型の字面〉
〈非連体修飾型ターム〉 ::= 〈名詞変数〉 〈非連体修飾の格助詞および相当表現〉
| 〈非連体修飾型の字面〉
| 〈副詞〉

```

以下に具体例を示す:

例 6.1 N1 が N2 から/より N3 に/へ/まで 旅する

この日本語パターンは、3つの見なし格要素と1つの述語で構成されている。格助詞に対応する部分には、「/」という選択記号が使われている。他のパターンの例を幾つか示す:

例 6.2 N1 が N2 に/の後ろから 付く (ていく)

例 6.3 N1 は/の 心が 暖かい

例 6.4 N1 が N2 の 記憶に 新しい

例 6.5 N1 が N2 に ぼったり 遭う

例 6.2 の2つめのタームは、「N2 に」または「N2 の後ろから」と解釈する。また、「(ていく)」は、後続表現であり、述語とあわせて、「付いていく」という表現と適合する。例 6.3 は、1つめのタームに連体修飾の格助詞があるため、このパターンの見なし格要素は、1つだけである。例 6.4 では、「N2 の 記憶に」が1つの見なし格要素である。例 6.5 には、副詞の見なし格要素がある。

6.2.2 記述子の変形の規則

日本語語彙大系の文型パターンは柔軟性が低い。それは、入力文が構文解析されることを前提としており、その構文情報とパターンを照合する設計のためだと考えられる。すなわち、入力文の側の柔軟性を高めている。本研究では、入力文の解析はあえて形態素のレベルに留めておき、パターンの側の柔軟性を高める方法をとる。

日本語語彙大系の文型パターンを、次の基準に従って変形する。

- パターンの先頭に、接続詞の離散記号「/cnj」と連用節の離散記号「/y」を付与する。
- タームの直前に、離散記号「/tk」を付与する。
- 変数で構成されるタームは、任意化する。(例) N1 は → [N1 は]
- タームの助詞の選択記号を変形する。(例) N1 は/の → N1 (は|の)
変数の場合は、変数以降を選言にする。字面の場合は、個別に対処する。
- 見なし格要素の直前に、連体修飾付き格要素・副詞の離散記号「/bf」を付与する。
- 見なし格要素の直後に、任意の読点「[,]」を付与する。
- 見なし格要素が2個以上存在するときは、順序を任意化する。(例) {N1 は, N2 に}
- 述語の直前に、連体修飾付き格要素・副詞の離散記号「/bf」を付与する。
- 述語に標準表記の記号を付与する。(例) N2 に ' 行く '
- 後続表現を、関数に置き換える。
- 末尾に文末表現の離散記号「/tail」を付与する。

第4章から比べて新しく定義している記述子は、「/cnj」、「/b」、「/tail」である。

「/cnj」は、入力文の先頭にある接続詞を読み捨てるための離散記号である。たとえば、「そして、彼は帰った。」という文と、「/cnjN1 は V2.kako。」というパターンとは適合する。「そして、」が「/cnj」で読み捨てられるからである。

「/b」は、連体修飾表現と格要素の離散記号である。たとえば、「/bv1.kako。」というパターンに対し、文「昨日買った鯖を焼いた。」、文「新鮮な鯖を焼いた。」、および、文「鯖を焼いた。」は適合する。下線部が「/b」と適合するためである。連体修飾節や連体修飾要素の離散記号「/tk」は係り先の名詞を含まないが、「/b」は係り先の名詞および格助詞を含む。

「/tail」は、述語の後続表現の離散記号である。たとえば、「v1/tail。」というパターンと「焼いている。」、「焼きました。」、「焼こうかな。」などの文は適合する。下線部が「/tail」と適合するためである。第4章の実験において、関数の汎化に効果があることを受けて導入した。

6.2.3 変形の実例

例 6.1 から例 6.5 までのパターンを変形した結果を以下に示す:

- /cnj/y{/bf[/tkN1 が[,]], /bf[/tkN2 (から|より)[,]], /bf[/tkN3 (に|へ|まで)[,]]} /bf'旅する' /tail
- /cnj/y{/bf[/tkN1 が[,]], /bf[/tkN2 (に|の後ろから)[,]]} /bf'付く'.teiku /tail
- /cnj/y[/tkN1 (は|の)] /tk 心が[,] /bf'暖かい' /tail
- /cnj/y{/bf[/tkN1 が[,]], /bf[/tkN2 の] /tk 記憶に[,]} /bf'新しい' /tail
- /cnj/y{/bf[/tkN1 が[,]], /bf[/tkN2 に[,]], /bf/tk ばったり[,]} /bf'遭う' /tail

個別に対処に必要な部分として、格要素の選択がある。たとえば、「引き分けと/に」を「引き分け(と|に)」に変形した。約 1.4 万パターンのうち対処すべきは 80 種類であった。一方、述語の後続表現を関数に置き換えるものは、上記の例において「付く(ていく)」を「付く.teiku」とすることである。こうした置き換えは 41 種類であった。これらの個別対処は、簡単な書き換え規則により対処した。

以上により、文型パターンを自動的に変換することができた。図 6.1 の文型パターン辞書には、変換後のパターンを登録しておく。

6.3 適合パターンの選択

パターン言語処理では、1つの入力文に対して、複数のパターンが適合することがある。適合パターンを選択する問題に対して、適合ユニットから適合の仕方に関する特徴量(フラグを含む)を算出し、ヒューリスティクスにより選択を行う。

6.3.1 選択に用いる特徴量

本方式において、次の点から特徴量を求めることにした。

- 意味解析に関すること
 - 意味属性制約を満たすか? (true/false)
 - 無生物主語の文型パターンが適合する際、入力文に主語が明示されているか? (true/false)
 - 非線形性の強さ。(字面記述子の適合した単語数)
 - パターンの適合性。(文型パターンに適合した入力文の単語の数)
- 情緒推定に関すること
 - 情緒主が主体であるか? (true/false)
- 辞書の構造に関すること
 - 思いつきやすさ。(文型パターンの枝番号)

「意味解析」について説明する。非線形言語モデルでは、線形要素に対して、統語的かつ意味的な制約を課すことで、意味解析を行おうとしている。ここで、意味的な制約について、本章で示した情緒推定方式では、統語的な制約の充足は SPM で保証されているが、意味的な制約の充足は保証されていない。そこで、適合結果に対して、意味属性制約の充足についての特徴量を測る必要がある。たとえば、表現「その会社が橋を千代川に掛ける」に、次の2つのパターンと統語的に

適合したとする:

- N1 (主体) が N2 (装身具) を N3 (首) に/へ 掛ける wear の意味
- N1 (主体) が N2 (橋) を N3 (場所, 場) に/へ 掛ける build の意味

前者のパターンは, N2 と N3 の意味属性制約が満たされないまま適合し, 後者のパターンはいずれの変数も意味属性制約が満たされている. パターンに対応する意味を見ると, 後者のパターンの方が適切である. こうして, 意味属性制約により適切なパターンを優位にすることができる.

「無生物主語」については, 日本語の理解のしやすさに関わるものである. 無生物主語を持つ文型パターンは, 無生物主語を持つ入力文と適合すると, 適切さが高く, 無生物主語を持たない入力文と適合すると, 適切さは低いと予想される. たとえば, 表現「疲れる」に, 次の2つのパターンが適合したとする:

- N1 (962 機械) が N2 (*) で/から 疲れる fatigue の意味
- N1 (人) が N2 (*) に 疲れる tired の意味

前節で, 離散記号や任意記号を導入したので, 格要素を持たない文も, これらのパターンには適合する. 通常「疲れる」という表現からは「人が疲れる」や「体が疲れる」という解釈をやすく, 「機械が疲れる」という解釈はしにくいであろう.

「非線形性の強さ」については, 慣用句を表すパターンの適合を特徴付ける効果がある. 慣用句のパターンの場合, 字面でパターンが記述される. そこで, 字面の記述子に適合した単語数が, 非線形性を表す特徴量になる. たとえば, 表現「へまを遣る」に, 次の2つのパターンが適合したとする:

- N1 (主体) が へまを 遣る bungle の意味
- N1 (主体) が N2 (人間活動, 抽象物) を 遣る do の意味

「へま」は「抽象物」の下位概念なので, 両パターンは, 統語的にも意味的にも適合する. しかし, 前者のパターンの方が, 非線形要素として適合する単語数が多く, 鋭く意味解析ができています.

「文型パターンの適合性」については, 変数や字面という制約をより多く満たしながら適合したユニットを優先するという考え方である. たとえば, 表現「田舎に帰る」は, 「N1 (人, 動物) が N2 (場所, 場, 抽象物, 人間活動) から/より N3 (場所, 場) に/へ 帰る」という1つのパターンであっても, 離散記号が格要素と対応するために, 複数の適合の仕方がある.

- N3 = 田舎 となる場合
- N3 = なし となる場合

この例の場合, 前者の適合の仕方が妥当である.

一方, 「情緒推定に関わること」として, 情緒主が「主体」であることという条件付けは自明である.

「辞書の構造に関すること」については, 同一パターンに複数の情緒属性セットを付与するとき, 分析者が思いついた順に付与されたと仮定すると, 枝番の小さいものの方が, 人が言語理解の際に思いつきやすいものだと考えられることから設けた. たとえば, 「N1 が/から N2 に/と/へ 挨拶する」というパターンには, 次の2つの情緒属性セットが対応付けられている.

- EC000003-01: 情緒名: 好ましい, 原因: 信頼, 情緒主: N2, 情緒対象: N1
- EC000003-02: 情緒名: 嫌だ, 原因: 不信, 情緒主: N2, 情緒対象: N1

知人から挨拶してもらおうと好ましく感じるが, 見ず知らずの人から挨拶されると好ましくは感じないということを, これらのパターンは表している.

6.3.2 ヒューリスティクス

上記に挙げた特徴量を適合ユニットから算出した後に、特徴量を比較しながら、以下のヒューリスティクスにより選択を行う。

情緒推定に関するヒューリスティクス

Step-1: 情緒主が主体であること

語義解析に関するヒューリスティクス

Step-2: 意味属性制約を充足すること

Step-3: 無生物主語の文型パターンに適合する際、主語が入力文に明示されていること

Step-4: 非線形性が強いこと

Step-5: パターンの適合性が高いこと

辞書の構造に関するヒューリスティクス

Step-6: 文型パターン ID の枝番が小さいこと

Step-7: 文型パターン ID の番号が小さいこと

これらは予備的な実験により作成した。最後の Step-7 は、いかなる場合でも推定結果を 1 つに絞り込むためのものである。以上の他に、「情緒対象が明示されること」というヒューリスティクスも試みたが、効果はみられなかった。将来、文脈情報を取り込む場合や、構文解析情報と協調的に処理を行う場合には、これらのヒューリスティクスを見直す必要がある。なお、ヒューリスティクス以外の方法として、パターンベースの機械翻訳では多変量解析を用いる方法が提案されている[75],[76]。今後、特徴量の種類を増やすとともに統計的手法による選択を試みる価値がある。

6.4 情緒推定の過程

具体例を参照しながら、パターン言語処理に基づく情緒推定を実行する様子を説明する。

6.4.1 入力文の表層レベルの解析

文を入力すると、形態素解析、文節境界の解析、意味属性の付与が行われる。

例として「鯉を釣った。」を入力すると、図 6.2 のように解析結果が得られる。単語ごとに通番が付与されている。単語の前の「/」と「+」は文節の境界かどうかを表している。4桁の数字は品詞コードである。{ } 内は意味属性コードである。ここでは、日本語語彙大系の一般名詞意味属性および用言意味属性のコードを掲載している。

```
INPUT=spm_work_file.txt=0x3f=0x1f
鯉を釣った。
1. /鯉(1100, {NI:543,NI:842})
2. +を(7430)
3. /釣っ(2384, 釣る, 釣っ, {NY:23,NY:31})
4. +た(7216)
5. +。(0110)
6. /nil
```

図 6.2 表層レベルの解析例

6.4.2 情緒属性付き文型パターン辞書

情緒属性付き文型パターン辞書の具体例を図 6.3 に示す。パターン ID , 文型パターン, 意味属性制約, 情緒名, 原因, 情緒主, 情緒対象をタブ区切りで1行に並べたものを1つの単位として, 本辞書を構成している。この辞書から, データベースを生成し, 実行時には効率良く運用する。

EC000001-00 のパターンのように, 情緒属性の無いものがある。これらは適合時には, 情緒が「なし」という判定になる。EC009033 は, 00 から 02 まで3つの枝番号を有する。情緒属性が異なっていることが分かる。

パターンID	文型パターン	意味属性制約	情緒名	原因	主	対
EC000001-00	/cnj/y{/bf[/tkN1 が [,]],/bf[/tkN2 を [,]]} /bf'愛好する'/tail	N1(4 人) N2(1001 抽象物 1560 行為)				
EC000002-00	/cnj/y{/bf[/tkN1 が [,]],/bf[/tkN2 を [,]],/bf[/tkN3(から よ り)[,]]} /bf'愛護する'/tail	N1(4 人 535 動物) N2(4 人 535 動物) N3(*)	喜び	保護	N2	N1
EC000003-00	/cnj/y{/bf[/tkN1(が から)[,]],/bf[/tkN2(に と へ)[,]]} /bf'挨拶する'/tail	N1(4 人) N2(4 人)	好まし い	信頼	N2	N1
EC000003-01	/cnj/y{/bf[/tkN1(が から)[,]],/bf[/tkN2(に と へ)[,]]} /bf'挨拶する'/tail	N1(4 人) N2(4 人)	嫌だ	不信	N2	N1
	(中略)					
EC009030-00	/cnj/y{/bf[/tkN1 が [,]],/bf[/tkN2 を [,]],/bf[/tkN3 から [,]]} /bf'吊る'/tail	N1(3 主体 533 具体物) N2(533 具体物) N3(388 場所 533 具体物 2610 場)				
EC009031-00	/cnj/y{/bf[/tkN1 が [,]],/bf[/tkN2 を [,]],/bf[/tkN3 に [,]]} /bf'吊る'/tail	N1(3 主体 533 具体物) N2(533 具体物) N3(388 場所 533 具体物 2610 場)				
EC009032-00	/cnj/y{/bf[/tkN1 が [,]],/bf[/tkN2 を [,]]} /bf'釣る'/tail	N1(4 人) N2(542 魚 介)	喜び	獲得	N1	N2
EC009033-00	/cnj/y{/bf[/tkN1 が [,]],/bf[/tkN2 を [,]],/bf[/tkN3 で [,]]} /bf'釣る'/tail	N1(3 主体 535 動物) N2(3 主体) N3(*)	喜び	獲得	N1	N2
EC009033-01	/cnj/y{/bf[/tkN1 が [,]],/bf[/tkN2 を [,]],/bf[/tkN3 で [,]]} /bf'釣る'/tail	N1(3 主体 535 動物) N2(3 主体) N3(*)	怒り	嘘	N2	N1
EC009033-02	/cnj/y{/bf[/tkN1 が [,]],/bf[/tkN2 を [,]],/bf[/tkN3 で [,]]} /bf'釣る'/tail	N1(3 主体 535 動物) N2(3 主体) N3(*)	嫌だ	嘘	N2	N3
	(以下省略)					

図 6.3 情緒属性付き文型パターン辞書(一部)

6.4.3 SPM による構造照合型パターン検索

表層レベルの解析データを SPM に与えると, 文型パターン辞書から全ての適合の仕方が検索される。検索結果は, パターン ID , パターンの適合状態, 変数の適合状態, および, 各種記憶値で構成された適合ユニットを単位としている。

例として、図 6.2 のデータを与えると、図 6.4 のとおり検索される。ここには、14 件の適合ユニットが列挙されている。7種類のパターンが適合し、各パターンにおいて、2通りの適合の仕方があることが分かる。1パターンにつき2通りの適合の仕方があるのは、格要素の離散記号に「鰹を」が適合する場合としない場合の違いによるものである。

「 PATTERN=」の行に注目する。「 EC009032-00 」というのがパターン ID である。その次の「 [N2, を, '釣る'] 」が適合したパターン記述子である。その次の「 [1, 2, 3] 」は、適合した単語の番号である。行末の「 3 」は適合した単語数である。

次の行を見ると、「 N2=」とある。変数に代入された情報が書かれている。最初の「 [1] 」は変数に適合した単語の番号である。次の「 {1} 」は意味属性制約の係り先となる単語の番号である。行末の「 1 」は適合した単語数である。

