

992031 文型パターンによる情緒推定のための 対話文解析

計算機工学講座 池原研究室 杉坂 岳志

概要

近年，知能エージェントやロボットが人間と自然に対話することを目指して研究が進められている．タスク指向対話において対話相手の情緒を推定することは，対話システムと相手との親和性を向上させ，対話を効果的に行うために重要である．そのために，発話文に明示されない相手の情緒を推定する必要がある．情緒推定を行う手段として，発話理解により対話文の心的状態を抽出し，情緒生起の原因が記述された対話コーパスより生成された情緒推定のためのルールベースと照合する方法がある [9]．

本研究では，日本語文の主要な情報を伝える命題表現と，命題表現の内容に対する感じ方や態度を指す様相表現（モダリティ表現）からなる重層構造に着目し，命題情報は結合価文法により解析し，様相表現は日本語文型辞典 [4] を使用して詳細な意味・機能を調査し，それらを統一的に解析することで発話行為の種類，心的状態の抽出規則となる対話文文型パターンを試作する．文型パターン作成のため，児童向けの寓話における勧誘の対話を対象に 36 対話 528 文を収集し，479 種類の文型パターンを作成した．動作確認として，収集対話文 528 文を形態素解析にかけ，解析に失敗したものは除外し，対話文文型パターンと照合を行う．結果として 98.1 % の正解率を得た．よって対話文文型パターンの有効性を示すことが出来た．

目次

1	はじめに	1
2	対話文からの情緒推定	2
2.1	情緒の定義	2
2.2	テキスト処理に関する情緒推定	2
2.3	情緒推定要因に基づく推定	3
2.3.1	コーパス	3
2.3.2	発話行為	4
3	対話文の解析	5
3.1	対話文の構造	5
3.2	命題表現	5
3.2.1	一般名詞意味属性	5
3.2.2	用言意味属性	6
3.2.3	結合価文法	6
3.3	様相表現	7
3.4	対話文文型パターン	7
4	対話文文型パターンの作成	8
4.1	対話事例の収集と整理	8
4.2	対話文文型パターンの作成	9
4.3	作成結果	10
5	対話文文型パターンの動作確認	11
6	考察	12
6.1	前処理失敗文の考察	12
6.2	動作確認失敗の考察	12
7	おわりに	13

目 次

1	『喜び/悲しみ』の特徴フレーム	2
2	情緒注釈付きコーパスの一例 (簡略表記)	3
3	一般名詞意味属性体系 (上位属性)	5
4	用言意味属性体系	6
5	対話事例の一部	8

表 目 次

1	結合価パターンの例	6
2	意味機能の例	7
3	発話行為の表層的特徴を表すキーワード語の例	9
4	対話文文型パターンの一部	10
5	対話文文型パターンによる発話行為の認識率数	11

1 はじめに

親和性のある対話処理のために、言語理解により対話相手の情緒を推定する必要がある [1],[2]。テキスト処理を対象にした情緒・感情の推定には様々な方法がある。

中でも、対話処理においては、情緒・感情の原因を深く理解して推定することも必要であることから、対話相手の心的状態に基づき情緒を推定する方法が開発されている [3]。しかし、対話文から心的状態を抽出する方法が問題として残されている。

対話文の解析には、発話行為とその内容を認識する必要がある。日本語文の場合、発話行為に関連する「依頼」「許可」などの意味・機能は文型に表れやすく、[4] はそれらを文型辞典として収録している。

本研究では、寓話における勧誘の対話を対象に、[4] を参考にして対話文を解析するための文型パターンを作成する。動作確認の結果、98.1% の文型パターンが正しく発話行為を認識し、対話文解析への有効性が確認できた。

本論文の構成は以下の通りである。まず、第1章で研究の背景として、対話処理のための心的状態に基づく情緒の推定方法について紹介し、本研究の目的である対話文文型パターンについて述べる。そして、第2章で情緒推定過程を具体的に述べ、情緒推定のために必要な発話行為の定義をする。第3章では、発話行為を導き出すための対話文文型パターンの構成について説明する。次に第4章で対話文文型パターンの作成過程と結果について、第5章で文型パターンを使ったクローズドテストの実施とその結果について示す。第6章ではテストの結果について考察をし、最後に第7章で今後の課題を提案する。

