

雑談対話における言外の情報を推定するためのデータ収集と分析

Collection and analysis of implicit information in chat-oriented dialogue

光田航 東中竜一郎 牧野俊朗 松尾義博
Koh Mitsuda Ryuichiro Higashinaka Toshiro Makino Yoshihiro Matsuo

日本電信電話株式会社 NTT メディアインテリジェンス研究所
NTT Media Intelligence Laboratories, Nippon Telegraph and Telephone Corporation

In order to realize a chat-oriented dialogue system that can communicate with users smoothly, it is necessary for the system to understand not only what is actually said by the users but also the implicit information conveyed in user utterances. To realize a system that can understand such implicit information, we conducted data collection of implicit information, in which multiple annotators annotated implicit information on human-human chat-oriented dialogue data. Then, we analyzed the collected data to examine the types of implicit information conveyed in chat-oriented dialogue. This paper describes our data collection and the results of our analysis.

1. はじめに

雑談は人間の会話の約 60% を占めるという調査結果や [小磯 15], タスク指向対話の使い勝手を改善するために有効とされることから, タスク指向対話とは異なる, 非タスク指向対話 (いわゆる, 雑談) を行う対話システムの研究が注目を集めつつある [Ritter 11, Higashinaka 14, Higashinaka 15].

しかし, 雑談対話システムの実現には課題が多い. たとえば, 既存の雑談対話システムは, ユーザの発話に対して, キーワードやテンプレートなどの表層的な情報や, 単語や言い換えの情報を利用して, 応答発話の選択, 生成を行うことが多い [Higashinaka 14]. しかし, 実際には, 表層のマッチングや言い換えだけでは適切に対応できない場合が存在する. その端的な場合が, 発話に「言外の情報」が含まれる場合である. 人間同士の対話では, 数少ないやり取りだけで相手の意図を理解したり, 協調的な振る舞いをしたりできる. これは発話の背後にある言外の情報をうまく汲み取って対話を行っているからである. 雑談対話システムがよりスムーズにユーザと対話をするためには, このような言外の情報を適切に理解する必要がある.

我々は, 言外の情報を理解することができる雑談対話システムを目指している. この目標を達成するためには, まず, ユーザ発話から, システムが理解すべき言外の情報を整理しておく必要がある. 本研究では, 雑談対話において, 対話システムがユーザ発話から理解すべき情報を明らかにするために, 雑談対話における各発話に対して, 話し手が伝えたい, もしくは, 聞き手に伝わると考えられる情報 (言外の情報) を収集し, そのデータの分析を行った. 具体的には, 人間同士の雑談対話を収録したテキストチャットデータにおける各発話に対して, 複数の作業者に言外の情報を記述してもらい, 収集した言外の情報の分析を行った. 本稿では, 言外の情報の収集手順, および, 言外の情報の類型化について報告する.

2. 関連研究

雑談対話から含意認識を行う研究として, Zhang らの Conversation Entailment の研究がある [Zhang 09, Zhang 10]. Zhang らは, Dialogue Segment (対話の断片) と仮説 (人手

で作成) が与えられたとき, その真偽を判定するという課題に取り組んでいる. 言外の情報を扱っている点は本研究に近いが, 仮説が数種類の情報に限定されており, 一般的な言外の情報を扱っているわけではない. また, 本研究が扱う言外の情報は, 含意よりも広い概念である. 含意関係認識の定義は, テキストから仮説を推論できるかが問題となる. しかし, 我々の扱う言外の情報は話し手から聞き手に伝わる情報すべてを含む. たとえば, 話し方から伝わる情報だったり, 一般常識を踏まえた推論によって得られる情報も含む.

我々が扱う言外の情報は, 発語媒介行為 [Austin 62] や会話の含み [Grice 75] などの対話理論やプラン認識と談話プランに基づく対話モデル [Litman 87] で議論されている事象が含まれている. 本研究は, これらの理論で扱われている事象が, 実際の対話データにおいて, どのように出現しているか, また, 対話システムがどの程度理解すべきかを, 体系的・定量的に調査したものであると位置づけることができる.