```

PATTERN=EC009032-00=[N2, を, '釣る']=[1, 2, 3]=3
N2=[1]={1}=1
-----
PATTERN=EC009033-00=[N2, を, '釣る']=[1, 2, 3]=3
N2=[1]={1}=1
-----
PATTERN=EC009033-01=[N2, を, '釣る']=[1, 2, 3]=3
N2=[1]={1}=1
-----
PATTERN=EC009033-02=[N2, を, '釣る']=[1, 2, 3]=3
N2=[1]={1}=1
-----
PATTERN=EC009034-00=[N2, を, '釣る']=[1, 2, 3]=3
N2=[1]={1}=1
-----
PATTERN=EC009034-01=[N2, を, '釣る']=[1, 2, 3]=3
N2=[1]={1}=1
-----
PATTERN=EC009034-02=[N2, を, '釣る']=[1, 2, 3]=3
N2=[1]={1}=1
-----
PATTERN=EC009032-00=['釣る']=[3]=1
-----
PATTERN=EC009033-00=['釣る']=[3]=1
-----
PATTERN=EC009033-01=['釣る']=[3]=1
-----
PATTERN=EC009033-02=['釣る']=[3]=1
-----
PATTERN=EC009034-00=['釣る']=[3]=1
-----
PATTERN=EC009034-01=['釣る']=[3]=1
-----
PATTERN=EC009034-02=['釣る']=[3]=1
=====

```

図 6.4 文法レベルの適合結果の例

6.4.3.1 意味属性制約の充足の検査

適合ユニットの中には文型パターンが指定する線形要素への意味属性制約が満たされていないものも含まれている。意味属性制約が満たされていなくても正しい適合であることは、全く無い訳ではない。そこで、制約が充足されたか否かのフラグを付与することとして、意味的に不適合であっても即座に削除することを行わない。

図 6.4 を例にする。EC009032-00 というパターンは、簡単に書くと「 N1 (4 人) が N2 (542 魚介) を 釣る」である。英語パターンは「 N1 catch N2 」である。変数 N2 に「 542 魚介」という意味属性コードまたはその下位概念の意味属性コードの単語が代入されているならば、制約が充足となる。この図の1件目はそれが満たされている。次に、EC009033-00 というパターンは、「 N1 (3 主体) が N2 (3 主体 535 動物) を N3 (*) で釣る」である。「 N1 attract N2 with N3 」という英語パターンが対応することから分かるのとおり、「感情動作」として解釈する場合のパターンである。パターン ID には「-00 」や「-01 」の枝番が付けられている。枝番が異なる場合は、文型パターンは同一であるが、情緒属性が異なることを表している。次のパターン EC009034-00 は、「 N1 (3

主体)が N2 (134 客)を N3 (*)で釣る」である。この場合、「鯉」は「134 客」の意味属性制約を満たさない。以上のようにして、意味属性制約の充足を検査すると、図 6.5 が得られる。PATTERN の行末の「T」は充足したことを、「F」は充足しないことをそれぞれ表す。

