

目的論的意味理解に基づく対話システムへのモデル提案

A proposition for a model of conversational systems based on teleological semantic understanding

菅原 朔*¹
Saku Sugawara

相澤 彰子*^{2*1}
Akiko Aizawa

*¹ 東京大学
The University of Tokyo

*² 国立情報学研究所
National Institute of Informatics

In this paper, we investigate a simplified model for character-based dialogue system. Inspired by Millikan's teleosemantics, we primarily focus on 'directive' (vs. 'descriptive') and 'direct' (vs. 'indirect') perspectives of utterances, and extend Dore's 'primitive speech acts'. We also conducted a preliminary study using dialogues extracted from Japanese novels to show the potential of the proposed model.

1. はじめに

既存の対話システム研究の発話生成には、[江頭 12] のような話題提供を目的とした雑談対話システムにおけるユーザー入力の属性分類や、[目黒 14] のようなルールベースと統計的手法の組み合わせによる手法がある。これらを発展させ、さらに自然で単純な応答を生成するためには、人間の発話の行為的側面を分類して、発話生成に用いられる属性やルールを言語行為の観点から精緻化・簡略化することが有用であると考えられる。そこで本研究では、これまで主に哲学・言語学の分野で提示されてきた表象や言語記号に関わる諸概念を用いて、対話システムにおける発話の理論的な枠組みを整理し、入力として与えられる発話の行為的な側面を分類するためのモデルの提案を試みる。

本稿における発話の分類は大きく次の二つの観点に基づく。第一に、発話を (1) 何らかの事態の内容を記述する側面 (記述的側面) と、(2) 行為として何らかの振る舞いをする側面 (指令的側面) の二つに区別する。第二に、発話を理解する段階として、(1) 言語表現が意味する事態を即座に認識するような直接的な段階 (直接的な言語行為) と、(2) さらに推論や前提知識との比較を挟んで内容の確認を行うような間接的な段階 (間接的な言語行為) の二つを想定する (2 節)。

本稿では、指令的かつ直接的な発話だけに着目して、発話行為の分類を試みる (3 節)。この分類は、会話要素の中でシステムが比較的あいまい性なく解釈できる行為だけに注目しており、このような簡略化したモデルを用いることで、入力として与えられた発話から行為的な側面を抽出して、内容に関する間接的な推論を挟まずに応答を生成する会話システムの設計指針が得られる。これにより、たとえば推論の困難な発話や内容に曖昧さがある発話に対して、破綻のない自然な応答を行い、相手による更なる説明を促す、などが可能になると考えられる。

さらに本稿では、提示した行為分類に基づいて振り分けのためのパターンを作成して、小説から引用した会話文を振り分けるテストを行い、分類の整合性能を評価する (4 節)。なお、本研究で対象とするのは文字表現のみを伝達手段とする対話である。これは、文字表現のみのメディアでは、身振りや音声で表現していた側面が例えばオノマオバや顔文字のように言語化・記号化され扱いやすくなる、という分析上の利点によるものである。

連絡先: 菅原 朔, 東京大学大学院 情報理工学系研究科 コンピュータ科学専攻, saku.sugawara@gmail.com

2. 発話の分類のための観点

2.1 記述と指令の区別

生物学の哲学や言語哲学の分野におけるルース・ミリカンなどの目的論的な意味論では、言語記号は志向的表象 (intentional representation) の一種であると定義される [Millikan 84]。志向的表象とは、文字通り何らかの事態を志向的に表すことを達成すべき機能として持つ表象のことであり、生物の認知システム内の知覚情報や、システム外に作り出されて個体間で伝達されるようなサイン (その一例が人間の扱う文字である) などが相当する。そして志向的表象の意味は、表象を認識して解釈する側が持つ、生存などの目的に基づいて定義される。

志向的表象には、何らかの事態を記述する側面と同時にそれに応じた行動を導く指令的な側面がある。たとえばハチのダンスは、記述的には蜜の位置を表し、指令的には仲間のハチにそこへ向かえよという行動を表している。このように一般には記述的側面と指令的側面は独立ではないが (ハチのダンスの例のように両者の側面が分化していないものは、オシツオサレツ表象と呼ばれている)、文章や信念といった言語的表現においては切り分けることが可能である [Millikan 04]。たとえば紙に書かれた「地球は丸い」という文は、ある事実を記述するだけで、それを読んだ者に何らかの動作を要請するものではない。また「逃げろ!」という発話は、聞き手に避難を指示するという側面を持つが、何らかの事態ではなく、強いて言うならば聞き手が逃げろということについて記述している。なお、ここで言う指令という語は、相手に何らかの具体的な行為を命令するという意味ではなく、相手の何らかの応答・行為に結びついていくという意味で緩さがあるという点に注意するべきである。

