

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

徳久 雅人 前田 浩佑 村上 仁一 池原 悟

鳥取大学工学部知能情報工学科 〒 680-8552 鳥取県鳥取市湖山町南 4-110

E-mail: { tokuhisa, s032049, murakami, ikehara } @ike.tottori-u.ac.jp

あらまし 本稿は、心的状態を表す対話行為タグを付与したテキスト対話コーパスを構築した結果について報告する。心的状態は、テキスト対話から対話者の情緒を推定するために必要とされる情報である。従来の対話行為タグでは、心的状態の表し方に曖昧性があるという問題があった。そこで、「文字通りに解釈できること」を念頭に言語表現に対する心的状態の分類基準を定めて、対話行為タグを定義する。約3万文のテキスト対話に対話行為タグを付与し、その信頼性を評価する。

キーワード テキスト対話, 対話行為, 心的状態, コーパス, 文末表現, 情緒, 感情

Construction of Text-Dialogue Corpus with Dialogue Act Tags Representing Mental States

Masato TOKUHISA Kousuke MAETA Jin'ichi MURAKAMI Satoru IKEHARA

Department of Information and Knowledge Engineering, Faculty of Engineering, Tottori University

4-110 Koyama-Machi-Minami, Tottori, 680-8552, Japan

E-mail: { tokuhisa, s032049, murakami, ikehara } @ike.tottori-u.ac.jp

Abstract This paper reports results of the construction of text-dialogue corpus annotated with dialogue act tags that represent mental states. The mental state is required for emotion reasoning from dialogues. The dialogue act tags of mental states, however, suffer from the ambiguity. We design the dialogue act tags distinguished by explicitness in language expression. Then, we annotate the tags to our text-dialogue corpus including 30,000 sentences to evaluate the reliability of them.

Keyword text-dialogue, dialogue act, mental state, corpus, suffix expression, emotion, affect

1 はじめに

本稿は、心的状態を表す対話行為タグを付与したテキスト対話コーパスを構築するとともに、付与したタグの共通性を評価する。

従来より、対話における感情に関する言語表現の分析は、タグ付きコーパスの構築を通じて行われてきた。しかし、使用されるタグは粒度が荒いため、表現される事態と情緒の関係を分析することには使用しにくいという問題がある。情緒の分析には、情緒主の持つ生理状態、ゴール・プラン、認識などの心的状態に関する情報が必要である。そのために粒度の細かいタグを付与したコーパスの試作が行われている[1]。しかし、言語表現の表すことのできる心的状態は、区別が難いため、タグ付与が不安定になるという問題がある。

そこで、本稿では、言語表現の表すことのできる範囲で心的状態を表すように対話行為タグを設計し、テキスト対話に対話行為タ

グを付与することを目的とする。

タグ付与は、複数の付与者により行う。付与者間のタグの共通性、付与者間で定めた正解タグに対する各付与者の正解率、および、正解タグに対する他者による同意率について評価する。

2 心的状態を表す対話行為タグの設計

2.1 情緒推定に必要な心的状態

心的過程をシミュレーションするために、心のモデルが提唱されている[2]。そこに示される主たる過程は、次の通りである：

知覚器を通じて外界を認識したり、生理状態を知覚したりすることで、欲求が生じる。生じた欲求に対してゴールを立て、プランの候補を立てる。プランの実行時の様子を予測し、プラン候補を評価する。評価の良い候補を選択し、プランを生成する。プランを行動に移すと、認識と協調しながら効果器を動作させる。一連の実行には過去の記憶を使用

する。これらの過程に対して情緒の生じる状況と判断されると情緒が生起する。情緒は、行動に割り込みをかけたり、プランを変更させたり、または、ゴールの変更をさせる。さらに、これらの様子は自身が監視しており言語でまとめあげるとともに構造化する。これはエピソード的な記憶の蓄積につながる。