```

PATTERN=EC009032-00=[N2,を,'釣る']=[1,2,3]=3=T
N2=[1]={1}=1
-----
PATTERN=EC009033-00=[N2,を,'釣る']=[1,2,3]=3=T
N2=[1]={1}=1
-----
PATTERN=EC009033-01=[N2,を,'釣る']=[1,2,3]=3=T
N2=[1]={1}=1
-----
PATTERN=EC009033-02=[N2,を,'釣る']=[1,2,3]=3=T
N2=[1]={1}=1
-----
PATTERN=EC009034-00=[N2,を,'釣る']=[1,2,3]=3=F
N2=[1]={1}=1
-----
PATTERN=EC009034-01=[N2,を,'釣る']=[1,2,3]=3=F
N2=[1]={1}=1
-----
PATTERN=EC009034-02=[N2,を,'釣る']=[1,2,3]=3=F
N2=[1]={1}=1
-----
PATTERN=EC009032-00=['釣る']=[3]=1=T
-----
PATTERN=EC009033-00=['釣る']=[3]=1=T
-----
PATTERN=EC009033-01=['釣る']=[3]=1=T
-----
PATTERN=EC009033-02=['釣る']=[3]=1=T
-----
PATTERN=EC009034-00=['釣る']=[3]=1=T
-----
PATTERN=EC009034-01=['釣る']=[3]=1=T
-----
PATTERN=EC009034-02=['釣る']=[3]=1=T
=====

```

図 6.5 意味属性制約の充足フラグ付き

6.4.3.2 情緒属性の付与

パターン ID をキーとして文型パターン辞書を参照し、情緒属性セットを得る。そして、情緒属性セットを適合ユニットに付与する。

図 6.5 を例にする。EC009032-00 のパターンには、「情緒主：N1」、「情緒対象：N2」、「情緒名：喜び」、「原因：獲得」という情緒属性のあることが辞書より分かる。変数には、適合結果より代入されている単語を展開する。変数に単語が代入されていないときはゼロ代名詞として「φ」を出力する。こうして、全ての適合結果に対して情緒属性の具体化をする。

6.4.3.3 その他の特徴量の付与

その他の特徴量として、無生物主語であるか否か、非線形性の強さ、適合単語数、文型パターン ID の枝番号がある。無生物主語については、文型パターンの「は」あるいは「が」で終わるタームについての意味属性制約から判別する。残りの特徴量は、適合ユニットに含まれる情報から直接得られる。

例として、図 6.5 の適合ユニットから特徴量を求めた結果を表 6.1 にまとめる。表中の、M は適合した単語の総数、N は非線形要素として適合した単語の数、S は意味属性制約が充足されたか否か (T または F) を表す。O は無生物主語のパターンに対するフラグで「1, 0, -1」の3値のいずれかをとる。無生物主語のパターンでない場合は「0」となり、無生物主語のパターンである場合は、主語が成功裏に適合していると「1」、そうでなければ「-1」となる。A は情緒主となる変数に主体の意味属性の単語が適合しているかどうかを表すフラグで「T,U,F」の3値のいずれかをとる。格要素が省略されている場合は「U」となり、そうでない場合は、主体の意味属性の単語が適合して

いると「T」となり、そうでなければ「F」となる。

なお、「可否」の列は、ヒューリスティクスを満たすものには「可」、そうでないものには「否」を付与している。

表 6.1 各適合ユニットにおける特徴量

#	パターン ID	原因	情緒名	情緒主	情緒対象	M	N	S	O	A	可否
1	EC009032-00	獲得	喜び	φ	鯉	3	2	T	0	U	可
2	EC009033-00	獲得	喜び	φ	鯉	3	2	T	0	U	可
3	EC009033-01	嘘	怒り	鯉	φ	3	2	T	0	F	否
4	EC009033-02	嘘	嫌だ	鯉	φ	3	2	T	0	F	否
5	EC009034-00	獲得	喜び	φ	鯉	3	2	F	0	U	否
6	EC009034-01	嘘	怒り	鯉	φ	3	2	F	0	F	否
7	EC009034-02	嘘	嫌だ	鯉	φ	3	2	F	0	F	否
8	EC009032-00	獲得	喜び	φ	φ	1	1	T	0	U	否
9	EC009033-00	獲得	喜び	φ	φ	1	1	T	0	U	否
10	EC009033-01	嘘	怒り	φ	φ	1	1	T	0	U	否
11	EC009033-02	嘘	嫌だ	φ	φ	1	1	T	0	U	否
12	EC009034-00	獲得	喜び	φ	φ	1	1	T	0	U	否
13	EC009034-01	嘘	怒り	φ	φ	1	1	T	0	U	否
14	EC009034-02	嘘	嫌だ	φ	φ	1	1	T	0	U	否

6.4.4 情緒推定結果の出力

適合ユニットの絞り込みと順位付けが行われた後は、最も優位である適合ユニットに付与された情報より、情緒の原因、情緒名、情緒主、情緒対象を参照し、出力する。

表 6.1 によると、#1 が出力される。「鯉を釣った。」からは「獲得」による「喜び」が推定された。情緒主はこの文の主語であり、ここでは省略(φ)されている。情緒対象は「鯉」である。

出力されなかったものについて見てみると、#2 も「可」ではある。#2 のパターンは「attract」の語義での「釣る」である。「鯉」を「惹き付けた」という点で喜んでいる。#3 は、情緒主が「鯉」となる場合で、パターンの語義は#2 と同一である。したがって、「惹き付けられた」ということにより「鯉が怒っている」という解釈になる。

言語理解を行う上で、曖昧性をどの段階で解消するのが重要である。ここで示したヒューリスティクスは、1文の単位で情緒推定(タイプ1の情緒推定)を行うためのものである。照応・省略解析、談話構造解析、非情緒的な意味の解析などと協調する場合には、ここで示したヒューリスティクスでは適切ではなく、協調の仕方に応じた変更が必須である。しかし、表 6.1 から想像できるように、幾通りかの解釈の道筋を立てる情報が、#1 から#14 の中に秘められていることから、将来、協調処理に寄与することが期待できる。

6.5 動作確認

基本的な動作の確認を行う。1点目は、多義のある用言について、動作の様子を調べる。2点目は、簡単な作例において、動作を確認する。

6.5.1 語義解析と情緒推定

本研究の情緒推定方法は、表層構造を参照する方法であるため、多義のある用言の使われた

入力文であっても情緒推定が正しく動作することが望まれる。

比較的多義のある用言「当たる」について動作を確認する。「当たる」には 38 種類のパターンがあることから、多義性が高い。実行結果を表 6.2 にまとめる。表中の M, N, S, O, A は表 6.1 と同一の意味である。表中の「同」は、情緒推定結果に同意できそうか否かを、ok と ng で表している。

#1 の「風にあたった。」については、気持ちの良い風であったのか、寒い風であったのかに依存するが、同意できる。なお、次点として「不快による悲しみ」も存在するので、文外情報との協調により絞り込みが可能であると予想される。#2 から#4 については期待した通りの結果である。

表 6.2 「あたる」を含む文の情緒推定

#	入力文	適合パターン	原因	情緒名	情緒主	対象	M	N	S	O	A	同
1	風にあたった。	EC000383-00	内的な快	喜び	φ	風	3	2	T	0	U	ok
2	宝くじにあたった。	EC000389-00	獲得	喜び	φ	宝籤	3	2	T	0	U	ok
3	鯖にあたった。	EC000384-00	内的な不快	悲しみ	φ	鯖	3	2	T	0	U	ok
4	岩にあたった。	EC000390-00	外的な不快	悲しみ	φ	岩	3	2	T	0	U	ok

6.5.2 ショートストーリーでの情緒推定

次に簡単な出来事を並べてみて、1文ごとの情緒推定を試みる。文章は日記の作例である。結果を表 6.3 にまとめる。推定結果に対する同意率は約 80% (11/14) である。

しかし、幾つか問題点がある。#6 と#12 は「帰る」という語義は同じであるが、状況が異なる。通常では「家に帰る」ことは「自由による好ましい」が推定されてもおかしくはない。しかし、ここの2つの例では文脈の影響が強いため正しく推定できたとはいいいにくい。#6 では、忘れ物を取りに帰る行動なので、自由になることには繋がらず、情緒推定は誤りである。#12 では以前の文の情緒が継続していると考えると、「しょんぼり」とした様子が想像できるので、情緒推定は正しくないかもしれない。しかし、気持ちの継続は「気分」と分類されるため、「家に帰った。」ということで気持ちが切り替わるかもしれない。そのため「？」と判定した。

人間が言語表現から情緒を推定する際には、情緒と気分の区別が容易ではないかもしれない。今後、情緒の継続性について、文脈の理解と共に進める必要がある。

表 6.3 ショートストーリーに対する情緒推定

#	入力文	適合パターン	原因	情緒名	情緒主	対象	M	N	S	O	A	同
1	朝6時に目が覚めた。	EC005354-00	内的な快	好ましい	φ	φ	3	3	T	0	U	ok
2	朝ごはんを食べた。	EC008113-00	五感(味覚)	好ましい	φ	朝ごはん	3	2	T	0	U	ok
3	顔を洗った。	EC000618-00	補助	好ましい	φ	φ	3	2	T	0	U	ok
4	学校へ行った。	EC001029-00		なし	φ	φ	1	1	T	0	U	ok
5	宿題を忘れた。	EC014664-00	喪失	悲しみ	φ	宿題	3	2	T	0	U	ok
6	家に取りに帰った。	EC002544-00	自由	好ましい	φ	φ	3	3	T	0	U	ng
7	学校に遅刻した。	EC008280-00	その他	嫌だ	φ	φ	3	2	T	0	U	ok
8	先生が叱った。	EC005477-00	説教	嫌だ	φ	先生	3	2	T	0	U	ok
9	英語と数学を勉強した。	EC012574-00	補助	好ましい	φ	φ	3	2	T	0	U	ok
10	物理は試験があった。	EC000019-00		なし	φ	φ	3	2	T	1	U	ok
11	試験に落ちた。	EC002144-00	非難	悲しみ	φ	φ	3	2	T	0	U	ok
12	家に帰った。	EC002544-00	自由	好ましい	φ	φ	3	3	T	0	U	?
13	晩ごはんを食べて、	EC008113-00	五感(味覚)	好ましい	φ	晩ご飯	4	2	T	0	U	?
14	すぐに寝た。	EC010841-00		なし	φ	φ	1	1	T	0	U	ok

6.6 まとめ

本章では、パターン言語処理に基づく情緒推定の方式を実装した。第5章で構築した知識ベースを運用する上で、次の2点が問題となっていた。

(1) 文型パターンが単純であるため、入力文に適合し難いこと

(2) 複数の文型パターンが適合するため、適切なものを選択しなければならないこと

(1) については、第4章の結果を参考にして、文型パターンの拡張を行った。接続表現、修飾表現の自由化、格要素の自由化、格要素の順序の任意化、および、述語後続の表現の自由化である。(2) については、適合の仕方に対する特徴量を測り、ヒューリスティクスにより、複数の適合結果の中から選択する方法を示した。特徴量としては、意味解析の観点から、意味属性制約、無生物主語、非線形性、パターンの適合性の4点、情緒推定の観点から、情緒主の主体性の1点、そして、辞書構築の観点から、思いつきやすさの1点を立てた。ヒューリスティクスとして、これらの特徴量に対する優先度を定め、順に望ましい適合結果を抽出することで選択した。

こうして実装した後、動作の様子を具体的に示し、多義のある場合や簡単な文章において、良好に動作することを確認した。

第7章 実験

本章では、提案方式を実験により評価する。実験には情緒タグ付き日記コーパスを使用する。このコーパスの情緒タグから情緒推定の正解が決定される。その正解と本方式による情緒推定の結果を比較することで、本知識ベース、および、本方式を評価する。

7.1 目的と方法

本実験では、次の2点の評価を目的とする。

- **網羅性**：本研究では網羅性の高い知識ベースの構築を目標の1つとしている。情緒生起原因に基づくタイプ1の情緒推定方法が適用できる文のうち、実際に本知識ベースの文型パターンが適合した文の割合を評価する。
- **精度**：提案方式の出力した情緒と正解の情緒との間の一致率により、性能を評価する。一致率は、人間による情緒推定の一致率と比較する。

また、精度を求めた後に、本知識ベースによる情緒推定の誤り事例を分析する。本知識ベースの不備が原因で情緒推定に失敗した割合を明確にして、本知識ベースの信頼性を確認する。

本実験には、日記文(第 3.2.3 節参照)を使用する。この日記は、情感で章立てが構成されており、喜び、悲しみ、怒り、などの情緒の種類をカバーしているので、実験対象として適している。情緒推定は、日記の作者を情緒主として行う。情緒推定の正解とする情緒は、第 3.2.3 節で付与した情緒タグに基づき定める。すなわち、情緒推定の熟練者 a 者から f 者までの6名の多数決により定める。また、読み取られる情緒にばらつきのある場合も存在する。その際は正解の無い場合とする。これは、実験対象から外す。具体的には、6名のうち4名以上の合意の得られた情緒を正解とする。

7.2 実験 7-1：知識ベースの網羅性

日記コーパスは、1,642 節で構成されている。このうち、情緒生起原因の節とみなして情緒推定をすることのできる節は、68.3% (1,122 節) である。しかし、この中には、格要素上に情緒主が直接現れていない場合などのように、本方式では対象外となる節も含まれている。

本方式で対象とする節は、次の4つの条件を満たす節である：

- a) 述語数が1つであること(連体修飾節や引用節のあるものは除外)
- b) 情緒主が、用言の格要素に登場すること(ゼロ代名詞として登場しても良い)
- c) 情緒生起の原因を表すこと(情緒状態や情緒的な反応を表す節は除外)
- d) 情緒主に起こった事態であること

a) については、たとえば、「くだらないとは分かっている」(引用節)や「興味をひくような会社はまだない」(連体節)という表現であるが、こうした部分を単独の節として取り出したところで、その部分だけからでは情緒が読み取りにくい。そのため、コーパス作成時に節として分割しなかった。また、「～しに～する」や「～して～する」のように動詞が連続して複合動詞のようになる場合も事態が若干複雑になるため除外する。b)、c) については、第 5.2.2 節で既に述べたとおりである。d) は、c) の

補足である。すなわち、未然の事態に情緒生起原因を見い出して情緒を推定するには、本知識ベースの情緒を読み替える必要があるため本実験では対象外とする。具体的に d) に該当しない節の種類は、「否定」、「疑問」、「未来」、「願望」、「条件」、「仮想」を表す節である。以下に幾つか例を示す：

対象外1) 散歩に行けなかった。(否定の節)

対象外2) 一人で映画を見に行くことにした。(未来の節)

対象外3) 毎週末、こうして鍛えれば、(条件の節)

これらの客体的表現(「散歩に行く」、「一人で映画を見に行く」、「毎週末、こうして鍛える」)は、表現されている事態が発生していない。なお、読み替えとは、「不可能」、「寂しいことの計画」、「期待できる条件」と解釈することで、「悲しみ」や「期待」を推定することである。

1,642 節より条件を順に適用して、実験対象となる節を求めると次のようになる：

a) 埋め込み文や引用文のため対象外となった節 ... 477 節

b) 単文だが、情緒主が登場しないため対象外となった節 ... 510 節

c) 単文で情緒主が登場するが、情緒の状態や反応を表すため対象外となった節 ... 248 節

d) 情緒主の登場する単文で情緒の状態や反応を表すが、情緒主に起こった事態でないため対象外となった節 ... 76 節

したがって、条件を満たす節は、331 節となる。

この 331 節に対して本知識ベースの文型パターンを照合したところ、適合した節は 301 節、適合しなかった節は 30 節であった。以上の分析によりカバー率について以下のことが言える：

- 狭い範囲での情緒生起原因に基づくタイプ1の情緒推定は、日記コーパスの 20% (331/1,642)に適用可能な方式である。
- 本知識ベースの網羅性は、91%(301/331)である。

「狭い範囲」とは、情緒主が客体的表現に表れる場合に限定していることを示す。別の範囲として情緒主の現れていない場合があり、たとえば、「雨が降った。」のような表現である。文脈上、情緒生起原因と言えるため、広い範囲ではこれに含まれる。また、情緒の状態や反応を表す表現は、日記文の 5% ~ 6%と、情緒を推定するチャンスはあまり高くないため、20%という方式のカバー率は低いものではない。また、網羅性は 91%であるので、本知識ベースは良好な部類に属すると言える。

7.3 実験 7-2 :情緒推定の精度

情緒推定実験の対象は、実験 7-1 で得た 331 節のうち、人間が共通的に情緒推定することのできた 204 節とする。その基準は、6 名のうち 4 名以上が合意した情緒の得られる節とする。

本実験では、情緒推定の初心者が行った場合の精度、本知識ベースを手動で運用した場合の精度、および、第6章で示したとおり自動で運用した場合の精度について比較する。

7.3.1 情緒推定の様子

以下に、日記コーパスの一部を見ながら、情緒推定の様子を説明する。

- | | |
|-------------------------|--------------------------|
| 節1) 蒲郡のおばさんからみかんが1箱届いた。 | 《喜び》4名, 《好ましい》2名, 《なし》1名 |
| 節2) とても甘くて | 《喜び》3名, 《好ましい》3名, 《なし》1名 |
| 節3) おいしかった。 | 《喜び》4名, 《好ましい》4名 |

節1)は、「日記の作者へ」という格が省略されている。この場合、日記の作者が登場していると見なす。また、4名の合意する情緒として《喜び》があるので、節1)は、実験対象となる。一方、節2)および節3)は、情緒主が登場しないので、対象外となる。

節1)の文を照合すると、パターン「N1 が N2 に/へ/まで 届く」が適合結果として得られる。このパターンには情緒属性として「情緒名：《喜び》, 原因：〈獲得〉, 情緒主：N2, 情緒対象：N1」のセットが対応付けられているため、本知識ベースによる情緒推定では、《喜び》を出力することになる。

以下に幾つか具体例を示す：

- アリイが夕食に招待してくれた。
 - 正解の情緒:《喜び》4名／少数意見:《好ましい》2名,《期待》2名
 - 適合パターン:「N1 が N2 を N3 に/へ招待する」
情緒属性:「情緒名：《喜び》, 原因：〈厚遇〉, 情緒主：N2 , 情緒対象：N1」
 - 推定された情緒:《喜び》
- (垢すりの話) 長い台の上に横になった。
 - 正解の情緒:《なし》6名／少数意見:《期待》1名
 - 適合パターン:「N1 が N2 に横に成る」