2 対話文からの情緒推定

2.1 情緒の定義

[3]では、Plutchikの多因子分析論を基に「喜び/悲しみ、好ましい/嫌だ、驚き、期待、怒り、恐れ」の8つの基本情緒について分析し階層的にまとめられている。例えば「喜び/悲しみ」は、最上位の特徴に「現状態は前状態よりも好都合/不都合である」があり、幾つかの中位特徴があり、最下位では「対人関係の優劣関係において優位である/劣っている」という特徴などがある。このように8つの基本情緒について123個の最下位特徴を定義している。

(.î /Ô"Æ/•/ß
(, %¾Ä' Ö/~/°¾Ä' Ö/Æ/â,,¥-Ô,,ç/ÉÔ-Ô,,ç/' /ç/
Ä, ~ý-~(~â-~/' +÷/ÉÔ†÷;□†°-~/' +÷/ÉÔ†÷);□
è·~ý-~((ÜÉ, ...', % (%Ö°ó...ý%, (»x~·/É/'/æ/»x~·°a//;□ ~fl, </, <<~î /•;□~%ì ÄÉÔ Ä) ;□ •x†±(. '°-î μ•x†±);□ ...' "Ô•°†ì (.°è°/^°°° ;□†~Ä^Ó..'' ;□ ~>, æ/ì μ'ì)) ;□ ^-è°·Ø•, (^ç·Ö°Ö..-(~°Ö"è'-;□~·°¶/"è·¶;□ ¶□, ~/"ó¶□, ~;□^ç~¾/æ/^ç/è/<///);□ ~¥, ô·Ø•, (~¥-Ü/ô-ø;□P»è/"óæ;□ Ép¾¾"è,,†;□·Ý, î /, <°ü/,,;□, ü¶±/. ä¶ /%/, ~¾))

図 1: 『喜び/悲しみ』の特徴フレーム

2.2 テキスト処理に関する情緒推定

テキスト処理に関する情緒推定には、様々なアプローチがあげられているが、大きく2通りに分けることができる。1つはテキストに明示されている情緒を理解するものである。例えば副詞や形容詞といった、文中の表現から得られる「快・不快」的なイメージ値から情緒を推定する方法 [5] が挙げられる。もう1つはテキストに明示されていない情緒を推定するものである。この方法には『涙を流す 悲しいとき』のように、動作や仕草といった「情緒反応」から情緒を推定する方法と、「宝くじが当たった」という1つの<認識>を伝える発話から、<獲得による喜び>という情緒を推定する、というような情緒生起の原因事態を理解する方法の2種類がある。本研究では、情緒生起の原因事態を理解する方法を使用する。

2.3 情緒推定要因に基づく推定

2.3.1 コーパス

対話相手の情緒を推定するには、以下のように定義する。

- 1) 発話文から相手の心的状態を理解する。
- 2) もし心的状態に情緒生起の要因を見出せたならば、相手に情緒が生起する。

本研究で対話文から情緒を推定する過程は、情緒注釈付き対話コーパスに表示する [6]。図 1 の「かえるくん」の発話は、予測を伝える発話である。がまくんの対話相手は、予測からプランの評価が高まるため、期待の情緒が生起すると推定できる。

初期状態：
g101.1, 目標, 手紙を所持, 設定, [], c
p101.1, プラン, 手紙を受け取る, 採用, [], c
ar01.1, 関係, 目標 プラン, [g101.1, p101.1], [], c
- - - - -
かえるくん:「今日は誰かがお手紙くれるかもしれないよ。」
sa10.1, 信念-発話行為, 予測の伝達, [], [], p
pr03.1, 信念-予測, 今日は誰かががまくんに手紙をくれる,
[], [sa10.1], p
ar16.1, 関係, プラン 予測,
[p101.1, pr03.1], [p101.1, pr03.1], c
pr04.1, 予測, 今日は誰かががまくんに手紙をくれる,
[], [pr03.1], p
ev01.1, 評価, 可能性(手紙を受け取る), 高い,
[p11, pr04.1], p
em02.1, 情緒, 期待-成算, 生起, [g101.1, p101.1, ev01.1], p
がまくん:「馬鹿らしいこと言うなよ。」
sa11.1, 発話行為, 予測の否定, [], [], c
pr05.1, 予測, 今日は誰かががまくんに手紙をくれる,
[否定], [sa10.1, pr03.1, sa11.1], p
close(pr04.1, [pr04.1, sa11.1])
close(ev01.1, [ev01.1, pr04.1])
close(em02.1, [em02.1, ev01.1])
- - - - -