対話文から対話者の情緒を推定する方法として、対話文から対話者の上記の心的過程の状況を読み取ることが挙げられる。たとえば、「お腹がすいている。でも、財布を家に忘れてきたんだ。」という発話をしたときに、生理と認識の状態が直接的に読み取られ、間接的には「食事」というゴール、「学食でご飯を食べるか、生協でパンを買う」というプラン、「財布が無いために可能性が無い」という評価などが読み取れる。この状況においてゴールの達成が不可能であるため、対話者に「悲しみ」という情緒が生起していると推定できる。

このように情緒推定を行うためには、「生理状態」、「欲求」、「情緒」、「ゴール」、「プラン」、「予測」、「評価」、「認識」、「行動」、「記憶」、「言語」という心的状態を理解することが必要である。

2.2 心的状態を表すタグ

対話行為タグは、「行為タイプ」、「対象タイプ」、「補助属性1」、「補助属性2」の4つ組で構成する。表 2.1 に構成要素についてまとめる。また、以前の設計[1]と比べて、文脈に依存しない判断でタグが付与できるように変更している。

行為タイプは、対話行為として一般的な種類ものを挙げている。以前からの変更点は、「受諾」、「拒否」、「肯定」、「否定」を廃止し、「はい」、「いいえ」を導入したことである。対話において、「要求」された場合には「受諾」／「拒否」で、「確認」された場合には「肯定」／「否定」で応答するのだが、実際の対話においては、こうした判断がつきにくい場合があるので、粒度を荒くした。

対象タイプは、心的状態に関わるものを挙げている。以前からの変更点は、次のとおりである。

- 「ゴール」の廃止：「欲求」や「プラン」と区別が難しいので廃止した。
- 「予測」、「記憶」、「認識」の廃止：予測・認識される事柄や記憶された事柄に、生理状態、人物の行為、物事の属性が含まれることがある。この場合、判断が分かれるので廃止した。
- 「評価」の廃止：プランに関する事柄に対する属性(嗜好性、危険性、時間性、可能性、存在性)を「評価」としていた。しかし、述べられた属性がプランに係るのかどうかは、前後の文脈を広く見渡す必要がある。そのため、判断がすぐにはできない。
- 「行動」の廃止：「プラン」との区別ができないので廃止した。なお、「プラン」には、過去のプランや現在実行中のプランも該当するものとした。

表 2.1 対話行為タグの構成要素

分類	要素	説明
行為タイプ	質問	仮説が述べられていない疑問文
	確認	仮説が述べられている疑問文
	伝達	事態を述べる文
	要求	聞き手の行動を定める文
	はい	「はい」、「うん」など承諾を表す文
	いいえ	「いいえ」、「ちがう」など否定を表す文
	その他	挨拶や相づちなど上記以外のもの
対象タイプ	生理	身体の内面や表面で感じる事
	欲求	生理や社会的に起因して生じる願望
	情緒	喜び、悲しみなどを感じる事
	プラン	人物の行動の1つ以上の系列
	属性	物事に対する価値や性質
	その他	上記意外のもの
補助属性1	過去	対象は過去のものである
	非過去	対象は過去のものでない
補助属性2	否定	対象は想像上であり、否定する
	想像	対象は想像であり、否定しない
	実在	対象は実在した、あるいは、実在する

対象タイプの改編にともない、欠落する情報があるので、補助属性1と2を追加している。

- 「補助属性1」の追加：日本語において、過去の時制と過去でない時制という区別が可能である。「記憶」と「認識」の廃止に伴い追加した。
- 「補助属性2」の追加：否定文を表すために「否定」を設けた。行為タイプにおいて「否定」や「拒否」を廃止したため、ここに追加した。また、「予測」の廃止に伴い、「想像」を設けた。また、否定文では、想像上の事柄を述べた上で否定を表す単語を用いるので、本質的には「想像」である。すなわち、本来は「想像」vs「実在」という対立関係があり、「想像」の下位に「想像を否定」と「想像を肯定」という層があるとまとめられる。これを効率化して、ここでは、3値を設けている。