対話以外の分野ではあるが, Document Enrichment というタスクが提案されている [Zhang 14]. これは, 与えられた文書に対して, Wikipedia などの外部リソースを利用して, 元文書と関連する追加の知識を付与するタスクである. 知識は, < 項, 述語, 項 > の三つ組として表現され, これを推定し, 知識として文書に追加することで, 文書処理の性能が向上することが示されている [Zhang 15a, Zhang 15b]. Zhang らの研究は, 言外の情報を外部リソースから取得する有効な手法を提案しており, 今後, 我々も言外の情報を推定する段階において利用を検討したいと考えている.

3. 言外の情報の収集

我々の目標は, 表層や言い換えだけからでは得られない, ユーザ発話の内容 (言外の情報) を理解することである. そこで, 本研究では, 言外の情報を, 発話内容の言い換えとは異なる, 話し手が伝えたい, もしくは, 聞き手に伝わると考えられる情報として定義し, それらを収集する.

収集に際しては, 作業者に雑談対話のテキストデータを見てもらい, 言外の情報 (と考えられるもの) を記述してもらう. 言外の情報を記述するために利用する雑談対話テキストには, 東中らが収集した雑談コーパス [Higashinaka 14] を利用する. このコーパスには, 初対面の話者が 1 対 1 でテキストチャット形式の雑談を行なった対話が 3,680 対話収録されている. この

連絡先: 光田航, 日本電信電話株式会社 NTT メディアインテリジェンス研究所, 神奈川県横須賀市光の丘 1-1, 239-0847, mitsuda.ko@lab.ntt.co.jp

表 1: 収集された言外の情報の例. まとめ上げの数字は, 言外の情報が所属するグループを示す. アスタリスクは代表文を表す. 妥当性判定の結果は, ◎を 4, ×を 1 としたときの各作業によって付与された判定結果である.

話者: 発話	記述作業: 言外の情報	まとめ上げ	妥当性判定
1 A:「こんにちは」	—	—	—
2 B:「こんにちは」	—	—	—
3 B:「最近興味あることなどありますか?」	w ₁ :「B は A に興味がある」	1	4/2/4/3/3
	w ₂ :「B は話題を振りたい」	2	4/3/4/4/4
	w ₂ :「B は A の趣味などを知りたい」	3*	4/2/4/3/3
	w ₂ :「B は A がどんな人か興味がある」	4	4/2/4/3/4
	w ₃ :「B は A の趣味について聞きたい」	3	—
	w ₃ :「B は多趣味」	5	2/2/1/3/3
	w ₃ :「B は好奇心が多い」	6	4/2/2/3/3
	w ₄ :「B は A の趣味に関心がある」	7	4/2/4/4/3
	w ₅ :「B は A と話したい」	8	4/3/4/3/4
	w ₅ :「B は特に話題がない」	9	4/2/2/3/3
w ₆ :「B は A について知りたい」	10	4/2/4/3/3	
4 A:「食べることとかですかね。」	w ₁ :「A は食事が好きである」	1	—
	w ₂ :「A は食道楽」	2	4/3/3/4/4
	w ₂ :「A は食べ物に詳しい」	3	4/2/3/4/4
	w ₂ :「A は料理ができる」	4	2/2/2/4/3
	w ₃ :「A は食いしん坊だ」	5*	3/2/4/4/4
	w ₃ :「A は太っている」	6	2/2/2/3/3
	w ₄ :「A は食べることが好きだ」	1*	4/3/4/4/4
	w ₄ :「A はグルメの話題に関心を持っている」	7	4/3/3/4/4
	w ₅ :「A は B と話したい」	8	4/2/3/3/3
	w ₅ :「A はくいしんぼう」	5	—
w ₆ :「A は食べることが好きだ」	1	—	

中から, 30 対話をランダムに選択し, 本研究で利用する. 今回用いた, 雑談対話テキストの総発話数は 1,257 発話, 平均発話数は 42 発話である.