図 2: 情緒注釈付きコーパスの一例(簡略表記)

2.3.2 発話行為

本研究では、勧誘というタスク指向対話を題材としているので、[7]を参考に「発話行為の種類」を定め、情緒推定に向け「発話行為の内容」を心的状態とする。発話行為は、以下のラベルを組み合わせてとらえる。

- 発話行為の種類：質問，伝達，確認，肯定，否定，要求，受諾，拒否，その他
- 発話行為の内容：生理，欲求，情緒，目標，プラン，予測，評価，記憶，認識，行動，言語

なお「挨拶」「お礼」や「独り言」などはその他とする。

3 対話文の解析

3.1 対話文の構造

日本語文の命題情報と様相情報の構造に着目する。例えば図1の「かえるくん」の「予測の伝達」を表す発話は「誰かが君に手紙をくれる」という命題表現と「～かもしれないよ」という様相表現がある。

3.2 命題表現

命題表現は結合価文法によって解析する。例の命題表現の場合、結合価パターン「N1【3 主体】が/から N2【388 場所】【533 具体物】【1001 抽象物】を N3【3 主体】に/へ くれる」に適合する [8]。このことから、命題の意味が【19 所有的移動】ということがわかる(【】内は意味属性)。

3.2.1 一般名詞意味属性

一般名詞意味属性は、約40万語の名詞を最大12段の木構造で構成する2,710の意味属性に分類し、上位の意味属性は下位の意味属性を包含する。これにより命題の名詞を解析する。

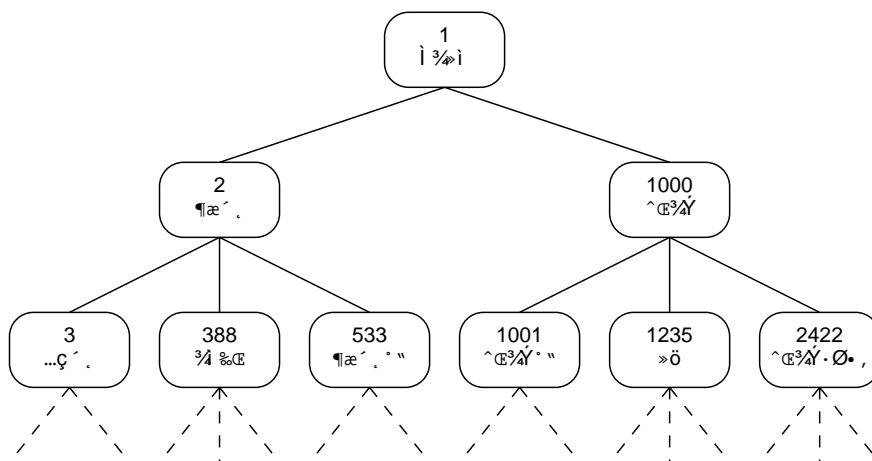


図 3: 一般名詞意味属性体系 (上位属性)

3.2.2 用言意味属性

用言意味属性は、最大4段の木構造で構成する36種類の意味属性に分類し、一般名詞意味属性と同様に上位の意味属性は下位の意味属性を包含する。例えば「歩く」という動詞には『18 物理的移動, 23 身体動作』の2つの用言意味属性が分類される。これにより命題の用言を解析する。