以下に、対話文に対するタグ付与の例を示す：

「昨日、雨が降ったよ。」〔伝達, その他, 過去, 実在〕

「昨日は雨だったの？」〔確認, その他, 過去, 想像〕

「うん、そうだよ。」〔はい〕

「10時まで空港に行くよ。」〔伝達, プラン, 非過去, 想像〕

「僕は、日替わりランチ。」〔その他〕

3 テキスト対話へのタグ付与

3.1 タグ付与の対象

対話行為タグの付与対象はテキスト対話とする。この対話は漫画ちびまる子ちゃんの第1巻から第10巻までから抽出したものであり、表情・情緒タグの付きコーパスとしてまとめている[3][4]。

3.2 タグ付与の方法

タグの付与者は、大学生・大学院生であり、上記の表の理解を

行い、タグ付与作業の前に1話ぶん程度を練習し、熟練者により正否の指導を受けた。

タグ付与は、台詞の文の単位に付与する。1つの箇所には2人の付与者が、独立にタグを付与する。このタグを「一時タグ」と呼ぶ。次に、2人の一時タグを比較して、両者の協議により「正解タグ」を決定する。したがって、1つの箇所に2つの一時タグと1つの正解タグが付与されることになる。

文脈の影響を受けにくいようにタグを設計しているが、完全に文脈に自由なわけではない。したがって、タグ付与の過程で前後の文を参照しながら、タグを付与する。また、1文に複数の述語が見られる場合には、主節に対してタグを付与する。

文脈を参照しても曖昧性が解消できず、タグが定め難いことがある。その場合、表 2.1 の上の方のラベルを優先することとする。

3.3 タグ付与の様子

表 4.1 に一時対話行為タグ2名ぶん、および、正解対話行為タグを付与したコーパスの一部を示す。通番を表す#の値は本稿の説明用である。なお、参考として、表情タグと情緒タグを載せる。表情が「普通」の箇所について表情タグと情緒タグを補強している。

さて、対話行為タグの具体例について説明しよう。たとえば、#1において、省略箇所を補完すると、X 者は、[確認, 属性, 非過去, 实在]という対話行為タグを付与し、Y 者も同一のタグを付与していることになる。このタグが正解タグと認められている。

#4 において、台詞がない。そこにはタグを付与しない。

#5 において、X 者の補助属性2には「想像」が付与されているが、Y 者には無い。協議の結果、「想像」を付与することが決まった。同様に、#7 においては、X 者とY 者の間で対象タイプが不一致であり、#9 においては、行為タイプが不一致である。これらも協議により正解タグが決められた。

#8 においては、話し始めが「やめてよ、」であるのだが、主節では「気持ち悪い。」となっている。付与者は、主節に対してタグを付与するので、[伝達, 情緒, 非過去, 实在]という結果となった。

3.4 タグ付与の結果

対象としたコーパスには、ナレータを除いた台詞が 29,537 文 (373,024 文字)であった。正解タグは、26,603 個であった。

対話行為タグの分布を表 4.2 に示す。タグは、次の組み合わせにより形成されるので 111 通りを想定していた。

- 「はい」、「いいえ」、「その他」は対象タイプをとらない。
- 対象タイプには補助属性がつく。
- 「質問」、「要求」は、「实在」の補助属性をとらない。
- 「要求」は「過去」の補助属性をとらない。

しかし、実際に使用されたのは 100 通りであった。全体の 31% が[その他]であり、情緒・プラン・属性・その他についての伝達が、10%台であった。8%が[要求, プラン]であった。具体例を表 4.3 に示す。なお、[その他]の具体例には、「ネエ。」、「はっ。」、「オハヨー。」などがある。

4 タグの信頼性の評価

4.1 評価基準

付与したタグの信頼性を、「共通性」、「正解率」、「同意率」により評価する。「共通性」は2人の付与者により別々に付与されたタグの一致について評価する。「正解率」は、正解タグと一時タグの一致について評価する。「同意率」は、新たな対話行為タグの解析者が正解タグを検査して、同意できたタグの割合で評価する。

「共通性」により本タグセットの判断の難しさを判断する。「正解率」により付与者の判定能力を判断する。「同意率」により正解タグの正しさを判断する。

表 4.1 対話行為タグ付きコーパスの例 (台詞は、[4] 第5巻・第2話より引用)