作業は 3 段階に分けて行った. 作業 1 は言外の情報の記述, 作業 2 は記述のまとめ上げ, 作業 3 は妥当性の判定とした. 各作業を行う作業は全て別の作業者とした. これらの作業の設計にあたっては, Conversation Entailment [Zhang 09], および, 矛盾する事態ペアのデータ収集手順 [Takabatake 15] を参考にした. 以下に, 各作業の詳細を示す.

作業 1: 言外の情報の記述

雑談対話テキストにおける各発話に対して, 作業者が言外の情報を自然文として記述する. 作業者は対話テキストを読み, 文脈を考慮した上で, 各発話に対して言外の情報を列挙し, 各情報を 1 文で記述する. 記述の際は, 対話と照らし合わせてほとんどの場合に成り立つと考えられる言外の情報のみを記述するようにし, 常識的に考えて当たり前である内容や, 述語項レベルで発話を補完した内容は書かないように教示する. 原則, 1 発話に対して 1 文以上記述し, 各文には必ず主語と述語が含まれるようにする. 記述は 6 名の作業者が行い, 各作業者が独立して全対話テキストに対して記述を行う.

作業 2: 記述のまとめ上げ

記述された言外の情報のまとめ上げを行う. 複数の作業者が記述した言外の情報には, 表層は異なっても同じ内容を示すものも含まれており, それらをあらかじめ人手でまとめておくことで, 以降の分析を効率的に進めることができる. 同じ内容を表すかどうかは, 記述された言外の情報に含まれる内容語, および, モダリティやテンスなどの情報が全て一致しているかどうかで判断し, 全て一致していれば, まとめ上げの対象とする. 言外の情報をまとめた場合は, まとめられた情報の中から代表となる言外の情報の一つを選択しておく. この代表となった情報は, 作業 3 に利用する.

作業 3: 妥当性の判定

まとめ上げられた言外の情報に対して, その妥当性 (どの程度成り立つか) を判定する. 記述された言外の情報で示される事象がどの程度妥当であるかは, 人によって判断が異なると考えられる. そこで, 複数の作業者が妥当性を判定する. 妥当性の判定は, ◎: ほとんどの場合に成り立つ (95%–100%), ○: おおた成り立つ (50%–95%), △: ときどき成り立つ・成りたつこともある (5%–50%), ×: ほとんどの場合に成り立たない (0%–5%) からいずれかを選択することで行う. 5 名の作業者が判定を行う. 各作業者は独立して, 全ての言外の情報に対して妥当性を判定する.

これらの収集手順に従い, 言外の情報を収集した結果, 12,723 個の言外の情報が記述され, それらが 11,533 個 (91%) にまとめ上げられた. 表 1 に収集された言外の情報, そのまとめ上げの結果, および, 妥当性判定の例を示す. 以降, 11,533 個の言外の情報を対象とした分析を行う.

4. 言外の情報の類型化

雑談対話を対象とすることもあり, 言外の情報の内容は多岐にわたると考えられる. したがって, 言外の情報を理解するためには, まず, 収集された言外の情報を類型化しておく必要がある.

言外の情報を類型化する方針として, あらかじめ何らかのクラスを用意するトップダウンの手法と, 同一の事例を集めた後にクラスを付与するボトムアップの手法が存在する. 本研究では, これまでにさまざまな提案がされてきた対話理論の知見が類型化に活用できるのではないかと考え, トップダウンの手法を採用する. 知見を活用する対話理論としては, 発話行為, 協調の原理, プラン認識, 談話構造, 共通基盤 (相互信念) を取り上げ, それぞれに対応するクラスを準備した. 該当しないものはその他とし, その中でさらに事例に基づき, クラスを定義した. 以下, 対話理論とそれに基づく類型について詳述する.