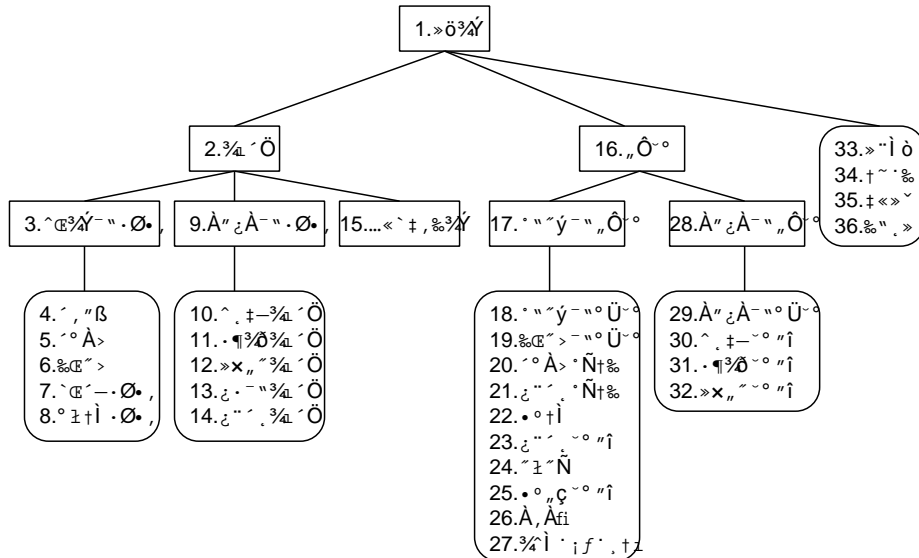


図 4: 用言意味属性体系

3.2.3 結合価文法

結合価文法とは用言と格要素(名詞+助詞)の意味的用法を記述したもので、日本語解析で発生する意味上の多義の解消を目的として開発されたものである。これにより支配される命題の意味的用法が明確になる。以下に結合価パターンの例を示す。

表 1: 結合価パターンの例

結合価パターン	用言意味属性
N1(4人 535 動物)が N2(2406 呼吸 652 涙)を止める	20 属性変化
N1(4人 800 薬品)が N2(1242 痛み)を止める	21 身体変化

3.3 様相表現

様相表現は日本語文型辞典 [4] により意味機能を調べる。この辞典では、文型を「文や節の意味，機能，用法に関わる形式」という広い枠組みで捉えており，日本語能力試験 1 級，2 級レベルで出題される文型や，新聞，雑誌，小説，シナリオ等から 3000 項目の表現を抽出している。例の「～かもしれないよ」には「推量」の意味機能を付与することができる。以下に様相表現の意味機能の例を示す。

表 2: 意味機能の例

様相表現	意味機能
～てくれ	依頼，指示，命令
～ましょう	勧誘，勧め
～だって	伝聞

3.4 対話文文型パターン

上記の命題表現と様相表現を総合的に判断することで，客観的な事態を予測したことを述べており「予測の伝達」という発話行為が認識できる。したがって，これらを別々に判断すると文全体の意味がとらえにくいと考えられる。そこで本研究では，発話行為を命題表現から得られるを格要素，用言とその意味属性，そして様相表現で構成される対話文文型パターンで解析することにする。なお，対話文文型パターンの構成において，対話文の文脈は考慮に入っていない。

4 対話文文型パターンの作成

4.1 対話事例の収集と整理

本研究では児童向けの寓話より勧誘の対話を収集する。収集文の選考基準は、話題が途中で変わらない、5ターン以上続く、台詞に単文の多い、とする。こうして、「二人は友達」、「ムーミン」などから36対話528文を収集する。

収集した対話文には、話し手、聞き手、および、発話行為についての情報を付与する。「ああ」、「えーと」などの情緒推定、および、勧誘の対話に関わらない台詞は、文型パターン作成の対象外とする。

~i -ö SP	HR	'æ:i	"fl~^„ô'ù
1	¥>¥C¥ñ¥C¥„ ~†	"É~ü/, ¥Ñi...¥~¥E/~†Ú/•/«/^/¿/«//i'	É%4t`
2	~† ¥>¥C¥ñ¥C¥„	/ロ/s;/	iβ
3	~† ¥>¥C¥ñ¥C¥„	//^/~ /â†Ú/•/«/^/¿/í ¥	É%4t` / . .
4	¥>¥C¥ñ¥C¥„ ~†	/%/i/~ .É/«/^/¿i¥	%0%i / . .
5	~† ¥>¥C¥ñ¥C¥„	¥>¥C¥ñ¥C¥„ ; /ì Á~ü/, ¥Ñi...¥~¥E/â/E/í /•/fl/~ ¥ ¥x¥	¥x¥0¥ó/ . .
6	¥>¥C¥ñ¥C¥„ ~†	/P/«/»/í /Ei"	¥x¥0¥ó/ . .
7	~† ¥>¥C¥ñ¥C¥„	¥>¥C¥ñ¥C¥„ / . ¥Ñi...¥~¥E/~ ; // / ^ / ~ / â†Ú/•// / i ¥ É%4	É%4t` / . .
8	¥>¥C¥ñ¥C¥„ ~†	.Ü-öi'	É%4t` / . .
9	~† ¥>¥C¥ñ¥C¥„	/ロ/s;/	iβ
10	~† ¥>¥C¥ñ¥C¥„	//^/~ /âi¥	É%4t` / . .