#	頁	コマ	話者	台詞	一時対話行為タグ(X者)	一時対話行為タグ(Y者)	正解対話行為タグ	表情タグ	情緒タグ
1	22	3	まる子	うちの樅の木は小さいね。	確認 属性	確認 属性	確認 属性	悲しみ	悲しみ
2			お姉ちゃん	仕方ないじゃん。	伝達 属性	伝達 属性	伝達 属性	幸福	悲しみ
3		4	お姉ちゃん	ぎゃあっ!	伝達 情緒	伝達 情緒	伝達 情緒	驚き, 恐れ(青)	驚き, 恐れ
4			まる子					驚き	驚き
5	23	1	お姉ちゃん	まる子、あんた樅の木の本に、金魚の死骸埋めたでしょ。	確認 プラン 過去 想像	確認 プラン 過去	確認 プラン 過去 想像	なし	
6		2	まる子	そうだよ。	はい	はい	はい	普通	なし
7				だって肥料になると思っ。	伝達 その他 想像	伝達 属性 想像	伝達 その他 想像		
8			お姉ちゃん	やめてよ、気持ち悪い。	伝達 情緒	伝達 情緒	伝達 情緒	嫌悪(汗)	嫌だ
9		3	お姉ちゃん	さあ飾り付けを始めるわよ。	伝達 プラン	要求 プラン	要求 プラン	幸福	喜び
10			まる子	ほいきた。	はい	はい	はい	幸福	喜び

表 4.2 対話行為タグの分布

	生理			欲求			情緒			プラン			属性			その他			空白
質問	0	1	-	0	0	-	0	1	-	6	113	-	0	10	-	0	55	-	-
	0	13	-	0	9	-	2	10	-	9	340	-	1	70	-	5	503	-	-
確認	0	13	2	1	2	2	1	12	6	8	141	43	0	26	8	4	36	15	-
	5	48	12	8	23	6	4	57	19	55	349	19	27	300	75	38	364	150	-
伝達	2	3	66	3	2	22	4	8	101	47	56	610	20	32	239	24	42	352	-
	18	61	391	53	32	245	31	81	2,480	293	1,079	681	270	335	2,489	205	268	1,660	-
要求	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	3	1	-	0	3	-	6	3	-	262	1,897	-	1	5	-	3	12	-	-
はい	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	616
いいえ	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	111
その他	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	8,348

※ 総数 26,603 件。縦軸は行為タイプ、横軸は対象タイプである。両タイプの組み合わせにより作られる 3x2 の枠は、上段が「過去」、下段が「非過去」、左が「否定」、中が「想像」、右が「実在」に該当。たとえば、〔伝達、プラン、過去、想像〕は 56 件。「-」は出現し得ないことを表す。

表 4.3 コーパスにおける対話行為タグと台詞の例

〔伝達、属性〕	〔伝達、プラン〕	〔情緒、伝達〕
だ…大丈夫。 まだまだだね。 すっごく綺麗だった。 花輪君、偉いね…。 ちっとも可愛くないー。	今から飲もう。 南の島へ行ったんだ。 その子が送ってくれたのさ。 今からタイル拾いをやるぞ。 しょうがねえ、やっつやるか。	あつ。 ヒェ〜。 アハハ。 怖くないぞ。 心配だな…。
〔伝達、その他〕	〔要求、プラン〕	
忘れ物だよ。 別にこれといって、いないけど…。 後で後悔しても、知らないぞ。 おもちゃも有るよ。 残すとお母さんに怒られる。	やまの屋へ行こう…。 サッカー部集合。 見してー。 買ってよ、買ってよ。 よそ見しながら食べちゃダメでしょ。	

表 4.4 一時タグの比較結果

巻	付与者	一致率	一致数	付与数	κ 値
1	H 者-L 者	59.3%	1,791	3,018	0.4998
2	H 者-M 者	65.7%	1,840	2,802	0.5803
3	I 者-N 者	64.2%	1,754	2,731	0.5804
4	I 者-O 者	58.4%	1,449	2,482	0.5259
5	J 者-O 者	59.6%	1,346	2,260	0.5500
6	J 者-L 者	61.5%	1,529	2,486	0.5652
7	J 者-N 者	64.1%	1,775	2,767	0.5948
8	K 者-N 者	57.9%	1,643	2,837	0.5353
9	K 者-M 者	55.9%	1,507	2,696	0.5204
10	K 者-O 者	57.2%	1,458	2,550	0.5287
総合		60.4%	16,092	26,629	0.5500