この観点から見ると、発話における指令的側面はオースティンによる発話内行為 [Austin 62] (発話自体が何らかの行為をしていると見なせるもの) に類似している。たとえば「雨が降っている」という発話は陳述という行為であり「雨が降っていますか?」という発話は質問という行為であると見なせる。本研究はこの点に着目し、発話が記述的内容として何を意味しているかではなく、指令的側面として聞き手に何を示すものであるかを分析する。指令的側面の分類、すなわち本研究で言う行為的側面の分類については、3 節で詳細に述べる。

2.2 直接性と間接性の区別

ミリカンは、心理学的な知見 [Gilbert 93] に基づいて、言語理解には知識や推論を介在させない直接的な知覚 (direct perception) の段階があると主張している [Millikan 04]。直接的な知

表 1: 発話における記述・指令と直接・間接の分類

	直接	間接
記述	命題的な意味内容	
指令	本研究	間接的言語行為

覚はいわば意味論的な写像関数 (semantic mapping function) の適用そのものであり、ミリカンはこれを翻訳 (translation) と呼んで推論 (inference) と対比させている。すなわち、通常の認知過程では翻訳に次いで知識との照合や他の信念との組み合わせによる推論が行われるのであり、その両者は区別されるべきだと提示しているのである。

実験によれば、人は過度の認知的重圧のもとにあるとき、たとえば話を聞きながら同時に 1000 から順に 3 ずつ引いていく作業をしなければならないとき、聞いたことをただ信じてしまう傾向が強い。他にも、「火星には宇宙人がいる」と小さな子供に言えばもしかしたら信じてしまう (想像してしまう) がもしれないのに対して、私たちがその言葉を間違いなく信じないであろう理由は、火星に宇宙人が存在しないことを示す証拠となるような知識を数多く持っているからである。これらの考察から、発話の直接的な即座の理解と、推論を経た間接的な理解に対しては、時間的な順序を定めることが実際的に可能であり、本研究においても、対話では推論や知識の介入を一切認めない発話理解が規定できるものとする。このとき、対話システムは入力として与えられた発話を相手の言語行為として無批判に解釈する。また、一切の推論を認めないことから、文字表現から直接的に指示される事態以外の理解を導くことはできない。たとえば、

例 1:

- A 「明日お茶でも行かない？」
B 「明日は出かける用事があるんだ」

という応答関係において、B が誘いを断っていることを A は推論できない。B が「行けない」と明言しない限り、A は B が来ないことを理解しない。このことは、オースティンとサールの用語に従えば、間接言語行為を含めない発話内行為のみを扱うということの意味する (表 1: 記述・指令と直接・間接の分類)。

上記の例 1 における対話システムの運用を考えると、システムの発言 A (後述する表 4 の分類では「返事の要求」) に対してユーザー B は「行けない」(後述の分類では「返事」という明示的な返答をしていないのだから、システムは「返事がもらえていない」と判断することになる。

2.3 二種の区別による利点

発話の指令的かつ直接的な側面のみを扱うという観点は、やりとりの内容に関する誤認識が許されないようなシステムの設計において特に重要となる。すなわち、対話システムの目的によっては、ユーザの発言内容をシステムが確率的に推論するよりも、再び同じ旨の質問を繰り返して明示的な応答をユーザから引き出すことで、間接言語行為の意味解釈の不確かさを回避するという対話戦略が好ましい場合も想定される。このような場合に、解釈の範囲を直接的な応答に絞ることによって、対話における曖昧さを排除することが可能になる。

たとえば顧客に商品を選択させる業務的対話システムを考える。「商品 X でよろしいですか」という「返事の要求」に対してユーザーが「X は素晴らしいよ」と「評価」を述べるに留

まる場合、発話内容の推論を行う以前の段階として行為的側面だけ判別すれば、返事をしていない発話だと即座に見なして質問を繰り返すことができ、やりとりを明示的に進めることが可能になる。すなわち、

例 2:

- システム「購入は商品 X でよろしいでしょうか？」(質問)
ユーザー「X は素晴らしいよ」(質問への返事ではない)
システム「よろしいでしょうか？」(再度返事を要求)
ユーザー「はい、よろしく」(質問への返事)