図 5: 対話事例の一部

4.2 対話文文型パターンの作成

文型パターンは字面と変数で記述する．

字面には，命題表現の格助詞，様相表現の字面 [4]，および，発話行為の表層的特徴を表すキーワード語である．キーワード語の一部を以下の表に示す．

表 3: 発話行為の表層的特徴を表すキーワード語の例

キーワード語	発話行為
～をよろしく	プランの要求
～てる頃だよ	予測の伝達
嫌	プランの拒否
ふうん	プランの確認

変数には，名詞句「*NP*」，動詞「*V*」，形容詞「*AJ*」，形容動詞「*AJV*」，話し手「*SP*」，聞き手「*HR*」がある．変数化した名詞句，用言には意味属性を制約条件として付与する．

例えば「お前どこにいるんだ？」(認識の質問)に対応する文型パターンは「*HR1* どこに *V1* んだ？」であり，*V1* の意味属性制約は「*V1*=【4 存在】」となる．

なお「こんにちは」「またね」などの定型句の文は，字面のみで対話文文型パターンとする．

4.3 作成結果

479種類 528個の対話文文型パターンを作成した．その一部を以下の表に示す．

表 4: 対話文文型パターンの一部

発話行為	文型パターン	意味属性制約
プランの要求	NP1の V1までを NP2に V2てみせてよ	NP1=860, V1=26, NP2=1048, V2=23
プランの拒否	AJV1よ	AJV1=5
行動の伝達	NP1に SP1は NP2を V1ている所なんだから	NP1=125, NP2=1114, V1=23
認識の質問	なあと NP1って?	NP1=125

日本語文型辞典との比較 日本語文型辞典では、「～た＝タ形」などの基本表現，および「～てくれ＝依頼」「～ているうちに＝期間」などの意味・機能が紹介されてる．今回対象とした対話文に含まれている件数を調べたところ，377文(71.4%)であった．

「プランの要求」を認識するためには「依頼・指示・命令」「勧誘・勧め」の意味機能の語が手がかりとなる．同様に「目標の伝達」の場合は「欲求・強い願望」が「予測の伝達」の場合は「推量」が手がかりとなる．

発話行為ごとの HR, SP の特徴 文型パターンに含まれる HR(聞き手) および SP(話し手) の件数はそれぞれ 98件および 73件であった．両方が含まれるのは 9件であった．

発話行為の内容が「生理」「情緒」「目標」という対話文文型パターンには，変数に SPが使われることはあるが HRが使われることはなかった。「(私は) 疲れた」や「(私は) 眠りたい」と述べることはあるが「(あなたは) 疲れた」ということが無いということと関係があるといえる．また，発話行為の種類が「拒否」である場合も同様であった．

5 対話文文型パターンの動作確認

形態素解析済みの対話文 528 文を離散型句節タイプ対話文文型パターンと照合させ、発話行為の認識を実験した。その結果、パターン照合の前処理段階での失敗を除くと 98.1%の正解率となった。一方、語尾パターンの場合、多くが一意に決まらなかった。このことから、命題表現を組合せる利点が伺える。

表 5: 対話文文型パターンによる発話行為の認識率数

	正解数	誤り数	前処理失敗数
対話文文型パターン	465	9	54

6 考察

6.1 前処理失敗文の考察

表3において，前処理に失敗した54文のうち，形態素解析に失敗した文がのべ25文であった．例えば「私の側に近付かないで。」(プランの要求)に対応する対話文文型パターンは「*SP1*の*NP1*に*V1*ないで」，意味属性制約は「*NP1*=2699 そば，*V1*=18 物理的移動」であるが，形態素解析結果は以下ようになる．

;; 形態素結果表示

1. 私 (1710,8,37)/の (7410)
 2. 側 (1100,2666,2710,2645,2600,956)/に (7430)
 3. 近 (3100)/付 (1100,2239,2475,2273,2157)/かない (1410, 叶い)(2193, 適う,5)(2393, 適う,5)(2193, 叶う,5,22)(2393, 叶う,5,22)(1410, 適い,1272,2463)/で (7410)
- \$ 英文 近 付 is fulfilled near me.