情緒属性:なし
 - 推定された情緒:《なし》

上記の2件は、情緒推定に成功している。

次の例をみてみよう。

- 母の日に子供たちから花束をもらった。
 - 正解の情緒:《喜び》6名
 - 適合パターン「N1 が N2 (*)を N3 に/から/より貰う」
情緒属性:「情緒名：《好ましい》, 原因：〈厚遇〉, 情緒主：N1 , 情緒対象：N3」
- (以下は、別の候補)
- 情緒属性:「情緒名：《喜び》, 原因：〈獲得〉, 情緒主：N1 , 情緒対象：N2」
 - 情緒属性:「情緒名：《悲しみ》, 原因：〈内的な不快〉, 情緒主：N1」
 - 推定された情緒:《好ましい》

ヒューリスティクスが働き、〈厚遇〉による《好ましい》が選択されて情緒推定に成功している。次候補の〈獲得〉による《喜び》も推定されても良い。しかし、〈内的な不快〉による《悲しみ》は候補から削除されるべきである。この情緒が推定されるべきは、N2 の名詞意味属性が(病気)の場合である。日本語語彙大系の文型パターンは、英語パターンの区別のために意味属性による制約が付けられているので、“get”という訳語選択の上で、(病気)の意味属性制約は付けられていない。今後、緻密に意味属性制約を付けることで解決する。

次に明らかな誤りの例を示す。

- (パチンコの話) 2時間で4万円も擦ってしまった。
 - 正解の情緒:《悲しみ》5名／少数意見:《嫌だ》2名
 - 適合パターン「N1 (人, 動物, 人工物) が N2 (動物, 人工物) を擦る」(読み:こする)
情緒属性:「なし」
- (以下は別の候補)

- 適合パターン「N1 (人) が N2 (人間活動) で N3 (金銭) を擦る」(読み: する)
情緒属性:「情緒名:《悲しみ》, 原因:〈喪失〉, 情緒主: N1, 情緒対象: N3」
- 推定された情緒:《悲しみ》

動詞の読みが異なるため、第一候補では情緒推定に失敗してしまう。形態素解析時の曖昧性が問題である。同様の例として「行った」を「おこなった/いった」と解析する場合がある。なお、読みが「する」の場合、「マッチを擦る, ブラシで擦る, 胡麻を擦る」などの多義があるが、N3 の名詞意味属性(金銭)が適合するのはこのパターンだけであり、「する」の多義は適切に解消されている。

7.3.2 結果

実験結果を表 7.1 にまとめる。正解とする情緒と、各推定者の出力した情緒との一致率について、第 3.1.4 節の評価式を用いて評価した。

「初心者」とは、情緒の研究に携わっていない大学生である。情緒の種類については区別ができる。具体的には第 3.2.3 節における g 者である。日記文を読みながら情緒推定を行ったので、文脈を参照しながらの情緒推定となる。

「KB 手動」とは、本研究で構築した知識ベースを手動で運用して情緒を推定することである。「KB 自動」とは、第6章で実装した情緒推定方式により情緒を推定することである。知識ベースを利用する上で、文脈情報や述語以外の言語情報(たとえば、「～てしまった」)は参照しなかった。手動と自動の違いは、情緒主の区別ができていること、形態素解析や名詞意味属性制約の検査が手動であるため柔軟であること、である。

表 7.1 日記文に対する情緒推定の正解率

推定者(文脈参照有無)	再現率	適合率	一致率	(一致数	正解数	出力数)	κ 複合	κ 単独
初心者 (あり)	0.4882	0.4952	49%	(104	213	210)	0.2818	0.3014
KB 手動 (なし)	0.4835	0.4813	48%	(103	213	214)	0.2943	0.3696
KB 自動 (なし)	0.4178	0.4836	45%	(89	213	184)	0.2796	0.2944

この表より次のことが言える。

- 初心者と「KB 手動」を比較すると、正解率はほぼ同等である。
- KB 手動と「KB 自動」を比較すると、「KB 自動」は出力数が少なく、再現率が劣っている。

初心者が情緒推定を行った際、文章として節を読んだので、文脈が考慮に入れられており、省略・照応解析ができた状態であると言える。また、形容詞、副詞、モダリティなどの表現も考慮に入れられている。一方、「KB 自動」においては、それらの言語的情報まで参照しないのだが、精度をあまり落とさずに情緒推定ができていると言える。

7.4 誤り分析

情緒推定の誤り分析を通じて、本知識ベースの問題点、および、本情緒推定方式の問題点を明らかにする。

7.4.1 誤り分析の対象

情緒推定実験の対象となった節は 204 節であった。その内訳は情緒推定の結果によると次のように大別される:

- 情緒推定に成功:
 - 推定結果がそのまま正解の情緒と完全一致(85節)
 - 曖昧性が解消されれば正解の情緒と完全一致(17節)
- 情緒推定に失敗:
 - 文型パターンが不適合のため失敗(22節)
 - その他の失敗(80節)

誤り分析では、「その他の失敗」に属する80節を対象とする。

7.4.2 誤り分析の様子

誤り分析の様子を具体例とともに示す。

(誤り事例1) 家族で回転寿司へ行った。

- 正解の情緒:《好ましい》4名／少数意見:《喜び》2名
- 推定された情緒:《なし》

行き先に依存する問題である。既に先行研究である目良ら[8]や Liu ら[9]で議論されているとおり、格要素の間、および、格要素と用言の関係により解決される問題である。格要素の関係は個別性の高い関係である。たとえば、「回転寿司」が好ましいかどうかは個人に依存する。Liu らはコーパス(大量の Web ドキュメント)の上での共起により確率的に好ましいと思われているかどうかをとらえるモデルである。本知識ベースでは、情緒主と情緒対象を変数で捉える構造があるため、今後、共起に関するモデルを導入することで解決を図る。

(誤り事例2) 今日友達と行ってみた。

- 正解の情緒:《期待》6名／少数意見:《好ましい》1名
- 推定された情緒:《なし》

移動を表す「行く」は情緒生起原因とはならないが、「～てみた。」より情緒主が「試み」として行動していることが伺えるため、《期待》が読み取られたと思われる。これはタイプ2に属する情緒推定である。

(誤り事例3) (先行する文章: もう、何だよ、あの態度！ タクシーを降りてすぐに、)

会社名とナンバーを控えておいた。

- 正解の情緒:《怒り》4名／少数意見:《なし》2名、《嫌だ》1名
- 推定された情緒:《好ましい》

対象としている節「会社名とナンバーを控えておいた。」は、文脈上、《怒り》の反応である。本実験では、表情などの身体に関わる情緒反応は明確に対象外としたが、文脈上の一連の行動過程をみて情緒反応ととれる部分は対象外としていなかった。このように広義の情緒反応ととれる場合については、次の文脈に繋ぐ際に情緒生起の原因ととらえられるためである。ここで、《好ましい》と推定された理由は、「N1 が N2 に N3 を控える」というパターンにおいて、控える媒体となった N2 に対して目標実現の補助となると解析されたためである。簡単に言えば「(メモ帳は)役に立つなあ」と情緒主が思っていると解析している。ここで、文脈のない状態では、推定された情緒について同意が得られるものと思われる。したがって、この誤り事例は、文脈処理の問題である。

(誤り事例4) なべで指にやけどをし、

- 正解の情緒:《嫌だ》6名／少数意見:《悲しみ》2名
- 推定された情緒:《悲しみ》

この事例では、少数意見と一致する情緒が推定されたが、正解の情緒には一致しなかった。「なべで指にやけどをした」というとき、1cmにも満たない僅かな火傷のとき、悲しみに沈むことは想像しにくく、ちょっと嫌だなと思う程度ではないだろうか。タグ付与においては、そのような感じ方でタグが付与されたと考える。したがって、情緒の強さを考慮する必要がある。《悲しみ》、《恐れ》、《怒り》の3つは、非常に弱いときには、《嫌だ》のタグを付与しやすいのではないかと予想する。今後、情緒の強さの検討を行う必要がある。

(誤り事例5) メッセージを残しておいた。

- 正解の情緒:《期待》4名／少数意見:《なし》2名
- 推定された情緒:《なし》

「N1(*)がN2(*)をN3(*)に残す」という文型パターンにおいて、N2に該当する名詞の意味属性が「*」という任意を表すものであったため、「メッセージ」が想像できなかった。もし、「メッセージ」が想像できていれば、このパターンの表す事態が「連絡をとる行為のきっかけ」であるとわかり、〈成行き〉による《期待》が想像できたと思われる。ゆえに、意味属性制約が任意である文型パターンについては、具体的な名詞を代入しながら再検討すべきである。

(誤り事例6) 結局2時間半ずっと立ちっぱなしだった。

- 正解の情緒:《嫌だ》6名
- 推定された情緒:《なし》

「立つ」という動作そのものでは、情緒生起原因とはならないが、「2時間半～ばなしだった」という表現から、長い間の継続的な動作であることが伺える。情緒主の疲労が読み取られて《嫌だ》というタグが付与されたと思われる。このように、時間幅を表すような副詞的な表現、および、テンス・アスペクトについての文末の表現が無視できないことが分かった。これは、用言や格要素以外の情報を使いこなすことの問題である。用言に対する知識ベースの不備によるものではない。

(誤り事例7) 午前中に突然お客さんが来た。

- 正解の情緒:《驚き》5名／少数意見:《なし》1名
- 推定された情緒:《なし》

「突然」という表現から、予想外の出来事であったことが分かる。「来た」という用言だけでなく、副詞も考慮する必要がある。その他に、「明日」などの時詞が出現している際、「期待」や「恐れ」という予測に関連する情緒を推定するほうが適切なことがある。今後、情緒推定で重点を置くところを、用言から読み取られる情緒であるのか、副詞から読み取られる情緒であるのかについて、検討を深める必要がある。

(誤り事例8) そう告白された。

- 正解の情緒:《驚き》6名
- 推定された情緒:《喜び》

「告白する」という行為は、対人関係の深さを変化させる効果があるため、「喜び」と推定されたことは不自然ではない。しかし、望まない相手の場合には「喜び」とはならず、この事例では文脈上、望まない相手であったため正解の情緒に「喜び」は挙げられていない。一方、「驚き」について、この行為が日常生活では頻度が多くないことより、「驚き」の情緒属性が付与されることは適切であるが、付与されていなかった。それは、「驚き」の特徴フレームにおいて〈対人関係〉の中位特徴がなかったためである。「驚き」の特徴フレームの弱点を補うために、対人関係に関する用言について見直しを行う必要がある。

7.4.3 誤り分析の結果

80 節についての誤り分析の結果を表 7.2 にまとめる。この表に示す割合は、実験対象とした節に対する割合である。

表 7.2 誤り分析の結果

#	不正解となった原因	割合 (該当節数 / 実験対象の節数)
1	内容(情緒主と他の格要素との関係, 日常性, 個人の嗜好性)が考慮されなかったため	16% (32 / 204)
2	主体的表現が考慮されなかったため	5% (11 / 204)
3	文脈(前後の文の情報)が考慮されなかったため	4% (9 / 204)
4	情緒の強さが考慮されなかったため	4% (8 / 204)
5	パターンへの情緒属性が不足していたため	3% (6 / 204)
6	テンス・アスペクト等が考慮されなかったため	2% (5 / 204)
7	副詞が考慮されなかったため	2% (4 / 204)
8	その他	2% (5 / 204)

この表より次のことが言える。

- 本知識ベースの不備による誤りは、#5 の 3%であった。
- 本方式の能力不足による誤りは、#1 の 16%であった。
- 本方式外のことが問題であった節は、21%であった。表 7.2 における#2 ～ 4 および#6 ～ 8 が該当する。

こうして、本研究で提案した手法は良好に動作しており、今後、格要素の関係などに改良を加えればよいことが分かった。そして、構築した知識ベースは誤り率が 3%であることから、その信頼性が確認できた。

7.5 まとめ

日記文を使用した情緒推定実験によると、正解の情緒と本システムの出力の情緒の一致率が 45%であった。本知識ベースの手動運用による情緒推定の場合、一致率が 48%であったことから、性能の低下は大きくなく、本システムの動作は良好であったと言える。なお、初心者による情緒推定の一致率が 49%であることから、初心者と比べて約9割の性能である。

本研究では、情緒生起原因に基づくタイプ1の情緒推定の方法を、言語の表層構造解析結果を基に実現することを目指していた。以上の実験により、この目標が達成できたと言える。

第8章 検討

情緒生起原因に焦点を当てたタイプ1の情緒推定を、パターン言語処理に基づく方式で実現した。この方式は、基本的な言語処理に対して、情緒的な意味情報を提供できることを狙っている。そこで、本章では、省略解析に対して情緒的な意味情報が提供できること、および、その意味情報が利用できることを簡単な例の範囲で示す。

8.1 省略解析の原理

文章の理解は、文と文の間の整合性があることが前提となっている。情緒についての整合性も当然前提に含まれる。ある文での情緒状態と、文脈上関わりのある文での情緒状態は、断りのない限り変化しないものと予想される。そこで、情緒状態の整合性と、情緒主や情緒対象などの情緒属性を根拠として、省略表現の解析ができるのではないかと予想される。

図 8.1 を参照しながら、情緒推定を省略解析に応用する過程について説明する。この図には2文で構成される文章がある。1文目「マウンテンバイクが当たった。」に対して情緒推定を行うと、「所有的移動」と「結果」という2通りの意味解析の結果が得られ、前者には「獲得による喜び」という情緒推定が行われ、後者には「外的な不快による悲しみ・嫌だ・怒り」という情緒推定が行われる。一方、2文目「嬉しい！」に対しては、情緒を直接表現する用言を収録した辞書[77]から「喜び」が推定される。感情形容詞であるので、情緒主は、デフォルト値として「話し手」となる。

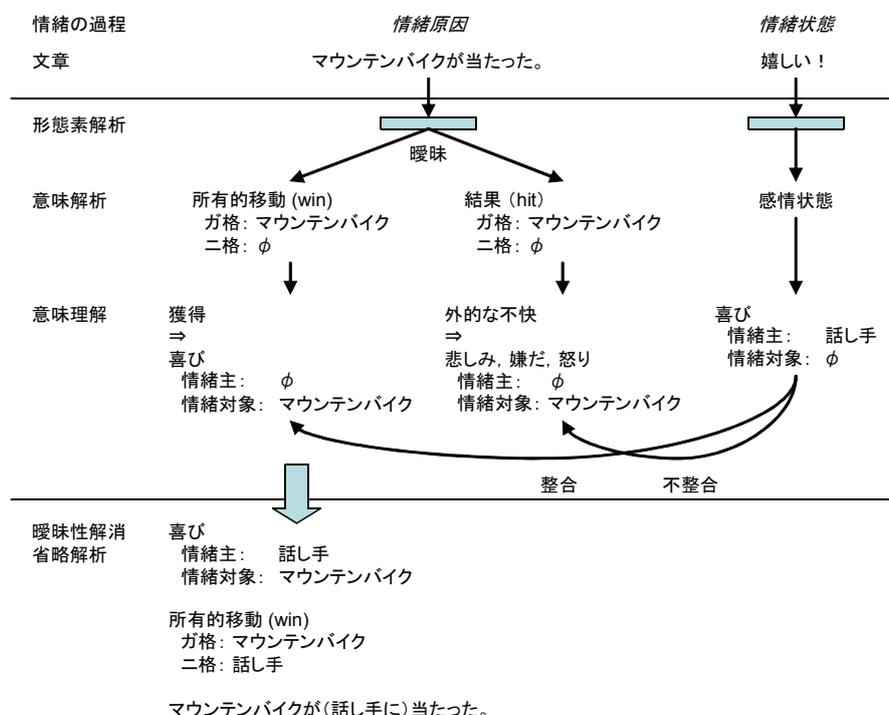


図 8.1 情緒推定を省略解析に使用する様子

ここで、情緒についての整合性があるはずなので、「獲得による喜び」となる意味解析が優勢となる。整合性は、情緒名のみならず、情緒主や情緒対象にも及ぶので、マウンテンバイクが当たったのは「話し手」に対してであることが解析される。

8.2 省略解析の方法

省略解析を次の手順で行うことを検討する。文章を節単位に分解して、入力するものとする。

Step-1: 各節において、情緒推定を行う。

Step-2: 情緒属性から、推論規則を生成する。

Step-3: 省略解析をする節について推論を実行する。

Step-1 の情緒推定は、提案方式を用いる場合と、Dyer のように語彙辞書を用いる場合がある。情緒原因を表す節に対しては、提案方式による推定が情緒推定の結果を出力し、情緒状態を表す節に対しては、語彙辞書による推定が結果を出力する。

Step-2 の推論規則は、情緒の整合性を検査するための推論規則であり、「規則」と「ファクト」の集合である。1文が支持する関係が1つの「規則」に対応し、1文が確定させる事柄が「ファクト」に対応する。たとえば、「マウンテンバイクが当たった」に対しては曖昧性があり、「獲得」ならば「喜び」であり、「外的な不快」ならば「悲しみ」などである。

Step-3 の推論実行では、Step-2 で生成した推論規則を用いて、節の曖昧な語義解析結果について質問する。たとえば、「マウンテンバイクが当たった」に対しては「誰が何を獲得したのか」、あるいは、「誰が外的な不快なのか」という質問を論理形式で行う。

8.3 解析の具体例

次の日記文において省略解析を試みる。ここで、節2において誰にマウンテンバイクが当たったのか、節5において誰に何を配達してくれるのかという点が表記されていない。情緒推定を利用してこれらを解析する。

- 1> 数ヶ月前に雑誌の懸賞に応募した。
- 2> 今日、マウンテンバイクが当たったという
- 3> 通知が届いた。
- 4> マジで？って感じ！
- 5> 来週末に配達してくれるということだった。
- 6> めちゃくちゃ嬉しい！

(日記文は、[64]より引用)

図 8.2 解析対象のテキスト

8.3.1 Step-1 情緒推定の実行

図 8.2 の各節について情緒推定を行う。図 8.3 に具体的な結果を示す。「EC」で始まる ID は、情緒原因に着目した情緒推定の結果であることを表す。ID のあとに、情緒名、そのあとに、情緒原因の名称、情緒主、情緒対象を出力する。「no match」は適合パターンの無いことを表す。情緒主と情緒対象が、具体的に表現されていない場合、該当する格要素の名称を出力する。

「直接表現」は情緒状態を表す表現から、情緒を推定したことを表す。その結果、同様に情緒名と情緒対象を出力する。この推定の方法は、本研究で示したパターン言語処理の方式で実現可能

である。異なる点は使用する知識ベースである。具体的には, [77]で構築した知識ベースを使用することで情緒推定が可能である。

1>	数ヶ月前に雑誌の懸賞に応募した。				
e1.1	EC001827-00	期待	その他	ガ格	懸賞
2>	今日、マウンテンバイクが当たったという				
e2.1	EC000382-00	喜び	獲得	ニ格	マウンテンバイク
e2.2	EC000387-00	悲しみ	外的な不快	ニ格	マウンテンバイク
e2.3	EC000411-00	嫌だ	外的な不快	ニ格	マウンテンバイク
e2.4	EC000411-01	怒り	外的な不快	ニ格	マウンテンバイク
3>	通知が届いた。				
e3.1	EC0009694-00	喜び	獲得	ニ格	通知
e3.2	EC0009695-00	なし			
4>	マジで?って感じ!				
e4.1	no match				
5>	来週末に配達してくれるということだった。				
e5.1	EC011073-00	好ましい	協力	ニ格	ガ格
e5.2	EC011073-01	喜び	獲得	ニ格	ヲ格
6>	めちゃくちゃ嬉しい!				
e6.1	直接表現	喜び		話し手	ガ格

図 8.3 情緒推定の結果

8.3.2 Step-2 推論規則の生成

情緒状態からの推定結果からは、「情緒」の「ファクト」として生成する。情緒名をファクト名とし、情緒主と対象を引数とする。下記の例では、感情形容詞であるので、情緒主はデフォルトで「話し手」であると判定した。

(例) 「めちゃくちゃ嬉しい!」(情緒状態の表現. [77]より判定可能)

生成したファクト: `glad('話し手', _)` .

情緒原因からの推定結果からは、「原因 → 情緒」という「規則」として生成する。原因名は、約 120 種類であり、形式が固定されている。したがって、変換テーブルを参照することで、原因名から論理の述語の記述を生成することができる。「原因 → 情緒」という規則にした理由は、情緒の生起が確認されるとき、情緒原因の発生があったことが真であると推論されるためである。

(例) 「通知が届いた。」(情緒原因の表現. 本研究の方式で判定可能)

生成した規則: `acquire(X, Y) :- glad(X, Y), Y='通知'.`

図 8.3 から生成した規則とファクトを図 8.4 に示す。