4.1 対話理論に基づく類型

発話行為に基づくクラス

発話行為論における発話媒介行為は言外の情報の代表的なものであろう。これは、言外の情報が発話によって引き起こされる行為の内容を示している場合に該当する。たとえば、下記の例では、A が発話を通じて、B にゆず胡椒を推薦するという行為を行っており、それが言外の情報として伝わっている。

- (例 1) A:「(ゆず胡椒は) しゃぶしゃぶにも使えますよ！」
→「A は B にゆず胡椒の良さを知ってもらいたい」

協調の原理に基づくクラス

言外の情報が、グライスの言う、「会話の含み」によって聞き手に伝わる情報を出している場合に該当する。下記の例では、料理の話題において、食事が好きであるということ、助詞「は」で強調して伝えることで、料理が好きではなく、あまりしないということが言外の情報として伝わっている。

- (例 2) A:「食べることは大好きです」→「A は自分であまり料理をしない」

プラン認識に基づくクラス

話者はプランを保持し対話を進めていると考えられるが、発話にはこのプランが陽に出現しない。すなわち、発話の背後にあるプランに該当する内容が言外の情報として考えられる。プランは前提条件 (precondition)、行為本体 (body)、分解 (decomposition)、効果 (effect) の 4 種類からなり、前提条件は計画実行のための条件、行為本体は行為自身、分解は行為を複数の行為に分解したもの、効果は計画によって達成される目標である。この 4 種類のそれぞれについて、言外の情報が考えられる。

プラン認識に基づく言外の情報の例を下記の例 3-例 6 に示す。例 3 では、A が妻であると述べていることから、前提条件を考えれば、A が結婚していることが言外の情報としてわかる。例 4 では、富士山に行くためには一般的に車が必要となるため、A が車を運転するだろうことがわかる。例 5 では、海外に行くためにはパスポートを申請する必要があるため、A がパスポートを持っていることがわかる。例 6 では、ジムに行く人の目的は、運動をして体を動かすことなので、A が体を動かしたいということがわかる。

- (例 3) A:「私は転勤妻なんです」→「A は結婚している」(前提条件)
(例 4) A:「気が向いたら、富士山の方に行ったりしますね」
→「A は車を運転する」(行為本体)
(例 5) A:「私は海外にもたまに行きますね」→「A はパスポートを持っている」(分解)
(例 6) A:「本当はジムに行きたいです」→「A は体を動かしたい」(効果)

談話構造に基づくクラス

言外の情報によって、対話の構造 (談話構造) が変化する場合がある。たとえば、話者が話題を変えたいということを言外の情報として伝えることで、実際に話題が変わる。このように、言外の情報の中には、談話構造に影響を与える目的のものがあり、そのような場合、談話構造に基づくクラスとする。下記の例 7 では、車の話題において、A が車の運転経験について話した後、B に車の運転経験を尋ねている。この発話によって、話題は A に関する車の話題から B に関する車の話題へと変化する。例 8 では、A が出身地について話を切り出すことで、B に出身地について話したいということが伝わり、話題が出身地に関する内容へと変化する。

- (例 7) A:「車乗られますか」→「A は B にも自分のことを話してほしいと思っている」
(例 8) A:「私は北国のほうの出身なのですが、」→「A は B に出身地の話がしたい」

共通基盤に基づくクラス

言外の情報が、共通基盤の形成に関わる場合である。共通基盤 (もしくは、相互信念とも言う) とは、共同活動で必要となる、A と B の双方が信じている命題を指す。共通基盤の形成は、表面には現れにくい、付帯的コミュニケーションによってなされるとされる。すなわち、これは言外の情報である。

下記の例 9 では、B が最近ニュースを見て、国際情勢が心配であると話した後、A が自分も同様のニュースを見たこと B に伝えることで、B に共感を示しているということが言外の情報として伝わる。A と B の間では、ニュースを見た結果として、国際情勢が心配であるという命題が共通基盤となっている。例 10 では、B が関東は数年振りの水不足であると伝えた後、A が理解を示すことで、最近異常気象が起きているという命題が共通基盤として形成される。