タグの一致を数値化する際、「κ 値」と「一致率」を使用する。κ 値は[5]を用い、一致率は次の式を用いる。

$$\langle \text{一致率} \rangle = \frac{\langle \text{一致したタグの数} \rangle}{\langle \text{台詞の数} \rangle} * 100 (\%)$$

4.2 共通性

一時タグの間の一一致を調べた結果を表 4.4 に示す。次のことが言える：

- 第5巻から7巻までが、総合と比べて一致率が良好である。J 者は、8人の付与者の中で、最も対話行為タグの定義に詳しい者であり、揺れが少なかったためであると予想される。
- 以前の結果[1]では、2者の一致が 57%であったが、本稿の結果では、60.4%であったことより、タグの定義の変更による効果があったと思われる。

4.3 正解率

正解タグと各者の一時タグを比較した結果を表 4.5 に示す。次のことが言える：

- 総合の κ 値が 0.7318 であり、それを上回った付与者は、I 者、J 者、K 者である。これら3者は安定的にタグが付与できていたと言える。

表 4.5 正解タグと一時タグの比較結果(κ 値順)

順位	巻	付与者	一致率	一致数	付与数	κ 値
1	2	M	89.7%	2,514	2,802	0.8751
2	3	I	86.4%	2,358	2,730	0.8364
3	7	J	84.2%	2,329	2,767	0.8191
4	10	K	81.9%	2,088	2,550	0.7978
5	4	I	82.4%	2,045	2,482	0.7949
6	5	J	80.7%	1,825	2,262	0.7813
7	6	J	80.7%	2,005	2,486	0.7798
8	9	K	79.5%	2,143	2,696	0.7745
9	8	K	78.6%	2,229	2,837	0.7613
10	1	L	79.1%	2,388	3,018	0.7474
11	6	L	73.5%	1,828	2,486	0.7037
12	8	N	72.7%	2,063	2,837	0.6987
13	7	N	72.6%	2,010	2,767	0.6932
14	3	N	73.5%	2,006	2,731	0.6883
15	4	O	69.8%	1,732	2,482	0.6556
16	9	M	68.2%	1,840	2,696	0.6531
17	1	O	68.1%	1,736	2,550	0.6470
18	1	H	71.9%	2,167	3,016	0.6437
19	2	H	70.6%	1,979	2,802	0.6361
20	5	O	65.8%	1,488	2,261	0.6212
総合			76.6%	40,773	53,258	0.7318

- 変動の大きかった者は M 者である。κ 値が 0.8751 と最大値を得た巻と、0.6531 と総合を下回った巻があった。
- 総合と比べて下回った付与者は、H 者と O 者である。しかし、κ 値は 0.6 を越えており、それほど悪い結果ではない。

4.4 同意率

正解タグに対して付与者とは別の者が検査し、同意の得られた割合を表 4.6 に示す。この表より次のことが言える：

- コーパスよりランダムに抽出した 3,404 件の対話行為タグに対して、3,017 件については同意が得られた。したがって、正解タグへの同意率は 89% である。
- 第1巻において、同意率が 84% であった。第1巻の担当者は、H 者と L 者であるが、表 4.5 によると H 者は下位レベル、L 者は中位レベルである。そのため両者の協議で作られた正解タグには誤りが残ったと考えられる。しかしながら、第4巻と1ポイント差であり、著しく劣位であるわけではない。

以上より、本稿の対話行為タグの信頼性が確認された。

表 4.6 正解タグの検査結果

巻	同意率	同意数	サンプル数
1	84%	239	286
2	86%	287	332
3	86%	381	445
4	85%	339	397
5	92%	284	309
6	96%	239	250
7	95%	237	249
8	88%	363	413
9	92%	340	371
10	88%	308	352
総合	89%	3,017	3,404