```

% e1

% e2 (conflict)
acquire(X,Y) :- glad(X,Y), Y = 'マウンテンバイク'.
surface_displeasure(X,Y) :- sadness(X,Y), Y = 'マウンテンバイク'.
surface_displeasure(X,Y) :- disliking(X,Y), Y = 'マウンテンバイク'.
surface_displeasure(X,Y) :- anger(X,Y), Y = 'マウンテンバイク'.

% e3 (conflict)
acquire(X,Y) :- glad(X,Y), Y = '通知'.

% e4

% e5 (combination)
support(X,Y) :- liking(Y,X).
acquire(X,Y) :- glad(X,Y).

% e6
glad('話し手',_).

```

図 8.4 推論規則の生成結果

8.3.3 Step-3 省略解析の実行

生成した規則とファクトを用いた推論を Prolog 上で実行する。「来週末に配達してくれるということだった。」に対して、「support(X,Y)」と「acquire(X,Y)」という2つの述語が対応している。ここで、誰に何が配達されるのかを調べるためには、「?- acquire(X,Y).」という質問をする。実行結果を図 8.5 に以下に示す。

```

1 ?- acquire(X,Y).

X = '話し手',
Y = 'マウンテンバイク';

X = '話し手',
Y = '通知';

X = '話し手'
2 ?-

```

図 8.5 省略解析の実行例

図 8.5 には、解析結果が3通り示されている。つまり、省略解析の結果は次のようになる。

- 解1。「来週末に**マウンテンバイクを話し手に**配達してくれるということだった。」
- 解2。「来週末に**通知を話し手に**配達してくれるということだった。」
- 解3。「来週末に**φを話し手に**配達してくれるということだった。」

ここでは、解の選択が問題である。その解決には情緒以外の情報による絞り込みも必要である。

たとえば、「通知」は既に届いているので、「来週末に通知を配達してくれる」というのは矛盾している、という判断が必要である。

省略解析は、情緒の意味情報だけで解決できるものではないが、有効な候補を挙げることができると言える。

8.4 情緒の意味情報のバリエーション

上述の例では、情緒状態から推定される情緒属性を事実とすることで、容易に推論の成立することが確かめられた。しかし、情緒状態を明示する表現は、文章中で多く存在する訳ではないため、情緒の意味情報を得るためにより広範囲に情緒を扱う知識が必要である。以下にその具体例を示す。

まずは、情緒原因のみで構成される文章や情緒反応を含む文章における省略解析の手順を例示しよう。これらでは、「原因→情緒」や「反応→情緒」という規則が作られるのだが、1つを仮説として取り上げる。すなわち、情緒を仮の事実として、推論を行うことで、本章での手順で省略解析が可能であると予想される。

例1. 「マウンテンバイクが当たった。痛かった。」

情緒推定：痛い —(原因)→ 外的な不快による悲しみ・嫌だ・怒り

仮説：悲しみ（これを仮説とするので、Prolog には事実として登録する。）

例2. 「マウンテンバイクが当たった。許せない。」

情緒推定：許せない —(反応)→ 怒り

仮説：怒り

また、タイプ2の情緒推定として文末表現を参照することが考えられる。文末表現は、話し手の情緒を表す可能性があるが、情緒状態を明示する場合ほど確定的な証拠とは言い難い。そのため、上述と同じく仮説を立てた省略解析とすることが望ましいであろう。以下に具体例を挙げる：

例3. 「マウンテンバイクが当たっちゃったよ〜！！」

情緒推定：「ちゃったよ〜！！」—(文末表現)→喜び

仮説：喜び

さらに、文章の上で、情緒の整合性が通常通りではないことを表す場合、すなわち、接続表現のある場合について例を示す：

例4. 「マウンテンバイクが当たった。でも、嬉しくない。」

情緒推定：嬉しくない—(直接表現)→《喜び、否定》

整合性の検査：「でも」より前の部分では、後の部分の情緒が反転する

仮説：喜び（前の部分から生成した「規則と事実」に対して、反転した情緒、すなわち、《喜び》を事実として登録する。）

「でも」の後の節では、「嬉しくない」という情緒状態を表す表現があることから、「喜び」が否定されている。「でも」の前の節では、逆説であるため「喜び」として解釈されることが通常であることをこの文章の話し手は意識していたと読み取ることができる。つまり、「マウンテンバイクが当たった」という表現に対して「喜び」が伴うという捉え方で、規則を生成する必要がある。

8.5 先行研究との比較

Dyer は、感情語彙(Affect Lexicon)とゴールの関係を解析する方法を提案した。そこでの問題点は次の2点である。

- ゴールなど非情緒的な情報の解析にコストがかかる(すなわち、深い言語情報を使用)。
- それらの情報の内部表現の設計が難しい。

これに対して、本研究の手法では次の利点がある。

- 表層的な言語情報から情緒原因を抽出し情緒を推定するため、解析コストが安い。
- 情緒原因が約 120 種類にカテゴリ分けされており、推論規則を作成する上での内部表現の種類に目処が立っている。

情緒原因を抽出する際、幾つかの候補が生じるのだが、Step-2 において推論規則にそれらを含ませることにより、Step-3 での最終的な推論過程により整合性のとれるものが選択されるので、情緒的に辻褃の合う語彙解析が行われていることになる。言語理解過程の曖昧性を無理に解消するのではなく、解析に必要な情報の得られた段階で解消する方法が成り立つ。また、内部表現の種類について、第 5.3 節で示したように、基本的な日本語の用言の語義に対してそのカテゴリで網羅的に説明できていることから、目処が立っているといえる。これらは、Dyer の問題を回避している。

一般の省略・照応解析において、シソーラスや格関係など非情緒的な意味情報が参照されていたり、統計的・確率的な情報が参照されているが、本節での手法は情緒的な意味情報も使用可能なことを提案している。

第9章 結論

本研究は、計算機による言語意味理解のための基礎技術の1つとして、言語表現から話者や登場人物の情緒を推定する方式を提案した。そして、情緒推定に必要な言語知識ベースを構築するとともに、提案方式の精度を実験的に確認した。

まず、情緒推定の定義と基礎技術としての位置付けを次のように設定した。「言語表現からの情緒推定」の定義を、「与えられた言語表現から、ある人物(情緒主)の情緒を(推定者が)読み取ること」とした。言語理解の基礎技術として、言語理解の基本的な構造解析のタスクに、情緒的な内容の意味的情報を提供することを目指した。このことにより、表層レベルの言語的情報を起点として情緒推定が行えること、および、情緒属性(情緒の種類、情緒主、情緒対象、情緒の原因)が出力できることが必要となった。

次に、情緒推定の方法、評価尺度、および、実現方法について提案した。

情緒推定の方法については、情緒の過程(情緒生起原因の発生、情緒状態、情緒反応の発生)、情緒主(登場人物、話し手、聞き手、傍観者)、解釈対象の種類(客体的表現、主体的表現、文脈、対話性、物語性)の組により、その種類を次の5つのタイプに分類した。

タイプ1: 客体的表現に基づき、登場人物の情緒を推定する方法

タイプ2: 主体的表現に基づき、話し手の情緒を推定する方法

タイプ3: 文脈に基づき、登場人物・話し手の情緒を推定する方法

タイプ4: 対話性に基づき、話し手・聞き手の情緒を推定する方法

タイプ5: 物語性に基づき、傍観者の情緒を推定する方法

これらのタイプのうち、タイプ1と2が基礎的な方法である。先行研究においては、情緒生起原因を対象とするタイプ1の方法について、表層レベルの言語的情報から情緒推定する方式が確立できておらず、そのための網羅的な言語知識ベースの開発が実現できていないことが分かった。

評価尺度については、人間による情緒推定の能力を参考にすることにした。まず、情緒推定は主観的判断と考えられがちである。しかし、言語処理では人々が共通的に理解することを機械処理で実現することが1つの目標であると言えるので、本研究では「複数の推定者に共通的に読み取られる情緒」を情緒推定の正解とした。そこで、計算機による情緒推定を追求する前に、人間による情緒推定の能力を、情緒タグ付きコーパスの構築を通じて調査した。テキスト対話、および、日記文に、複数の推定者により情緒タグを付与する作業を行い、情緒タグの共通性、すなわち、推定者間での情緒タグの一致率について集計した。その結果、表情付きテキスト対話においては、一致率が 65.2% (κ 値=0.444 ~ 0.546) であり、表情なしテキスト対話においては、一致率が 60.5% ~ 58.0% (κ 値=0.348 ~ 0.472)、そして、日記文においては、一致率が 52.3% (κ 値=0.318 ~ 0.416) であることが分かった。さらに、日記文において、情緒の過程の種類ごとに、出現頻度を調査した。その結果、日記文の節(1,642 節)のうち、情緒生起原因を表す節は 68.3%、情緒状態を表す節は 15.7%、情緒反応を表す節は 5.1%、その他の節は 11.4% であることが分かった。

以上のことから、情緒生起原因を対象とするタイプ1の方法の実現とそのため言語知識ベースの構築を行うことが重要であることが分かった。

情緒推定を実現するために、「非線形言語モデル」に基づく「パターン言語処理」を応用した情緒推定方式を提案した。この方式は、あらかじめ文型パターンと情緒の関係を対応付けた知識ベースを構築しておき、実行時に入力文と適合した文型パターンから情緒を推定する方式である。この方式を実現する上で、次の課題があった：

- 文型パターンの記述子の設計
- 文型パターンと表現の照合方式の開発
- 文型パターンへの情緒の関係の対応付け

文型パターンの記述子を設計するにあたり、表層レベルの言語情報を使用することが前提であるため、与えられる言語表現は、形態素情報であり、構文情報ではない。そのため、文型パターンと言語表現の照合において不適合となりやすかった。そこで、文型パターンの側に柔軟性を持たせるために、離散記号、任意記号、時制・相・様相の関数などの記述子を導入した。その結果、離散記号・任意記号の導入の前と後を比較すると、文型パターンの数に換算して約 20 倍に相当する適合の性能をこれらの記号がもたらすことが確認された。また、意味的に正しくパターンに適合させるための意味属性制約について、「A の B」型名詞句の日英機械翻訳実験に利用したところ良好な精度の得られたことより、意味属性制約の効果が確認された。この実験において半自動的に意味属性制約を与えることができた点は、今後の応用に役立つと言える。

一方、複雑な記述子で構成される文型パターンに対して、与えられる言語表現を照合する方式が問題となった。そこで、数十万件規模で文型パターンを収録した知識ベースから、与えられる言語表現に適合するものを全て検索するツール *SPM* を開発した。これは *ATN* を基本とする方式であるが、絞り込み前処理、join 動作、履歴機能という3つの高速化を施すことにより、実効的には言語表現の長さ n に対して $O(n) \sim O(n^2)$ で動作し、ページングとしては良好に動作する部類に属することを確認した。

文型パターンに情緒の関係を対応付けることで知識ベースを構築した。日本語の基本的な用言を網羅するために、「日本語語彙大系」の文型パターンを参照し、「情緒名」、「原因」、「情緒主」、「情緒対象」という情緒属性を対応付けることにした。この作業は抽象度が高いため全て手作業で行うことにした。ここで、情緒原因となる事態を表す文型パターンに対して対応付けを行うのだが、情緒原因を判定する基準が必要となった。

そこで、情緒生起原因についての概念分析を行った。分析は、ボトムアップ分析とトップダウン分析で行った。ボトムアップ分析では、約 1,000 件の情緒生起に関する事象を分類し、「生理的な快」、「獲得」、「反抗」などの特徴を見出した。トップダウン分析では、心理学や認知科学における文献を参考にして、その特徴の階層構造の見直しと構成要素の補充を行った。その結果、《喜び》、《悲しみ》、《好ましい》、《嫌だ》、《驚き》、《期待》、《恐れ》、《怒り》という 8 つの基本情緒に対して、125 個の最下位層の特徴、51 個の中位層の特徴、および、9 個の最上位層の特徴を得た。

日本語語彙大系の文型パターン(14,819 個)の 1 件 1 件に対して、上記の情緒生起原因についての特徴を参照しながら、手作業で情緒属性を対応付けた。この作業は、3 回繰り返し行った。2 回目の対応付け作業の後、検査者が情緒生起特徴を参照しながら総検査を行ったところ誤り率は 14.2%であったが、3 回目の対応付けの際に全て修正した。こうして知識ベースを構築した。本知

識ベースの 52.2%のパターンには情緒生起原因が対応付けられ、残りのパターンは該当しないことが分かった。

こうして構築した言語知識ベースを使用して、パターン言語処理に基づく情緒推定の方式を実装した。日本語語彙大系の文型パターンは、記述子に柔軟性が無いので、離散記号や任意記号の導入により柔軟性を高めた。また、入力文とパターンの照合における曖昧性に対して、7つの特徴量に基づくヒューリスティクスで解決を行った。

動作実験では、情緒タグ付き日記コーパス(1,642 節)を使用して、本知識ベースの網羅性、および、情緒推定の精度を評価した。網羅性について、本方式で処理すべき節(331 節)のうち、91%は本知識ベースの文型パターンが適合したことより確認できた。精度について、情緒推定の初心者が 49%の一致率であったことに対して、本知識ベースを手動で運用した場合が 48%、自動で運用した場合が 45%であった。さらに、誤り分析によると、本知識ベースの対応付け誤りによる誤り率が 3%であることから本知識ベースの精度が確認された。本方式の不備について、格要素の解析に依る誤り率が 16%、曖昧性解消に依る誤り率が 8%であった。格要素の問題は既存の手法により解決策が考えられる。一方、その他の誤りについては、用言以外の言語情報(副詞、時制・相・様相の表現、文脈など)を併用する必要のあるものであり、誤り率が 21%であった。実験結果より、本研究の提案方式は、通常の人間に匹敵する情緒推定の能力を持つことが確認できた。また、構築した言語知識ベースは、精度の良いことが確認できた。

最後に、情緒的な意味情報が語義解析や省略解析に使用可能であることを簡単な例を通じて示した。本研究の方法は、情緒主の登場する言語表現に限定しているため、事態を表すための中間的な内部表現の問題が回避され、かつ、プラン認識などの深い言語理解を使用せずに実現できた。試行的な段階であるが、情緒的な意味情報を言語理解の基本的な処理タスクに提供できることを示すことができた。

以上により、表層レベルの言語的情報から情緒を推定する方法を確立することができた。そして、そのための言語知識ベースを日本語の基本的な用言の範囲において網羅的に構築することができた。深い言語理解を行うために、従来研究では、情緒状態を表す表現について情緒的な意味情報が使用されていたのだが、本研究の方式により、情緒原因を表す表現というより広い範囲にまで情緒的な意味情報が提供可能となった。従って、新しい言語意味理解の基礎技術を開発することができたと言える。

謝辞

本研究を進めるにあたり御指導下さいました鳥取大学池原悟教授に心より御礼申し上げます。概念分析をはじめとする研究の基礎を御指導下さいました九州工業大学岡田直之名誉教授に心より御礼申し上げます。両先生の下で研究の職を授かることができたことは幸運の極みです。

本論文をまとめるにあたり御指導下さいました九州工業大学遠藤勉教授，鳥取大学増山博教授，鳥取大学菅原一孔教授，鳥取大学村上仁一准教授に深く感謝致します。

本研究は，筆者が九州工業大学に在籍していた頃(同大学大学院博士前期課程 1993 年)より進めました。特に，日頃より励まして下さいました乃万司教授ならびに甲斐郷子助手(現在，同大学情報科学センター准教授)に深く感謝致します。種々の最新情報技術関連のイベントに起用して下さいました中村順一教授に深く感謝致します。研究面において，筆者が井の中の蛙とならぬよう身近な視点から助言を下さいました乾健太郎助教授(現在，奈良先端科学技術大学院大学准教授)に深く感謝致します。計算機管理および研究室運営にご協力下さいました松元隆二技官，および，荻野友子秘書に感謝致します。研究室において試作された多くの分析データを本研究で使用しています。その作成に携わりました知能工学講座のメンバー，とりわけ，吉田太輔先輩，徳久良子君，田川峰和君に感謝致します。

鳥取大学に着任(2002 年 4 月)してからは，研究室運営にご協力下さいました山本理絵秘書に感謝致します。研究室において試作された多くの分析データを本研究で使用しています。その作成に携わりました計算機工学講座のメンバー，とりわけ，田中努君，守谷有司君，安井敏君，松浦大樹君，岡本一輝君，金澤佑哉君，石上真理子君，松浦祥悟君，山形亮君，平井雅幸君に感謝致します。漫画コーパスの原文を入力して下さいました鳥取シルバー人材センター田中勝弘氏・東弘之氏に感謝致します。大規模パターン辞書の作成に関して CREST プロジェクトを通じてご協力頂きました NTT アドバンステクノロジーの方々，とりわけ，技術会議で筆者の意見を汲み取って下さいました阿部さつき氏，小船園望氏，竹内奈央氏に感謝致します。また，本論文をまとめる終盤において，筆者の研究班の運営と本論文の細部に渡るチェックに協力して下さいました本講座の黒住亜紀子君，前田浩佑君，水田理夫君に感謝致します。

本研究ではとりわけ「日本語語彙大系(岩波書店)」，「ちびまる子ちゃん(集英社)」，「英語で日記を書いてみる(ベレ出版)」の図書を研究対象とし，かつ，本論文および発表文献において引用致しました。著作者の方々へ敬意を表します。それから，ツールに関して，鳥取大学池原研究室と NTT の間での契約の下で日本語形態素解析器 ALT-JAWS を利用させて頂きました。また，奈良先端科学技術大学院大学松本研究室より公開されている日本語形態素解析器 chasen および日本語構文解析器 cabocha の両一式を利用させて頂きました。Prolog 処理系では SWI-Prolog を使用させて頂きました。これらの関係者の方々に敬意を表します。

参考文献

- [1] 黒橋禎夫: “言語コンピューティング.” 人工知能学会誌, Vol.22, No.5, pp.711-718, 2007.
- [2] 河原大輔, 黒橋禎夫: “自動構築した格フレーム辞書と先行詞の位置選好順序を用いた省略解析.” 自然言語処理, Vol.11, No.3, pp.3-19, 2004.
- [3] 飯田龍, 乾健太郎, 松本祐治, 関根聡: “最尤先行詞候補を用いた日本語名詞句同一性指示解析.” 情報処理学会論文誌, Vol.46, No.3, pp.831-844, 2005.
- [4] Pfeifer, R. and Nicholas, D.W.: “Toward Computational Models of Emotion.” Selected and Updated Papers from the Proceedings of the 1982 European Conference on Progress in Artificial Intelligence, pp.184-192, John Wiley & Sons, 1985.
- [5] Frijda, N.H. and Swagerman, J.: “Can Computer Feel? Theory and Design of an Emotional System.” *Cognition and Emotion*, Vol.1, No.3, pp.235-257, 1987.
- [6] Elliott, C.: “The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System.” Ph.D. Thesis, The Institute for the Learning Sciences Technical Report #32, Northwestern University, 1992.
- [7] Dyer, M.G.: “In-Depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension.” The MIT Press, 1983.
- [8] 目良和也, 市村匠, 相沢輝昭, 山下利之: “語の好感度に基づく自然言語発話からの情緒生起手法.” 人工知能学会論文誌, Vol.17, No.3, pp.186-195, 2002.