という対話が可能になる。将来的には、本例のような言語行為の指令・応答関係を談話構造における隣接ペアとして捉え、<挨拶-挨拶> や <疑問-肯定/否定> のような定型として学習させることが可能である。人間の発話理解は文法における品詞の機能分類と同規模のルールでなされていると考えられるが、この学習によって、少ないルールで適切な発話を導くことが可能になることも考えられる。

3. 発話行為と属性の分類

3.1 行為分類の検討

オースティンの言語行為論では、発話内の力に応じて発話が五種類に分類されている (表 2: Speech Acts [Austin 75], 説明は [石崎, 伝 01] に依った) が、サールが [Searle 79] で指摘しているように、この分類は発話内行為の分類というよりも英語の動詞の分類を行なっていると言わざるを得ない。重要なのは、オースティンにおいては発話内行為の分類と発話内行為動詞の分類が一致しないということであり、サールはこの観察を間接言語行為の議論に発展させている。一方、本研究では前述のように推論を経ない直接的な側面だけに注目するため、サールの言うような間接性は扱わない。[Searle 79] の第一章結論部で指摘されているように、オースティンの行為分類には、発話内行為としての目標が同一であり、強弱などの属性にこそ違いがあるような例がいくつも挙げられる。この考察から、本研究では発話の言語行為におけるラベルとして、ある程度抽象化された「行為」と態度としての「属性」のふたつを与えることを提案する。

行為の分類として参考になるのは、[Dore 75] に見られるような言語発達における言語行為 (primitive speech acts) である (表 3: Primitive Speech Acts [Dore 75])。本研究では、学習過程の子供に特有の行為である「ラベル付け」「反復」「練習」を除外し、「抗議」を「評価・提示」として拡張、さらに「動作の申し出」を追加することで、Dore の分類を発展させる形で対話における社会的・相互的なやりとりを包含できるような分類を検討した。

具体的には、まず対話を成立させるまでの過程として「呼びかけ」と「挨拶」を想定し、さらに対話が成立してから各発話の次に何が起きるか (対話が継続されるか、別の動作が起きるか) で発話行為を区別した。

対話を継続する場合、返事を要請するかどうかでさらに区別する。要請する場合が「返事の要求」であり、これに対する応答として「返事」が発話される。一方で要請しない場合が「評価・説明」の提示であり、これに対する応答は内容に依存するため一意に定まらない。「返事の要求」は疑問などの形で相手の明示的な応答を要求するが、「評価・説明」はそのような明示的な要求を含まない。発話は明示的でない仕方でも何らかの行為の要求を含意する場合がありますが、その解釈には推論が必要であることから、本研究では「評価・説明」に分類する。

表 2: Speech Acts [Austin 72]

分類	説明
判定宣告型 (Verdictives)	何らかの証拠・理由に基づく判定の行使
権限行使型 (Exercitives)	影響力の主張・権力の行使
行為拘束型 (Commissives)	義務の引き受け・意図の宣言
態度表明型 (Behabitives)	他人の行動に対する一定の態度の表明
言明解説型 (Expositives)	理由・議論・伝達作用の明確化

表 4: 本研究の Speech Acts

行為	内容
呼びかけ	相手に呼びかける, 会話への注意を求める
挨拶	儀礼的・慣習的であり, 感謝や謝罪を含む
返事	肯定や否定といった応答
返事の要求	返事や評価を求める
動作の要求	相手の動作を求める
動作の申し出	自分の動作を提示する
評価・説明	相手や何らかの事態に対する評価や説明を提示する

表 3: Primitive Speech Acts [Dore 75]

行為	内容
呼びかけ (Calling)	相手に呼びかける
挨拶 (Greeting)	挨拶をする
抗議 (Protesting)	反意や拒否を示す
返事の要求 (Requesting answer)	情報を返すことを求める
動作の要求 (Requesting action)	動作を求める
返事 (Answering)	他者の質問に返事をする
ラベル付け (Labeling)	事物や出来事に名前をつける
反復 (Repeating)	発声を繰り返す
練習 (Practicing)	語や発声のパターンを再現する

表 5: 分類例: 分類ごとの発話の分布

行為	発話数
呼びかけ	0
挨拶	1
返事	14
返事の要求	38
動作の要求	3
動作の申し出	0
評価・説明	34
計	90

また, 行為が発生する場合, その行為を相手に求めるか自分が行うかで区別できる. 相手が遂行するよう求める場合が「動作の要求」であり, 自分が遂行することを示す場合が「動作の申し出」である.