「近付く」を正しく形態素解析できていないため，対話文文型パターンは成立しない．

文中の名詞句，用言に意味属性が無い文がのべ31文であった．例えば「妹ってのはめんどくさいものだなあ。」(認識の伝達)に対応する対話文文型パターンは「*NP1*ってのは *AJ1*ものだなあ」であるが，日本語語彙大系に「めんどくさい」の用言意味属性が存在しないため対話文文型パターンが成立しない．

これらの失敗は，問題の部分を字面に変換することで文型パターンとすることができるとは，汎用性が失われ，1文のみにしか対応しない文型パターンとなり得る．

6.2 動作確認失敗の考察

照合を誤った文は，9文中6文が同じ単語の繰り返しを含んでいた(例：何日も何日も，何時間も何時間も前だよ)．これは単語の繰り返し処理を吸収する照合アルゴリズムを追加することで対処可能である．

残り3文は，対話文文型パターンや形態素解析に異常が見られなかった点から，クローズドテストの実行プログラムに何らかの不都合が生じたと思われる．

7 おわりに

寓話における勧誘の対話事例を対象に，日本語文の命題表現と様相表現の構造に着目した発話行為を認識する対話文文型パターンを 479 種類作成した．文型パターンの動作確認では，98.1 %の認識率となり，様相表現のみにより発話行為を認識した場合より高い効果を得た．これにより命題表現と様相表現を統一的に判断する手法が有効であると確認できた．

今後は，対話事例のさらなる収集と文型パターンの作成により，文型パターンの汎化を進める．そして用言句の変数化によるパターンの効率化，対話文の文脈を考慮に入れた文型パターンの作成，情緒推定システムへ [9] の組み込みが課題である．

謝辞

本研究を進めるに当たり，終始に渡り御指導頂きました徳久雅人助手に深謝いたします．

また，種々の御助言を頂きました池原悟教授，および村上仁一助教授に心から御礼申し上げます．そして，同じ情緒をテーマに研究を進め，研究に関し相談にのって下さった有田真康さん，情緒注釈付き対話コーパスに関する作業支援ツールの強化に当たって下さった吉村英展さん，および様々な場面で御助力頂いた計算機工学講座池原研究室の皆様深く感謝の意を表します．

参考文献

- [1] E. André, M. Klesen, P. Gebhard, S. Allen and T. Rist: “Integrating models of personality and emotions into lifelike characters”, *Affective Interactions* (Ed. by A. Paiva), Vol. 1814 of *Lecture Notes in Artificial Intelligence*, Springer, pp. 150–165 (2000).
- [2] F. de Rosis and F. Grasso: “Affective natural language generation”, *Affective Interactions* (Ed. by A. Paiva), Vol. 1814 of *Lecture Notes in Artificial Intelligence*, Springer, pp. 204–218 (2000).
- [3] 徳久, 中野, 山下, 岡田: “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”, *信学技報*, **TL2001-25**, pp. 21–28 (2001).
- [4] 砂川, 駒田, 下田, 鈴木, 筒井, 蓮沼, ベケシュ, 森本: “日本語文型辞典”, くろしお出版 (1998).
- [5] 佐伯: “副詞および形容詞による感情表現性の判定”, 鳥取大学工学部知能情報工学科卒業論文, (2003).
- [6] 吉村: “情緒注釈付き対話コーパスに関する作業支援ツールの強化”, 鳥取大学工学部知能情報工学科卒業論文, (2003).
- [7] 飯田, 有田: “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*, **31**, 6, pp. 810–821 (1990).
- [8] 池原, 宮崎, 白井, 横尾, 中岩, 小倉, 大山, 林: “日本語語彙体系”, 岩波書店 (1997).
- [9] 有田: “情緒推定のためのルールベースの強化”, 鳥取大学工学部知能情報工学科卒業論文, (2003).