- [9] Liu, H., Lieberman, H. and Selker, T.: “A Model of Textual Affect Sensing Using Real-World Knowledge.” IUI '03: Proceedings of the 8th International Conference on Intelligent User Interfaces, pp.125-132, 2003.
- [10] 池原悟: “言語で表現される概念と翻訳の原理.” 電子情報通信学会技術研究報告, TL2003-25, pp.7-12, 2003.
- [11] Narayanan, S.: “Towards Modeling User Behavior in Human-Machine Interactions: Effect of Errors and Emotions.” ISLE Workshop on Multimodal Dialog Tagging, 2002.
- [12] Litman, D. and Forbes, K.: “Recognizing Emotions from Student Speech in Tutoring Dialogues.” ASRU '03: IEEE Workshop on Automatic Speech Recognition and Understanding, pp.25-30, 2003.
- [13] 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: “日本語彙大系.” 岩波書店, 1997.
- [14] 独立行政法人情報通信研究機構: “EDR電子化辞書.” 2002.
- [15] 黒橋禎夫, 長尾眞: “京都大学テキストコーパス・プロジェクト.” 言語処理学会第3回年次大会発表論文集, pp.115-118, 1997.
- [16] 山下俊郎, 松山義則, 高野清純, 宮本美沙子, 原野広太郎, 高木俊一郎: “情緒・欲求・動機.” 児童心理学講座, Vol.6, 1969.
- [17] 伊藤正男, 梅本守, 山鳥重, 小野武年, 往住彰文, 池田謙一: “情動.” 岩波講座 認知科学, Vol.6, 1994.
- [18] Simon, H.A.: “Comments.” *Affect and Cognition*, The 17th Annual Carnegie Symposium on

- Cognition, Lawrence Erlbaum, pp.333-342, 1982.
- [19] Fiske, S.T. and Taylor, S.E.: "Social Cognition." 2nd ed. McGraw-Hill, 1991.
- [20] 中島義明, 安藤清志, 子安増生, 坂野雄二, 繁榊算男, 立花政夫, 箱田裕司(編): "心理学辞典." 有斐閣, 1999.
- [21] 岡田直之: "自然言語理解には情緒的機能も必要." 電子情報通信学会誌, Vol.70, No.9, pp.897-902, 1987.
- [22] Shanahan, J.G., Qu, Y. and Wiebe, J. (eds.): "Computing Attitude and Affect in Text: Theory and Applications." Springer, 2006.
- [23] 電子情報通信学会: "エンサイクロペディア電子情報通信ハンドブック." オーム社, p.147-152, 1998.
- [24] Picard, R.W.: "Affective Computing." MIT Media Laboratory, Perceptual Computing Section Technical Report, No.321, 1995.
- [25] Lehnert, W.G.: "Plot Unit and Narrative Summarization." *Cognitive Science*, Vol.4, pp.293-331, 1981.
- [26] P.エクマン, W.V.フリーゼン: "表情分析入門." 工藤力(訳), 誠信書房, 1987.
- [27] André, E., Klesen, M., Gebhard, P., Allen, S. and Rist, T.: "Integrating Models of Personality and Emotions into Lifelike Characters." *Affective Interactions*, Lecture Notes in Artificial Intelligence, Vol.1814, pp.150-165, Springer, 2000.
- [28] Plutchik, R.: "The Multifactor-Analytic Theory of Emotion." *Psychology*, Vol. 50, pp.153-171, 1960.
- [29] Plutchik, R.: "Emotions and Life: Perspectives from Psychology, Biology, and Evolution." American Psychological Association, 2002.
- [30] Ortony, A., Clore, G.L. and Collins, A.: "The Cognitive Structure of Emotions." Cambridge University Press, 1988.
- [31] ランドルフ・R・コーネリアス: "感情の科学 -心理学は感情をどこまで理解できたか." 齊藤勇(監訳), 誠信書房, 1999.
- [32] 福田正治: "感情を知る 感情学入門." ナカニシヤ出版, 2003.
- [33] Elliott, C., Rickel, J. and Lester, J.: "Lifelike Pedagogical Agents and Affective Computing: An Exploratory Synthesis." *Artificial Intelligence Today*, Lecture Notes in Computer Science, Vol.1600, pp.195-212, Springer, 1999.
- [34] Reilly, W.S.: "Believable Social and Emotional Agents." Ph.D. Thesis, Technical Report CMU-CS-96-138, School of Computer Science, Carnegie Mellon University, 1996.
- [35] Mueller, E.T.: "Story Understanding through Multi-Representation Model Construction." *Proceedings of the HLT-NAACL 2003 Workshop on Text Meaning*, pp.46-53, 2003.
- [36] Riloff, E.: "Information Extraction as a Stepping Stone toward Story Understanding." *Understanding Language Understanding: Computational Models of Reading*, pp.435-460, MIT Press, 1999.
- [37] 大野晋, 浜西正人: "類語新辞典." 角川書店, 1981.
- [38] 柴田武, 山田進: "類語大辞典." 講談社, 2002.
- [39] 中村明: "感情表現辞典." 東京堂出版, 1993.

- [40] 飛田良文, 浅田秀子: “現代形容詞用法辞典.” 東京堂出版, 1991.
- [41] 飛田良文, 浅田秀子: “現代副詞用法辞典.” 東京堂出版, 1994.
- [42] 飛田良文, 浅田秀子: “現代擬音語擬態語用法辞典.” 東京堂出版, 2002.
- [43] 東弘子: “感情表出文.” 自然言語処理, Vol.6, No.4, pp.45-65, 1999.
- [44] 寺村秀夫: “日本語のシンタクスと意味.” Vol.1, くろしお出版, 1982.
- [45] 松本和幸, 三品賢一, 任福継, 黒岩眞吾: “感情生起事象文型パターンに基づいた会話文からの感情推定手法.” 自然言語処理, Vol.14, No.3, pp.239-271, 2007.
- [46] グループ・ジャマシイ: “日本語文型辞典.” くろしお出版, 1998.
- [47] 時枝誠記: “国語学原論.” 岩波書店, 1941.
- [48] 遠藤大介, 齋藤真実, 山本和英: “係り受け関係を利用した感情生起表現の抽出.” 言語処理学会第12回年次大会発表論文集, pp.947-950, 2006.
- [49] ダイアン・ブレイクモア: “ひとは発話をどう理解するかー関連性理論入門ー.” 武内道子, 山崎英一(訳), ひつじ書房, 1994.
- [50] 土屋誠司, 吉村枝里子, 渡辺広一, 河岡司: “状況表現を考慮する発話文章からの感情判断手法の提案.” FIT2006 情報技術フォーラム, 一般講演論文集, 第2分冊, pp.231-234, 2006.
- [51] 南不二男: “現代日本語文法の輪郭.” 大修館書店, 1993.
- [52] 小林のぞみ, 乾健太郎, 松本祐治, 立石健二, 福島俊一: “意見抽出のための評価表現の収集.” 自然言語処理, Vol.12, No.2, pp.203-222, 2005.
- [53] 那須川哲哉, 金山博: “文脈一貫性を利用した極性付与評価表現の語彙獲得.” 情報処理学会研究報告, 自然言語処理, NL-162-16, pp.109-116, 2004.
- [54] Rosis, F. and Grasso, F.: “Affective Natural Language Generation.” *Affective Interactions, Lecture Notes in Artificial Intelligence*, Vol.1814, pp.204-218, Springer, 2000.
- [55] 松本和幸, 任福継, 黒岩眞吾: “語の意味情報を考慮した感情推定アルゴリズム.” 言語処理学会第11回年次大会発表論文集, pp.145-148, 2005.
- [56] 往住彰文: “認知的感情の構造と文学テキスト理解.” *Cognitive Studies*, Vol.8, No.4, pp.369-375, 2001.
- [57] 重野純: “感情認知における視聴覚情報の統合と文化の影響.” 視聴覚研究会資料, Vol.35, No.8, pp.507-512, 2005.
- [58] 乾孝司, 奥村学: “テキストを対象とした評価情報の分析に関する研究動向.” 自然言語処理, Vol.13, No.3, pp.201-241, 2006.
- [59] Okada, N. and Endo, T.: “Story Generation based on Dynamics of the Mind.” *Computational Intelligence*, Vol.8, No.1, 123-160, 1992.
- [60] 飯田仁, 有田英一: “4階層プラン認識モデルを使った対話の理解.” 情報処理学会論文誌, Vol.31, No.6, pp.810-821, 1990.
- [61] Cohen, J.: “A Coefficient of Agreement for Nominal Scales.” *Education and Psychological Measurement*, Vol.20, No.1, pp.37-46, 1960.
- [62] 徳久良子, 寺寫立太: “雑談における発話のやりとりと盛り上がりとの関連.” 人工知能学会論文誌, Vol.21, No.2, pp.133-142, 2006.
- [63] さくらももこ: “ちびまる子ちゃん.” Vol.1-10, 集英社, 1987-1993.
- [64] 石原真弓: “英語で日記を書いてみる 表現集編.” ベレ出版, 2002.

- [65] 三浦つとむ: “日本語はどういう言語か.” 講談社, 1976.
- [66] 池原悟: “自然言語処理の基本問題への挑戦.” 人工知能学会誌, Vol.16, No.3, pp.422-430, 2001.
- [67] Russell, S. and Norvig, P.: “エージェントアプローチ人工知能.” 古川康一(監訳), 共立出版, 1997.
- [68] Woods, W.A.: “Transition Network Grammars for Natural Language Analysis.” *Communications of the ACM*, Vol.13, No.10, pp.591-606, 1970.
- [69] Allen, J.: “Natural Language Understanding.” Addison-Wesley, 1995.
- [70] 池原悟: “意味類型パターン記述言語仕様書.” 日本語表現意味辞書等管理委員, 2007.
- [71] 池原悟, 阿部さつき, 徳久雅人, 村上仁一: “非線型な表現構造に着目した重文と複文の日英文型パターン化.” 自然言語処理, Vol.11, No.3, pp.69-95, 2004.
- [72] 中井慎司, 池原悟, 村上仁一: “階層構造を持つ属性の組とクラスで与えられる構造規則のクラス分類.” 言語処理学会第5回年次大会, pp.277-280, 1999.
- [73] Murray, E.J.: “Motivation and Emotion.” Prentice-Hall, 1964.
- [74] 戸田正直: “感情 一人を動かしている適応プログラム.” 東京大学出版会, 1992.
- [75] 岡田敏, 村上仁一, 徳久雅人, 池原悟: “多変量解析による最適文型パターンの選択方式.” 言語処理学会第11回年次大会発表論文集, pp.25-28, 2005.
- [76] 原真一郎, 村上仁一, 徳久雅人, 池原悟: “日英機械翻訳における多変量解析を用いた最適パターンの選択.” 言語処理学会第12回年次大会発表論文集, pp.268-271, 2006.
- [77] 黒住亜紀子, 徳久雅人, 村上仁一, 池原悟: “結合価パターン辞書からの情緒を明示する用言の知識ベース化.” 言語処理学会第13回年次大会発表論文集, pp.39-42, 2007.

研究業績一覧

主論文

- 1 Masato Tokuhisa, Naoyuki Okada: “A Conceptual Analysis of Emotional Words for an Intellectual, Emotional Agent.” Proceedings of Pacific Association of Computational Linguistics (PACLING1997), pp.307-315, 1997.
- 2 徳久雅人, 岡田直之: “パターン理解的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, Vol.39, No.8, pp.2440-2451, 1998.
- 3 Masato Tokuhisa, Ryoko Tokuhisa, Kentaro Inui, Naoyuki Okada: “Emotion Recognition in Dialogue.” G. Hatano, N. Okada and H. Tanabe (eds.), *Affective Minds*, pp.221-229, Elsevier, 2000.
- 4 徳久雅人, 守谷有司, 村上仁一, 池原悟: “意味属性の共起による「A の B」型名詞句の翻訳規則.” FIT2003, 情報科学技術フォーラム, 情報技術レターズ, Vol.2, pp.87-88, 2003.
- 5 Masato Tokuhisa, Tsutomu Tanaka, Satoru Ikehara, Jin'ichi Murakami: “Emotion Reasoning based on Valency Patterns - A Prototype Annotation of Causal Relationships.” Human and Artificial Intelligence Systems --- From Control to Autonomy, Proceedings of the 4th International Symposium on Human and Artificial Intelligence Systems (HART2004), International Series on Natural and Artificial Intelligence, Vol.1, pp.534-539, 2004.
- 6 Masato Tokuhisa, Kumiko Endo, Yuya Kanazawa, Jin'ichi Murakami, Satoru Ikehara: “Evaluation of Pattern Generalization Effect under Development of Pattern Dictionary for Machine Translation.” Proceedings of the Conference of Pacific Association for Computational Linguistics (PACLING2005), pp.311-318, 2005.
- 7 Masato Tokuhisa, Jin'ichi Murakami, Satoru Ikehara: “Construction and Evaluation of Text-Dialog Corpus with Emotion Tags Focusing on Facial Expression in Comics.” Bogdan Gabrys, Robert J. Howlett, Lakhmi C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Proceedings of 10th International Conference (KES2006), Lecture Notes in Artificial Intelligence, Vol.4253, pp.715-724, Springer, 2006.
- 8 徳久雅人, 村上仁一, 池原悟: “漫画における表情に着目した情緒タグ付きテキスト対話コーパスの構築.” 自然言語処理, Vol.14, No.3, pp.192-217, 2007.

参考論文

学術論文

- 1 池原悟, 阿部さつき, 徳久雅人, 村上仁一: “非線型な表現構造に着目した重文と複文の日英文型パターン化.” 自然言語処理, Vol.11, No.3, pp.69-95, 2004.
- 2 金出地真人, 徳久雅人, 村上仁一, 池原悟: “結合価文法による動詞と名詞の訳語選択能力の評価.” 自然言語処理, Vol.11, No.3, pp.149-164, 2004.
- 3 池原悟, 徳久雅人, 竹内(村本)奈央, 村上仁一: “日本語重文・複文を対象とした文法レベル文型パターンの被覆率特性.” 自然言語処理, Vol.11, No.4, pp.147-178, 2004.

国際会議

- 1 Naoyuki Okada, Kentaro Inui, Masato Tokuhisa: “An Affective Approach to Human-Friendly Dialogue Systems.” Proceedings of Pacific Association of Computational Linguistics (PACLING1999), pp.3-15, 1999.
- 2 Naoyuki Okada, Kentaro Inui, Masato Tokuhisa: “Towards Affective Integration of Vision, Behavior, and Speech Processing.” Proceedings of the IEEE International Conference on Integration of Speech and Image Understanding, pp.49-77, 1999.