3.2 属性分類の検討

行為それぞれに対しては, 態度の付与が生じている. それは例えば強弱であったり, 推測のような曖昧さや断定のパラメータであったり, 尊敬語のような丁寧さ・粗雑さであったりする. 役割語のような言い回しも属性に含まれる. 属性の付与は辞書的な対応関係で解決できるが, 属性の有無は言語に依存する (たとえば日本語では敬語表現があるが英語では明示的には存在しない).

3.3 分類例

行為分類の妥当性について検証するため, 青空文庫の小説から抜き出した 45 組の発話ペアについて, 人出で表 4 で提案した 7 つの行為を割り当てた. 対話文は, 夏目漱石『こころ』『三四郎』『それから』青空文庫の 3 作品から 15 組ずつ, 直前・直後に地の文のない 2 連続の発話をペアとして抜き出した. 行為はひとつの発話に対してひとつ定めるものとする.

その結果は表 5 のようになった. 例を挙げる (出典は 1: 『こころ』, 2 と 3: 『それから』):

- 1: 「すぐお宅へお帰りですか」(返事の要求)
「ええ別に寄る所もありませんから」(返事)
- 2: 「何を見ているんです」(返事の要求)
「あててごらん下さい」(動作の要求)
- 3: 「ええ, ついその先の角です」(評価・説明)
「どうもありがとう」(挨拶)

行為ごとの発話では, 「返事の要求」という「情報の提供を求める行為」と「評価・説明」という「情報を提供する行為」に多くが分布しており, 会話の情報交換手段として良く機能していることがわかる. しかし, コーパスによってこの程度は変わると考えられ, 小説以外の対話文についても分類を試みる必要がある.

4. 分類器の構築と予備的評価

4.1 自動分類の方針

発話文中で観察される以下の 2 種類のパターンを手がかりに, 発話行為を分類することを試みる.

1. 形態素解析に基づく品詞とその文法的機能の組み合わせを参照して定めた文末表現
2. 挨拶や呼びかけなど慣習化された定型句

こうした辞書の構築は手作業で行わざるを得ないが, 項目としては膨大な数になるようなことはないと推測される.

4.2 人手によるパターンの作成と分類結果

実際に発話文からパターンを抽出して, 分類における有効性を評価する実験を行った. まず, 3.3 節と同様に抽出した発話 200 文程度を人手で解析し, 表 4 の 7 分類に対応するパターン 109 個を作成した. 次に, 作成したパターンを用いて, 抽出に使用していない新しい発話ペア 172 組 344 文について自動で分類を行った.

人手により作成したパターン数とその例を表 6 に挙げる. ここで「^」は文頭, 「\$」は文末をあらわす. たとえば「...でしょうか\$」というパターンは, 文末が「...でしょうか」となる文に合致する. パターンを用いた自動分類の結果を表 7 に示す. これらの分類の正確さについて, 人手で処理した正解の分類と比較を行い評価したところ, 約 81 パーセントの発話で分類に成功した (279/344).

4.3 分析

分類に成功した例と成功していない例を挙げる (夏目漱石『三四郎』青空文庫より引用):

成功例

- ・「どうもありがとう」(挨拶: 成功例)
- ・「ええ, 下宿したそうです」(返事: 成功例)
- ・「あててごらん下さい」(行為の要求: 成功例)
- ・「風邪だろう」(評価・説明: 成功例)

表 6: 作成パターン数とパターン例

行為	パターン数	パターン例
呼びかけ	2	^ねえ\$
挨拶	16	^こんにちは\$
返事	27	^ええ,
返事の要求	31	でしょうか\$
動作の要求	6	してくれないか\$
動作の申し出	4	てあげようか\$
評価・説明	23	と思います\$
計	109	

表 7: 実験結果: 発話の分類結果

行為	発話数 (うち成功数)
呼びかけ	0(0)
挨拶	1(1)
返事	35(34)
返事の要求	119(103)
動作の要求	2(2)
動作の申し出	0(0)
評価・説明	187(139)
計	344(279)

失敗例

- ・「そうか」(返事の要求: 失敗例 1, 文末表現だけで判断したため)
- ・「なんだ, それは」(評価・説明: 失敗例 2, 倒置に対応できていない)
- ・「じゃいっしょに行きましょうか」(返事の要求: 失敗例 3, 正しくは評価・説・明だと判断されるべきだと思われる)