5 不同意箇所の分析

5.1 不同意の分布

同意の得られなかった 387 カ所の内訳を表 5.1 にまとめる。〔その他〕、および、〔伝達、その他〕と誤って付与されていたことが目立つ。元々〔その他〕が多く付与されていたためである(表 4.2)。第 4.4 節のランダム抽出では、〔その他〕が 684 件抽出されており、〔その他〕に対する同意率は 70% である。

5.2 不同意の具体例

以下に誤り事例を示す。

1) 誤りが〔その他〕で正しくは〔伝達、情緒〕の場合

次に示す例は、〔その他〕が付与されていたが、〔伝達、情緒〕を付与すべきものである：

「どきっ！」、「うー。」、「こっ、これは！」、「あっ、こらまる子！」

表 5.1 同意されなかった正解タグの内訳

順位	件数	誤	正
1	88	〔その他〕	〔伝達・情緒〕
2	28	〔その他〕	〔はい〕
3	26	〔伝達、その他〕	〔伝達・属性〕
4	24	〔その他〕	〔伝達・属性〕
5	13	〔その他〕	〔伝達・その他〕
6	11	〔伝達、プラン〕	〔伝達・その他〕
7	10	〔その他〕	〔伝達・プラン〕
8	10	〔その他〕	〔要求・プラン〕
9	10	〔伝達、その他〕	〔伝達・プラン〕
10	8	〔その他〕	〔質問・その他〕

オノマトペの扱いを呈示していなかったため、判断が揺らいだものと思われる。

2) 誤りが〔その他〕で正しくは〔はい〕の場合

次に示す例は、〔その他〕が付与されていたが、〔はい〕を付与すべきものである：

「いいもん。」、「うん。」、「はいはい。」、「そうだね。」、「へー、そうだったのかあ。」

表面上、「肯定」や「受諾」を示しているならば、〔はい〕を付与すべきであるが、怒りながらの発話や、単なる相づちに近いものは付与されていなかった。また、格要素と述語の見られる台詞に〔はい〕を付与しにくかったものと思われる。

3) 誤りが〔伝達、その他〕で正しくは〔伝達、属性〕の場合

次に示す例は、〔伝達、その他〕が付与されていたが、〔伝達、属性〕を付与すべきものである：

「まだ工作の宿題もあるんだっ。」、「ほ…っ、これで天下泰平ね。」、「英語もペラペラだよ。」、「うーん、うまく書けたよ。」、「おとなしく勉強すりゃいいのに。」

存在性を述べる台詞は、〔伝達、属性〕を付与することとしていたが、「ある／なし」の動詞を参照すると「属性」という判断が行い難かったものと思われる。ダ文(名詞述語文)も同様に、属性という判断に至らなかったことと思われる。

また、「うまく書けたよ。」のように「可能」を表す台詞や、「～すればいいのに。」のように「義務・推奨」を表す台詞は、属性を表す台詞という判断と同時に、〔伝達、プラン〕という判断もできる。そのためタグの選択が曖昧になり、抽象度のある〔伝達、その他〕が付与されたものと思われる。

5.3 誤りを無くすためには

一般に、対話行為タグの定義を複雑にしすぎると、付与者が理解できないことが問題となる。本稿の対話行為タグの定義は、実際に使用したところ運用の可能なものであることが分かった。しかし、判断の揺らぐところが見られたので、次のような定義の補強が必要である：

- 病名、震え、発汗、心拍など身体の内部・表面の状態や変化

を言及するならば、対象タイプとして「生理」を使用する。

- 望む状態や事象を言及するならば、対象タイプとして「欲求」を使用する。
- 「泣く」のように情緒反応を言及するならば、対象タイプとして「情緒」を使用する。
- オノマトペにより感情が表されるならば、対象タイプとして「情緒」を使用する。
- 対象タイプ「プラン」には補助属性として「可能」、「義務」、「推奨」、「終了」を追加し、あえて「属性」の範囲とオーバーラップさせる。

また、複文や重文のように、複数の述語のある場合には、主節に対するタグ付与を作業基準に含めていた。しかし、倒置の表現において、惑わされることが見られた。そこで、次の作業基準の追加が必要である：