- 3 Satoru Ikehara, Masato Tokuhisa, Jin'ichi Murakami, Masashi Saraki: “Development of Semantic Pattern Dictionary for Non-linear Structures of Complex and Compound Sentences.” Proceedings of the ESSLLI'05 Workshop on Empirical Challenges and Analytical Alternatives to Strict Compositionality, pp.79-98, 2005.
- 4 Satoru Ikehara, Masato Tokuhisa, Jin'ichi Murakami, Masashi Saraki, Takashi Ikeda, Masahiro Miyazaki: “Semantic Pattern Dictionary for Translating Non-Linear Structures of Complex and Compound Sentences.” Proceedings of the Conference of Pacific Association for Computational Linguistics (PACLING2005), pp.170-179, 2005.
- 5 Satoru Ikehara, Masato Tokuhisa, Jin'ichi Murakami, Masashi Saraki, Masahiro Miyazaki, Takashi Ikeda: “Pattern Dictionary Development Based on Non-Compositional Language Model for Japanese Compound and Complex Sentences.” Computer Processing of Oriental Languages. Beyond the Orient: The Research Challenges Ahead, Lecture Notes in Computer Science, Vol.4285, pp.509-519, Springer, 2006.
- 6 Satoru Ikehara, Masato Tokuhisa, Jin'ichi Murakami: “Analogical Mapping Method and Semantic Categorization of Japanese Compound and Complex Sentence Patterns.” Proceedings of the 10th Conference of the Pacific Association for Computational Linguistics (PACLING2007), pp.181-190, 2007.

図書・解説

- 1 乾健太郎, 徳久雅人, 徳久良子, 岡田直之: “感情の生起とその反応.” 特集:感情のモデルと工学的応用の動向, 日本ファジィ学会誌, Vol.12, No.6, pp.741-751, 2000.
- 2 Naoyuki Okada, Kentaro Inui, Masato Tokuhisa: “Integrated Mind: What is Needed and How to Design it.” A.Kent(ed.), *Encyclopedia of Library and Information Science*, Vol.68, pp.180-209, Marcel Dekker, 2000.

研究発表

- 1 杉野昇史, 村上浩規, 目良和也, 徳久雅人, 岡田直之: “インソップワールド - - - 対話処理に向けて.” 電子情報通信学会技術研究報告, SP94-66, NLC94-35, pp.25-32, 1994.
- 2 徳久雅人, 岡田直之: “情緒の生起過程へのパターン理解的接近.” 電子情報通信学会技術研究報告, TL95-8, pp.9-20, 1995.
- 3 徳久雅人, 岡田直之: “知的エージェントの情緒生起アルゴリズムの開発.” 電子情報通信学会技術研究報告, TL96-8, pp.13-24, 1996.
- 4 徳久雅人, 田川峰和, 乾健太郎, 岡田直之: “基本情緒の強さの要因分析.” 電子情報通信

- 学会技術研究報告, TL98-14, pp.17-24, 1998.
- 5 木村宏, 徳久雅人, 目良和也, 甲斐郷子, 岡田直之: “対話における相手意図の理解と応答のためのプランニング.” 電子情報通信学会技術研究報告, TL98-15, pp.25-32, 1998.
 - 6 徳久雅人, 貞光康, 乾健太郎, 岡田直之: “統合的知能エージェントにおける並列分散処理手法.” 電子情報通信学会技術研究報告, AI98-78, KBSE98-48, pp.15-22, 1999.
 - 7 徳久良子, 徳久雅人, 乾健太郎, 岡田直之: “相手の情緒を理解する対話システムの構築に向けて.” 電子情報通信学会技術研究報告, TL99-7, pp.13-20, 1999.
 - 8 徳久良子, 乾健太郎, 徳久雅人, 岡田直之: “言語コーパスにおける感情生起要因と感情クラスの注釈づけ.” 人工知能学会研究会資料, SIG-SLUD-A003-2, pp.9-16, 2001.
 - 9 徳久雅人, 中野育恵, 山下智之, 岡田直之: “情緒を加味した深いタスク指向の対話理解のためのルールベースの構築.” 電子情報通信学会技術研究報告, TL2001-25, pp.21-28, 2001.
 - 10 日巻正寛, 徳久雅人, 岡田直之: “情緒を加味した勧誘の対話プランニング.” 電子情報通信学会技術研究報告, TL2001-38, pp.25-32, 2002.
 - 11 道祖尾太祐, 村上仁一, 徳久雅人, 池原悟: “日英対訳パターンの自動抽出に向けて.” 情報処理学会研究報告, 2003-NL-153, pp.113-118, 2003.
 - 12 金出地真人, 徳久雅人, 村上仁一, 池原悟: “結合価文法による動詞と名詞の訳語選択能力の評価.” 情報処理学会研究報告, 2003-NL-153, pp.119-124, 2003.
 - 13 池原悟, 佐良木昌, 宮崎正弘, 池田尚志, 新田義彦, 白井諭, 村上仁一, 徳久雅人: “機械翻訳のための日英文型パターン記述言語.” 電子情報通信学会技術研究報告, TL2002-48, pp.1-6, 2003.
 - 14 池原悟, 阿部さつき, 徳久雅人, 村上仁一: “非線形な表現構造に着目した日英文型パターン化.” 情報処理学会研究報告, 2004-NL-160, pp.49-56, 2004.
 - 15 池原悟, 徳久雅人, 村本奈央, 村上仁一: “日本語重文・複文を対象とした文法レベル文型パターンの被覆率特性.” 人工知能学会研究会資料, SIG-SLUD-A303-10, pp.57-64, 2004.
 - 16 竹信伸介, 徳久雅人, 村上仁一, 池原悟: “動詞節に修飾された名詞句の係り受け解析.” 情報処理学会研究報告, 2004-NL-163, pp.83-89, 2004.
 - 17 池原悟, 徳久雅人, 村上仁一, 佐良木昌, 池田尚志, 宮崎正弘: “非線型な重文複文の表現に対する文型パターン辞書の開発.” 情報処理学会研究報告, 2005-NL-170, pp.157-164, 2005.
 - 18 池原悟, 阿部さつき, 竹内奈央, 徳久雅人, 村上仁一: “意味的等価変換方式のための重文複文の統語的意味的分類体系について.” 情報処理学会研究報告, 2006-NL-176, pp.1-8, 2006.
 - 19 徳久雅人, 村上仁一, 池原悟: “重文・複文文型パターン辞書からの構造照合型文型パターン検索.” 情報処理学会研究報告, 2006-NL-176, pp.9-16, 2006.
 - 20 松浦大樹, 徳久雅人, 村上仁一, 池原悟: “心的状態の解析のための対話行為タグ付きテキスト対話コーパスの試作.” 電子情報通信学会技術研究報告, TL2006-62, pp.31-36, 2007.
 - 21 片山慶一郎, 村上仁一, 徳久雅人, 池原悟: “日英機械翻訳のための文型パターン辞書の圧縮に関する検討.” 情報処理学会研究報告, 2007-NL-178, pp.41-46, 2007.
 - 22 徳久雅人, 前田浩佑, 村上仁一, 池原悟: “心的状態を表す対話行為タグ付きテキスト対話コーパスの構築.” 電子情報通信学会技術研究報告, TL2007-45, pp.25-30, 2007.

- 23 徳久雅人, 岡田直之: “移動における認識と行動の協調処理.” 平成5年度電気関係学会九州支部連合大会講演論文集, p.672, 1993.
- 24 徳久雅人, 吉田太輔, 岡田直之: “機械処理のための基本的情緒の分析.” 平成6年度電気関係学会九州支部連合大会講演論文集, p.943, 1994.
- 25 吉田太輔, 徳久雅人, 岡田直之: “情緒を表す語の概念分析.” 電子情報通信学会秋季大会講演論文集, パネル討論: 思考過程の解明へのアプローチ, PA-3-6, pp.341-342, 1994.
- 26 徳久雅人, 岡田直之: “イソップワールド”における実験支援システムの構築.” 平成7年度電気関係学会九州支部連合大会講演論文集, p.795, 1995.
- 27 徳久雅人, 岡田直之: “知的エージェントにおける情緒モジュールの作成 - - - “喜び/悲しみ”の場合.” 平成8年度電気関係学会九州支部連合大会講演論文集, p.707, 1996.
- 28 桑原博文, 徳久雅人, 岡田直之: “対話における理解から応答へのプランニング.” 平成9年度電気関係学会九州支部連合大会講演論文集, p.593, 1997.
- 29 山野直樹, 成瀬隆之, 徳久雅人, 岡田直之: “行動と対話処理のための統合的プランナ.” 平成9年度電気関係学会九州支部連合大会講演論文集, p.595, 1997.
- 30 秀島悟, 徳久雅人, 岡田直之: “個性を持つ知能エージェントの構築.” 平成9年度電気関係学会九州支部連合大会講演論文集, p.596, 1997.
- 31 木村宏, 徳久雅人, 岡田直之: “対話理解のためのLFGパーザの改良.” 平成9年度電気関係学会九州支部連合大会講演論文集, p.594, 1997.
- 32 桑原博文, 徳久雅人, 岡田直之: “並列分散方式による知能エージェントの行動プランニング.” 平成10年度電気関係学会九州支部連合大会講演論文集, p.107, 1998.
- 33 徳久雅人, 岡田直之: “基本情緒における強度の要因分析.” 平成10年度電気関係学会九州支部連合大会講演論文集, p.109, 1998.
- 34 木村宏, 徳久雅人, 岡田直之: “並列分散方式による知能エージェントの対話理解.” 平成10年度電気関係学会九州支部連合大会講演論文集, p.108, 1998.
- 35 森口拓雄, 徳久雅人, 岡田直之: “移動・獲得・飲食で構成されるプランの模擬と評価.” 平成11年度電気関係学会九州支部連合大会講演論文集, p.533, 1999.
- 36 田川峰和, 徳久雅人, 乾健太郎, 岡田直之: “基本情緒の反応規則の作成 - - - 「驚き」・「期待」・「恐れ」の場合.” 平成11年度電気関係学会九州支部連合大会講演論文集, p.534, 1999.
- 37 森口拓雄, 徳久雅人, 岡田直之: “エピソードデータを用いた対話生成.” 平成12年度電気関係学会九州支部連合大会講演論文集, p.813, 2000.
- 38 中野育恵, 徳久雅人, 徳久良子, 岡田直之: “情感のある対話処理 - 非情緒状態の認識 - .” 平成12年度電気関係学会九州支部連合大会講演論文集, p.810, 2000.
- 39 田川峰和, 徳久雅人, 徳久良子, 岡田直之: “情感のある対話処理 - 情緒状態の認識 - .” 平成12年度電気関係学会九州支部連合大会講演論文集, p.811, 2000.
- 40 日巻正寛, 徳久雅人, 徳久良子, 岡田直之: “情感のある対話処理 - 発話の選択 - .” 平成12年度電気関係学会九州支部連合大会講演論文集, p.812, 2000.
- 41 徳久雅人, 中野育恵, 岡田直之: “情緒を加味した対話プランニング.” 電子情報通信学会総合大会講演論文集, シンポジウム: 心を持つロボット, 対話するロボット, SA-7-7, pp.512-513, 2001.

- 42 徳久良子, 乾健太郎, 徳久雅人, 岡田直之: “規模とコストを考慮した感情タグつき言語コーパスの作成方法.” 電子情報通信学会総合大会講演論文集, シンポジウム:心を持つロボット, 対話するロボット, SA-7-8, pp.514-515, 2001.
- 43 岡崎良久, 徳久雅人, 岡田直之: “割り込みの復帰を考慮した自律分散型プランニング.” 平成 13 年度電気関係学会九州支部連合大会講演論文集, pp.643, 2001.
- 44 古賀美雪, 徳久雅人, 岡田直之: “情緒を加味した勧誘における語尾選択.” 平成 13 年度電気関係学会九州支部連合大会講演論文集, pp.645, 2001.
- 45 日巻正寛, 徳久雅人, 岡田直之: “自己観察を行う知能エージェントにおける心的過程の要約.” 平成 13 年度電気関係学会九州支部連合大会講演論文集, pp.646, 2001.
- 46 中野育恵, 徳久雅人, 岡田直之: “勧誘の対話における情緒推定規則の検証.” 平成 13 年度電気関係学会九州支部連合大会講演論文集, pp.644, 2001.
- 47 尾高智大, 村上仁一, 徳久雅人, 池原悟: “構造的類似文の自動検出と縮退方法.” 言語処理学会第 9 回年次大会発表論文集, pp.1-4, 2003.
- 48 杉坂岳志, 吉村英展, 徳久雅人, 村上仁一, 池原悟: “情緒推定のための対話文の解析.” 言語処理学会第 9 回年次大会発表論文集, pp.43-46, 2003.
- 49 展瑜, 徳久雅人, 池原悟, 村上仁一: “日中機械翻訳における名詞訳語の選択.” 言語処理学会第 9 回年次大会発表論文集, pp.334-337, 2003.
- 50 佐伯美香, 徳久雅人, 村上仁一, 池原悟: “副詞および形容詞による感情表現性の判定.” FIT2003, 情報科学技術フォーラム, 一般講演論文集, 第 2 分冊, pp.117-118, 2003.
- 51 田中康仁, 村上仁一, 徳久雅人, 池原悟: “構造的類似文検索アルゴリズムを応用した日本語文型パターン抽出法.” FIT2003, 情報科学技術フォーラム, 一般講演論文集, 第 2 分冊, pp.89-90, 2003.
- 52 村上仁一, 佐伯亮, 徳久雅人, 池原悟: “結合価パターンを用いた音声認識.” FIT2003, 情報科学技術フォーラム, 一般講演論文集, 第 2 分冊, pp.255-256, 2003.
- 53 岡田敏, 村上仁一, 徳久雅人, 池原悟: “DP を使用した重文複文の日英翻訳の精度.” 言語処理学会第 10 回年次大会発表論文集, pp.233-236, 2004.
- 54 前田春奈, 村上仁一, 徳久雅人, 池原悟: “パターンを使用した重文複文の日英翻訳の精度.” 言語処理学会第 10 回年次大会発表論文集, pp.237-240, 2004.
- 55 道祖尾太祐, 村上仁一, 徳久雅人, 池原悟: “N-gram を利用した日英翻訳パターンの自動抽出.” 言語処理学会第 10 回年次大会発表論文集, pp.241-244, 2004.
- 56 藤本敬史, 村上仁一, 徳久雅人, 池原悟: “埋め込み文に対する日英機械翻訳規則の作成について.” 言語処理学会第 10 回年次大会発表論文集, pp.245-248, 2004.
- 57 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.
- 58 徳久雅人, 村上仁一, 池原悟: “文型パターンパーサの試作.” 言語処理学会第 10 回年次大会発表論文集, pp.608-611, 2004.
- 59 吉田真司, 徳久雅人, 村上仁一, 池原悟: “結合価パターンを用いた仮名漢字変換候補の選択.” 言語処理学会第 10 回年次大会発表論文集, pp.717-720, 2004.
- 60 古塩貴行, 徳久雅人, 村上仁一, 池原悟: “情緒注釈付き対話コーパスの誤り分析.” 人工知能学会全国大会, 2G3-02, 2004.

- 61 岡田敏, 村上仁一, 徳久雅人, 池原悟: “多変量解析による最適文型パターンを選択方式.” 言語処理学会第 11 回年次大会発表論文集, A1-7, pp.25-28, 2005.
- 62 金澤佑哉, 徳久雅人, 村上仁一, 池原悟: “文型パターンにおける時制・相・様相表現の汎化とその効果.” 言語処理学会第 11 回年次大会発表論文集, A1-8, pp.29-32, 2005.
- 63 佐伯美香, 徳久雅人, 村上仁一, 池原悟: “感情表現に着目した日本語副詞の日英対照分析.” 言語処理学会第 11 回年次大会発表論文集, A1-9, pp.33-36, 2005.
- 64 石上真理子, 徳久雅人, 村上仁一, 池原悟: “結合価パターンを用いた動詞句の翻訳可能性の調査.” 言語処理学会第 11 回年次大会発表論文集, A2-6, pp.364-367, 2005.
- 65 遠藤久美子, 徳久雅人, 村上仁一, 池原悟: “文型パターンにおける任意要素の記述方法とその効果.” 言語処理学会第 11 回年次大会発表論文集, A2-7, pp.368-371, 2005.
- 66 西山七絵, 村上仁一, 徳久雅人, 池原悟: “単文文型パターン辞書の構築.” 言語処理学会第 11 回年次大会発表論文集, A2-8, pp.372-375, 2005.
- 67 神野絵理, 徳久雅人, 村上仁一, 池原悟: “文型パターンによる日英翻訳のための名詞句パターン辞書の構築.” 言語処理学会第 11 回年次大会発表論文集, A2-9, pp.376-379, 2005.
- 68 徳久雅人, 村上仁一, 池原悟: “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.
- 69 吉岡篤志, 徳久雅人, 村上仁一, 池原悟: “名詞句パターン辞書を用いた日英機械翻訳の試作 - - - bi-gram による訳出選択の場合.” 平成 17 年度電気・情報関連学会中国支部連合大会論文集, pp.305-306, 2005.
- 70 安井敏, 徳久雅人, 村上仁一, 池原悟: “「AのB」型名詞句に対する連体修飾節の係り先の決定.” 言語処理学会第 12 回年次大会発表論文集, pp.125-128, 2006.
- 71 小林和晃, 村上仁一, 徳久雅人, 池原悟: “単文文型パターンからの言い換えの抽出.” 言語処理学会第 12 回年次大会発表論文集, pp.200-203, 2006.
- 72 楊鵬, 村上仁一, 徳久雅人, 池原悟: “結合価パターンを用いた日中機械翻訳方式の検討.” 言語処理学会第 12 回年次大会発表論文集, pp.264-267, 2006.
- 73 原真一朗, 村上仁一, 徳久雅人, 池原悟: “日英機械翻訳における多変量解析を用いた最適パターンを選択.” 言語処理学会第 12 回年次大会発表論文集, pp.268-271, 2006.
- 74 片山慶一郎, 村上仁一, 徳久雅人, 池原悟: “日本語文型パターンの縮退方法.” 言語処理学会第 12 回年次大会発表論文集, pp.568-571, 2006.
- 75 小林和晃, 村上仁一, 徳久雅人, 池原悟: “選択記号による文型パターン汎化の効果.” 言語処理学会第 12 回年次大会発表論文集, pp.572-575, 2006 .
- 76 前田春奈, 村上仁一, 徳久雅人, 池原悟: “用言意味属性を用いた適合文型パターンの絞り込み.” 言語処理学会第 12 回年次大会発表論文集, pp.576-579, 2006
- 77 吉岡篤志, 徳久雅人, 村上仁一, 池原悟: “日英対訳パターンを用いた名詞句翻訳.” 言語処理学会第 12 回年次大会発表論文集, pp.580-583, 2006.
- 78 中村聡, 村上仁一, 徳久雅人, 池原悟: “「意味」による文型パターン検索方式の最適化.” 言語処理学会第 12 回年次大会発表論文集, pp.632-635, 2006.