分類が正確でなかった例について原因を調べたところ, 以下の問題が観察された:

1. 文法的機能の解釈に推論が必要なものが存在する
失敗例 1,3 で挙げたような終助詞「か」などが該当する。これらの例では疑問表現ではなく「確認・念押し」のような表現として用いられているが, 疑問文との判別には内容的な推論が要求される。本来ならば「返事の要求」などの疑問表現であるのを「評価・説明」と分類してしまったケースもこれに該当する。
2. 形態素解析の時点で倒置表現の解消ができていない
失敗例 2 が該当する。倒置表現の解消のためには内容的判断が要求される場合もあるため, 本稿で用いたパターンに基づく分類法では対応に限界があると考えられる。
3. 定型文化できない短い表現を分類できない
たとえば「先生」といったような名詞のみの発話は, 単独での分類を試みる場合「呼びかけ」か「評価・説明」かが判断できない。これを解決するためには, 前後の発話に「返事」や「返事の要求」があるかを調べる必要がある。したがって, 隣接ペアによる学習で分類可能になることが期待される。

今回の実験では, 人手によるパターンの作成に用いた文数が少なく, 得られたパターンの数も 109 個と小規模なものであったが, 正解判定を行った範囲では比較的高い分類性能が得られた。これは, 使用したコーパスの発話に「返事の要求」と「評価・説明」という, パターンで捉えやすい分類が頻出していた

ことが主因と考えられる。より多くの由来のコーパス・データで判別パターンを洗練させれば, より高い精度の分類が可能になると予想できる。

5. おわりに

本研究では, 対話システムにおける発話の理論的な枠組みを整理し, 入力として与えられる発話の行為的な側面を分類するためのモデルの提案を試みた。実験の結果, 本稿で用いたコーパスの範囲では, 少ないパターン数で発話の行為的側面を比較的精度よく分類できることがわかった。本稿では簡略化のためパターンを用いたルールベースの判断が中心となったが, 文法的機能を中心にしたルールの作成も可能であり, その場合はさらにルールの数が減ることが期待される。

本稿では扱うことができなかったが, 行為分類に加えて態度の属性を付与した発話の組み合わせを学習させれば, たとえば「丁寧な呼びかけには丁寧な返事が返される」といったようなパターンが見出されることが期待される。直接的で簡単な返答であれば記述的側面の理解が無くても可能であり, 2.3 節で挙げたようにシステムの目的に応じた活用が期待できる。今後は, 本研究で提示した行為分類の精度を上げ, 応答パターンの学習などを行うモデルの構築を目指し, 段階的に推論や内容理解の研究へと繋げていきたい。

参考文献

- [Austin 62] Austin, J. L.: *How to Do Things with Words*, J.O.Urmson(ed.), Oxford, Clarendon Press (1962) (邦訳: 言語と行為, 坂本 百大 訳, 大修館書店 (1978)).
- [Dore 75] Dore, D.: Holophrases, speech acts and language universals, *Journal of Child Language*, Vol. 2, Issue01, pp. 21-40 (1974).
- [Gilbert 93] Gilbert, D. T., et al.: You Can't Believe Everything You Read, *Journal of Personality and Social Psychology*, 65, No. 2, pp. 221-233 (1993).
- [Millikan 84] Millikan, R. G.: *Language, Thought, and Other Biological Categories*, Cambridge, Mass., The MIT Press (1984).
- [Millikan 04] Millikan, R. G.: *Varieties of Meaning*, Cambridge, Mass., The MIT Press (2004) (邦訳: 意味と目的の世界, 信原 幸弘 訳, 勁草書房 (2007)).
- [Searle 79] Searle, J.: *Expression and Meaning*, Cambridge University Press (1979) (邦訳: 表示と意味, 山田 友幸 訳, 誠信書房 (2006)).
- [石崎, 伝 01] 石崎 雅人, 伝 康晴: 言語と計算 3 談話と対話, 東京大学出版会 (2001).
- [江頭 12] 江頭 勇佑, 柴田 知秀, 黒橋 禎夫: 雑談対話システムにおける強化学習を用いた応答生成モジュールの選択, 言語処理学会第 18 回年次大会論文集, pp. 654-657 (2012).
- [目黒 14] 目黒 豊美, 杉山 弘晃, 東中 竜一郎, 南 泰浩: ルールベース発話生成と統計的発話生成の融合に基づく対話システムの構築, 人工知能学会全国大会 (第 28 回) 論文集, 2M5-OS-20b-2 (2014) .