- 節(述語)の単位でのタグを付与する。

6 情緒推定に向けて

対話状態をとらえて情緒を推定する方法を実現する上で、心的状態を対話文から解析することが必要とされている[6]。心的状態は、対話者の持つ信念を表す空間に展開する。信念空間上に対話者の情緒の生じる状況が形成されることを検出することで、情緒を推定する。この原理に基づく情緒推定を机上シミュレーションした結果を図 7.1 に示す。

ここで、b1 は信念を表す述語である。第1引数は信じる者である。第2引数は心的状態である。recognition は認識を、oplan は他者に要求するプランを、plan は自身のプランを、evaluation は評価を、emotion は情緒をそれぞれ表す。また、第2引数で b1 の述語を使用することによって、対話相手の信念を入れ子で記述することができる。各述語には識別番号が付けられている。close は信じることを止めることを表す。

たとえば、3姉の発話で、姉は「まる子が、『まる子が部屋を片付ける』というプランを持つ」と信じる(b3)が、4まる子の発話のとおり、断られる。そのため、信念 b3 は棄却され、プランが cancel となる。この関係を検出して、姉に怒りが生じたと推定される(b6)。

この実現には、発話文から心的状態を抽出する知識、および、心的状態から別の心的状態や情緒を推論する知識が必要である。今後、前者の知識の開発を本コーパスを用いて行う。

7 おわりに

心的状態を表す対話行為タグについて、以前のもの[1]を改良した。実際にテキスト対話にタグ付与を実施した。最終的に得られたタグは、同意率が 89%であり、信頼できるタグが得られた。また同時に、タグの定義が妥当であったと言える。

```
1 姉：「ちょっと、まる子っ。」
   [その他]
2 姉：「また部屋を散らかしてっ。」
   [伝達, プラン, 過去, 実在]
   b1:b1 (姉, recognition::[まる子が部屋を散らかす])
3 姉：「さっさと片付けてよ。」
   [要求, プラン, 非過去, 想像]
   b2:b1 (姉, oplan::[まる子が部屋を片付ける, accept])
   b3:b1 (姉,
   b1 (まる子, plan::[まる子が部屋を片付ける, candidate]))
4 まる子：「嫌だよ。」
   [いいえ]
   b4:b1 (まる子, plan::[まる子が部屋を片付ける, cancel])
   close (b3)
   b5:b1 (姉,
   b1 (まる子, plan::[まる子が部屋を片付ける, cancel]))
   b6:b1 (姉, emotion::[怒り, 反抗, まる子])
5 まる子：「散らかってた方が便利でしょ。」
   [確認, 属性, 非過去, 想像]
   b7:b1 (まる子, evaluation::[散らかってた方が便利です])
   b8:b1 (まる子, emotion::[好ましい, 補助, 散らかってた方])
   b9:b1 (姉, emotion::[怒り, 反感, まる子])
```

図 7.1 対話の進行と信念空間の例 (机上シミュレーション)

謝辞

本研究は、博報『ことばと文化・教育』研究助成の下で行いました。タグ付与に協力して下さった本研究室メンバーの水田理夫君、黒住亜紀子さん、植村和久君、原田誠君、東真之介君、吾郷裕昭君、中道龍三君に感謝します。

参考文献

- [1] 松浦, 徳久, 村上, 池原: 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, 思考と言語, TL2006-62, pp.31-36, 2007.
- [2] Okada, N. and Endo, T.: Story Generation based on Dynamics of the Mind, Computational Intelligence, Vol. 8, No. 1, pp. 123-160, 1992.
- [3] 徳久, 村上, 池原: 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.
- [4] さくらももこ: ちびまる子ちゃん, Vol.1-10, 集英社, 1987-1993.
- [5] Cohen, J.: A Coefficient of Agreement for Nominal Scales, *Education and Psychological Measurement*, Vol.20, No.1, pp.37-46, 1960.
- [6] 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, 電子情報通信学会技術研究報告, 思考と言語, TL2001-25, pp.21-28, 2001.