- 79 田中努, 徳久雅人, 村上仁一, 池原悟: “情緒生起情報付き結合価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, pp.1151-1154, 2006.
- 80 古塩貴行, 徳久雅人, 村上仁一, 池原悟: “テキストからの情緒推定のための事象間関係の

- 解析に向けて。”言語処理学会第12回年次大会発表論文集, pp.1155-1158, 2006.
- 81 徳久雅人, 村上仁一, 池原悟: “文頭・文末表現を区別した確率言語モデルに基づく情緒推定。”言語処理学会第12回年次大会ワークショップ「感情・評価・態度と言語」論文集, pp.45-48, 2006.
- 82 徳久雅人, 松浦大樹, 村上仁一, 池原悟: “漫画における表情に着目した情緒注釈付き対話コーパスの作成。”電子情報通信学会2006年総合大会講演論文集, シンポジウム「円滑なコミュニケーション」, AS-4-6, pp.S-44-45, 2006.
- 83 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.
- 84 黒住亜紀子, 村上雄弥, 徳久雅人, 村上仁一, 池原悟: “結合価パターン辞書における情緒表現性のある用言の意味分析。”電子情報通信学会ソサイエティ大会講演論文集, 基礎・境界, p.168, 2006.
- 85 黒住亜紀子, 徳久雅人, 村上仁一, 池原悟: “結合価パターン辞書からの情緒を明示する用言の知識ベース化。”言語処理学会第13回年次大会発表論文集, pp.39-42, 2007.
- 86 金澤佑哉, 徳久雅人, 村上仁一, 池原悟: “記号と関数による文型パターン汎化効果の総合評価。”言語処理学会第13回年次大会発表論文集, pp.51-54, 2007.
- 87 石上真理子, 徳久雅人, 村上仁一, 池原悟: “関数・記号付き文型パターンを用いた機械翻訳の試作と評価。”言語処理学会第13回年次大会発表論文集, pp.67-70, 2007.
- 88 楊鵬, 村上仁一, 徳久雅人, 池原悟: “日中機械翻訳に対する結合価パターン翻訳方式の応用。”言語処理学会第13回年次大会発表論文集, pp.79-82, 2007.
- 89 徳久雅人, 村上仁一, 池原悟: “日記文における情緒生起の原因に着目した情緒推定の実験。”電子情報通信学会総合大会講演論文集, 基礎・境界, シンポジウム「円滑なコミュニケーションと感情」, AS-7-3, pp.S-79-80, 2007.
- 90 田村元秀, 村上仁一, 徳久雅人, 池原悟: “Webを知識源とする質問応答システムにおけるパターン方式とキーワード方式の比較。”電子情報通信学会ソサイエティ大会講演論文集, 基礎・境界, A-13-4, p.195, 2007.
- 91 徳久雅人, 池原悟, 村上仁一: “非線形言語モデルに基づく文型パターン型言語知識ベースの開発。”電子情報通信学会ソサイエティ大会講演論文集, 基礎・境界, シンポジウム「言語の意味と思考過程」, AS-5-1, pp.S-49-50, 2007.
- 92 黒住亜紀子, 徳久雅人, 村上仁一, 池原悟: “情緒状態と直結した情緒反応を表す用言の結合価パターン辞書の構築。”電子情報通信学会ソサイエティ大会講演論文集, 基礎・境界, シンポジウム「言語の意味と思考過程」, AS-5-2, pp.S-51-52, 2007.
- 93 前田浩佑, 徳久雅人, 村上仁一, 池原悟: “情緒を表す文末表現の書き換えの試行。”電子情報通信学会ソサイエティ大会講演論文集, 基礎・境界, シンポジウム「言語の意味と思考過程」, AS-5-3, pp.S-53-54, 2007.
- 94 水田理夫, 徳久雅人, 村上仁一, 池原悟: “重文・複文文型パターン辞書による意識の可能性。”電子情報通信学会ソサイエティ大会講演論文集, 基礎・境界, シンポジウム「言語の意味と思考過程」, AS-5-4, pp.S-55-56, 2007.

研究資料

- 1 村上仁一, 徳久雅人, 池原悟: “日本語英語の文対応の対訳データベースの作成.” 第7回 LACE 研究会報告, 2002 .
- 2 徳久雅人, 岡田直之: “タスク指向対話のための情緒推定 - - - 勧誘の場合.” 第7回 LACE 研究会報告, 2002.
- 3 徳久雅人, 池原悟, 村上仁一: “文型パターン化の方法と文型照合方式.” 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, 平成 14 年度研究状況報告会, 2003.
- 4 徳久雅人, 村本奈央, 池原悟, 村上仁一: “日英文型パターンにおける節の変数化条件について.” 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, 平成 14 年度研究状況報告会, 2003.
- 5 安井敏, 徳久雅人: “比喩の自動判定に向けて - - - 既存の比喩文との意味ベクトル比較による場合.” 第8回 LACE 研究会報告, 2003.
- 6 佐伯美香, 徳久雅人: “副詞と形容詞に着目した日本文の感情表現性の調査.” 第8回 LACE 研究会報告, 2003.
- 7 徳久雅人: “重文・複文の基本文型に対する文型パターン辞書のカバー率.” 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, 平成 15 年度研究状況報告会, 2004.
- 8 徳久雅人, 田中努, 村上仁一, 池原悟: “情緒の生起する原因を表す日本語の用言の分布について.” 「言語・認識・表現」第11回年次研究会, 2006.

付録1 文型パターン記述子

分類	名称	記述形式	役割・適合条件	記述例
字面	出現形字面	〈全角文字列〉	文型パターンにおいて非線形要素の字面を表す。照合対象の字面が出現形字面と一致するとき適合する。	N1 は AJ2。
	基本形字面	'〈全角文字列〉'	文型パターンにおいて、活用形に自由度のある字面を表す。照合対象の字面の基本形が基本形字面と一致するとき適合する。	'走る' [. kako] 'モナカ' を V1 ※「走った」や「もなかを食べる」と適合
変数	意味自由変数	〈変数名〉〈変数番号〉	文型パターンにおいて線形要素を表す。通常、変数名は品詞を表す。照合対象の部分表現が変数の表す条件と一致するとき、適合する。適合した部分表現は変数に代入される。	N1 が走る。 ※「犬が走る」と適合し、「N1=犬」が返される。
	意味制約付き変数	〈変数名〉〈変数番号〉〈意味属性制約〉	文型パターンにおいて、意味属性で限定された線形要素を表す。自由変数の適合条件に加え、意味属性制約を満たすとき適合する。	NP1 (IM: 123, IM: 456)
関数	ドット関数	. 〈関数名〉	通常、関数名は助動詞を表す。照合対象の部分表現が関数の表す条件と一致するとき適合する。	V1. kako ※「走った」と適合
	ハット関数	^ 〈関数名〉	直前の記述子に対する制約を表す。直前の記述子の被制約箇所が関数の表す条件と一致するとき適合する。	V1^renyo ます。 ※「走ります。」と適合 V1 が「走り」と適合し、それが連用形であるので、^renyo も適合
記号	文節境界記号	!	文型パターンにおいて、文節境界を表す。照合対象の当該部分表現に文節境界が存在するとき適合する。	N1 は ! N2 だ。 ※「これは本だ。」と適合
	非文節境界記号	+	文型パターンにおいて、非文節境界を表す。照合対象の当該部分表現に文節境界が存在しないとき適合する。	N1 + は N2 + だ。
	離散記号	/ 〈離散名〉	文型パターンにおいて、新たに挿入可能な表現を表す。照合対象の部分表現が離散記号の表す条件と一致するとき適合する。	/t N1 は /c 貸した。 ※「昨日買った本は友達に貸した。」と適合 /t は連体修飾節の部分表現と適合し、/c は格要素と適合

(続き)

分類	名称	記述形式	役割・適合条件	記述例
記号	選択記号	選択記号の前後いずれかのパターン記述が適合することを適合とする.	N1 (は が) V2。 ※ 「太郎は行く。」も「太郎が行く。」も適合
	任意記号	[...]	囲まれたパターン記述は、適合しても、しなくてもよいことを表す.	N1 は [すぐに] V2。 ※ 「太郎はすぐに行く」も「太郎は行く」も適合
	補完記号	< ... >	囲まれたパターン記述は、適合しても、しなくてもよいことを表す。パターン作成に用いた標本に存在しない要素を、パターンに追加する際に使われる.	<N1 は>嬉しい。 ※ 「私は嬉しい。」も「嬉しい。」も適合
	順序任意記号	{ ..., ..., ... }	中括弧内は、コンマで区切られて、パターン記述が並べられている。照合対象の部分表現がこれらの記述に順不同で全て適合することをこの記述子全体の適合とする.	{N1 が, N2 に, N3 を} V4。 ※ 「太郎が花子に本を貸す。」や「花子に本を太郎が貸す。」などが適合
	移動可能記号	\$〈番号〉 \$〈番号〉^ { ... }	文型パターンにおいて、位置変更の可能なパターン記述とその位置を表す。照合対象の部分表現がいずれかの位置でそのパターン記述と適合するとき適合とする.	\$1^ {N1 は} N2 を \$1V3 て、\$1V4. kako。 ※ 「太郎はそれを取って、喜んだ。」や「それを取って、太郎は喜んだ。」が適合。前者は、\$1=0 が返され、後者は\$1=2 が返される.
	まとめ記号	(...)	囲まれたパターン記述を照合処理上、1つの単位にまとめる。選択記号や記憶記号の及ぶ範囲を定める.	N1 (は が) V2 ※ 「太郎は行く」も「太郎が行く」も適合
	記憶記号	#〈番号〉	直後のパターン記述子に適合した部分表現を記憶する.	#1 (N1 は)、#2V3.kako ※ 「太郎は帰った」と適合し、#1=[N1, は], #2=[V3] が返される.

付録 2 情緒生起の特徴

コロン“:”の右側はラベル名, 左側は定義文である. ○付きの特徴は, 最下位特徴である. /は対極であることを表す.

喜び/悲しみ: 現状態は前状態よりも好都合/不都合である

生理的: 生理状態の現状態は, 前状態よりも好都合/不都合である

○内的な快/不快: 内的な生理状態が良く/悪くなった

○外的な快/不快: 外的な生理状態が良く/悪くなった

心理的: 心理状態の現状態は, 前状態よりも好都合/不都合である

目標実現: 目標実現に好都合/不都合な状態になった

情報収集: 目標実現に有効な情報を得た/得なかった

○思惑通り/思惑違い: 目標実現に有効な物事が予測した通り認識できた/できなかった

○発見/見落とし: 目標実現に有効で自分の知らなかった物事を認識した

/目標実現に有効な物事が行動途中で存在していたことを後で知った

○判明/不明: 目標実現に必要で有効な物事が, 努力の末分かった/努力の末分からない

計画: 目標実現に有効な計画を得た/得なかった

○立案/無計画: 目標実現に有効な計画が生成できた/できなかった

実行結果: 目標を実現する/しない結果である

○完遂/断念: 努力して計画を実行し, 目標を実現した/目標を実現できずに諦めた

○獲得/喪失: 目標実現に必要な物事を, 努力して手に入れた/なくした

○有効/無駄: 目標実現への行動結果は, 予測以上だった/効果がなかった

対人関係: 対人関係が好都合/不都合な状態になった

仲間意識: 仲間意識の現状態は前状態よりも好都合/不都合になった

○同意/反対: 他者は自分の意見を, 受け入れた/受け入れなかった

○同感/反感: 他者はある事に対して自分に賛同的/反抗的な感情を持った

○協力/非協力: 他者は自分の目標実現のための行動を支援した/しなかった

○仲直り/仲たがひ: 不仲になった者と再び仲良くなった

/仲良しだった者と不仲になった

優劣関係: 仲間内での優劣関係が好都合/不都合になった

または, 好都合/不都合な関係が発揮されている

○優越/劣等: 自分の能力を周囲の者より高く/低く自己評価した

○賞賛/非難: 自分の能力は他者に高く/低く評価された

○服従/反抗: 他者は要求の通りに行動した/しなかった

○厚遇/冷遇: 他者は自分の状態を配慮した/無視した行動をとった

○保護/見放し: 他者は自分を危険から守った

/他者は自分の危険を知りつつ何もしなかった

その他

好ましい／嫌だ：自分にとって有利／不利な状況になった

生理的：生理的に有利／不利な状況になった

- 内的な快／不快：内的な生理状態が改善した
- 外的な快／不快：外的な生理状態が改善した
- 五感：（視覚）明るい色彩のものだった／物を認識できないほどの明るさだった
（形や大きさ動き等が）整っていた／いなかった
（聴覚）音色や音量、リズムの統制がとれていた／とれていなかった
安定した自然環境を表す／危害を加えるものの存在を知らせる音だった
（嗅覚）綺麗／不潔，食料／敵などの存在を知らせる匂いだった
（味覚）食べることができる／できない味だった
（触覚）痛みや温度，固さ，手触り，衝撃は体の各部位に馴染む／限度を超えている

心理的：心理的に有利／不利な状況になった

目標実現：目標実現が有利／不利な状況になった

目標設定：目標の選択が有利／不利になった

- 自由／制限：条件や制約が減少／増加し，実行可能なプランの数が増加／減少した

情報収集：目標実現に有利／不利な情報だった

- 判明／不明：目標実現に必要で不明な物事が，努力の末分かった／分からなかった
- 教訓／迷信：目標実現の役に立つ／立たない情報だった
- －／嘘：目標実現のための誤った情報だった

実行結果：目標実現に有利／不利な結果だった

- 獲得／喪失：目標実現に必要な物事を手に入れた／なくした
- －／無駄：目標実現への行動結果は，得られるはずのものが得られていなかった
- 補助／妨害：目標実現の助け／妨げになった

対人関係：仲間集団の状況は有利／不利になった

仲間意識：仲間意識が高まった／下がった

- 信頼／不信：他者は信頼できる／できない行為をした
- －／露見：知られたいくないことが他者に知られた
- 同意／反対：他者は自分の意見を受け入れた／受け入れなかった
- 協力／非協力：他者は自分の目標実現のための行動を支援した
- －／一方的：他者の目標実現だけが進み，自分の目標実現は進まなかった
- 出会い／別れ：他者の数が増加／減少した
- －／仲間外れ：自分を仲間と認めてくれなかった

優劣関係：仲間内での優劣関係が有利／不利になった

- 優越／劣等：自分の能力を周囲の者より高く／低く自己評価した
- 賞賛／非難：自分の能力は他者に高く／低く評価された
- 厚遇／冷遇：他者は自分の状態を配慮／無視した行動をとった
- －／無礼：自分の下位の者の行動は，礼儀に外れていた
- 保護／見放し：他者は自分を危険から守った／他者は自分の危険を知りつつ何もしなかった
- －／説教：上位の者から制約を受けた

その他

驚き：予測していなかったことが起きた

生理的：予測していなかった生理的情報を得た

○五感急変：予測していなかった五感の急変が起きた

心理的：予測していなかった心理的情報を得た

情報収集：予測していなかった情報を収集した

○発見：予測していなかった物事を認識した

：予測していたことが起きなかった

心理的：予測していたことが起きなかった

目標実現：目標実現の過程が予測通りではなかった

情報収集：目標実現に必要な情報を予測した通りに収集できなかった

○思惑違い：目標実現のために予測していた物事を認識できなかった

実行結果：目標実現のための行動の結果が予測通りではなかった

○成功：可能性が低い行動にもかかわらず成功した

○失敗：可能性が高い行動にもかかわらず失敗した

その他

期待：好都合なことが起こると予測した

生理的：生理状態が好都合になると予測した

○内的な治癒：内的な生理状態が良好になると予測した

○外的な治癒：外的な生理状態が良好になると予測した

心理的：心理状態が好都合になると予測した

目標実現：目標が実現できると予測した

情報収集：収集した情報により、目標が実現できると予測した

○成行き：認識している事象の成り行きが、目標実現に有利であると予測した

○終了直前：もう少しで目標が実現できると予測した

計画：計画が立つことで、目標が実現できると予測した

○成算：成功する可能性の高い計画をたてた

対人関係：対人関係が好都合になると予測した

仲間意識：仲間関係が好都合になると予測した

○同意：自分の意見に他者は同調すると予測した

○同感：自分に賛同的な感情を他者は持つと予測した

○協力：目標実現に他者が協力してくれると予測した

○仲直り：自分と他者との関係が良くなると予測した

優劣関係：優劣関係が好都合になると予測した

○賞賛：自分のことを他者が高く評価すると予測した

○優越：他者よりも優れた結果を自分は出せると予測した

○服従：自分の要求に他者が従うと予測した

○保護：危害を及ぼすものから、他者が自分を守ると予測した

○厚遇：他者が自分を丁重にもてなすと予測した

その他

恐れ：危害を及ぼす事象・対象物を認識した

生理的：生理状態が悪化したり、直接身体に危害の及ぶ事象・対象物を認識した

- 内部異常：身体内部の状態が悪化すると予測した
- 外部損傷：身体表面の状態が悪化すると予測した
- 五感：五感で感じた刺激は身体に悪影響を及ぼすと予測した

心理的：心理状態が不都合になる(不都合なまま、あるいは、さらに悪化する)と予測した

目標実現：目標が実現できないと予測した

情報収集：目標実現の失敗につながることを認識した

- 注意事項：目標実現を困難にさせると既に分かっている物事を認識した
- 正体不明：目標実現を困難にさせないとは断定できない物事を認識した
- 存在不明：目標実現を困難にさせる物事が存在するかもしれない状況を認識した
- 危機一髪：目標実現の失敗につながる状況を認識した

計画：目標実現に成功する計画が生成できなかった

- 試行：成功の可能性の低い計画、初めて実行する計画を生成した
- 絶対絶命：目標実現の失敗を回避したり補ったりする計画が生成できない

実行結果：目標実現の失敗につながる実行結果だった

- 喪失：目標実現に必要な物をなくした

対人関係：対人関係が不都合になると予測した

仲間意識：仲間関係が不都合になると予測した

- 暴露：恥ずべきことや悪事などの秘密が他者に知られると予測した
- 叱責：自分の計画や行動の不備がとがめられると予想した

優劣関係：優劣関係が不都合になると予測した

- 非難：自分のことを他者が低く評価すると予測した

その他

怒り：規則や当然のことが守られていない

生理的：生理状態についての規則や当然のことが守られていない

○内的な不快：内的な生理状態が不快であってはいけないのに、現状は不快である

○外的な不快：外的な生理状態が不快であってはいけないのに、現状は不快である

心理的：心理状態についての規則や当然のことが守られていない

目標実現：目標を実現するための規則や当然のことが守られていない

目標設定：目標を設定する際の規則や当然のことが守られていない

○強制：義務でも無く、欲求を満たさない目標を実現しなければならない

情報収集：目標実現のための情報を集める際の規則や当然のことが守られていない

○嘘：目標実現に必要な情報が、真実でなければならないのに誤りである

実行結果：目標実現のための実行の際の規則や当然のことが守られていない

○失敗：目標実現のための行動で、成功しなければならないのに失敗した

○妨害：目標実現のために、自分は合法的に行動しているのに、他者はその行動を困難にした

○ずる：目標実現のために、自分は合法的に行動しているのに、他者は同じ目標のために違法の行動をした

対人関係：対人関係に関する規則や当然のことが守られていない

仲間意識：仲間としての規則や当然のことが守られていない

○反対：受け入れなければならない自分の意見を、他者は受け入れない

○反感：自分に賛同的な感情を持たなければならない物事に対して、他者はその感情を持たない

○仲間外れ：他者の圧力により、仲間集団として当然の行動や待遇を自分には行うことができない

○脱退：他者は、仲間の一員であるのに、仲間集団から抜けようとする

○暴露：他者は、自分の秘密事項を、知ろうとしたり、他者に好評する

○責任逃れ：仲間集団内での果たすべき事を他者は行わない

○責任転嫁：仲間集団内での果たすべき事を他者は自分にさせる

優劣関係：優劣関係に関する規則や当然のことが守られていない

○非難：周囲の者が自分の能力を不当に低く評価する

○反抗：自分の要求に対して従うべき立場の他者が従わない

○無礼：自分の下位の者の言動は礼儀に外れている

○不平等：自分の評価や報酬が、周囲と平等でない

その他