- (例 9) A:「それは私もニュースで見ました」→「A は B に共感している」
(例 10) A:「そうみたいです」→「A は異常気象だと思っていた」

4.2 その他のクラス

人物属性に関するクラス

話者の話し方や話す内容から、話者自身の人物属性が言外に現れることがある。そのような言外の情報は、人物属性に関するクラスとする。下記の例では、ランニングを日課しているということから、話者の真面目な性格が言外に現れている。

- (例 11) A:「ランニングは日課になっています」→「A は真面目な性格である」

因果関係に関するクラス

言外の情報が、常識的な因果関係によって推論可能な場合には、因果関係に関するクラスとする。プラン認識に基づくクラスとは異なり、対話の目的には関係しない、一般常識に照らし合わせた際に理解されるべき言外の情報を指す。下記の例では、発話は単に事実を述べているだけであるが、因果関係によって推論することで、外に出られなかったことが言外の情報として伝わっている。

- (例 12) A:「インフルエンザにかかり寝てましたー」→「A は外に出たくても出られなかった」

ドメイン知識に関するクラス

言外の情報が、特定のドメイン知識に関するもの場合には、ドメイン知識に関するクラスとする。たとえば、下記の例では、スキーのドメイン知識としてパラレルという滑り方があることや、話者がそれを知っているということによって、スキーに詳しいといったことが言外の情報として伝わっている。

- (例 13) A:「スキーはパラレルできますか?」→「A はスキーに詳しい」

客観的事実に関するクラス

言外の情報が、客観的な事実と照らし合わせて推論可能な場合には、客観的事実に関するクラスとする。下記の例では、「(今が) 残暑である」という発話内容から推論することで「今が夏の終わりである」という言外の情報が得られる。

- (例 14) A:「残暑がひどくて」→「今は晩夏である」

表 2: 言外の情報に対するアノテーションの結果

クラス	定義	頻度	割合	
対話理論	発話行為	言外の情報が発話によって引き起される行為の内容を表す.	4	4%
	協調の原理	言外の情報が「会話の含み」によって聞き手に伝わる情報を表す.	11	11%
	プラン認識	言外の情報が発話の背後にあるプランを表す.	21	21%
	談話構造	言外の情報が対話の構造 (談話構造) を変化させる.	17	17%
	共通基盤	言外の情報が共通基盤の形成に関わる.	9	9%
その他	人物属性	話し方や話の内容から現れる言外の情報が話者の人物属性を表す.	17	17%
	因果関係	言外の情報が常識的な因果関係によって推論できる.	12	12%
	ドメイン知識	言外の情報が特定のドメイン知識に関するものである.	5	5%
	客観的事実	言外の情報が客観的な事実と照らし合わせて推論可能である.	4	4%
合計		100	100%	

4.3 言外の情報に対するアノテーション

収集された言外の情報について、言外の情報のクラスのアノテーションを行った。アノテーションの際には、発話内容と言外の情報、および、それらの間の関係に基づいて、最も適切なクラスを付与するものとした。

アノテーション対象として、収集した言外の情報からランダムに 100 個の言外の情報を抽出した。アノテーションは著者 1 名が行った。そのアノテーション結果を表 2 に示す。表 2 には、対話理論に基づく 5 つのクラスと、その他の 4 つのクラスが示されており、各クラスに該当する言外の情報の定義が示されている。表 2 の 4 列目および 5 列目には、ランダムに抽出した 100 個の言外の情報のうち、各クラスに該当する言外の情報が出現した頻度と割合が示されている。

表から、今回の結果ではプラン認識に関するクラスが最も多く出現しており、今後言外の情報を推定するためには、プラン認識が重要であることがわかる。談話構造に関するクラスの頻度が高いのは、雑談において様々な話題が扱われるためだと考えられる。人物属性のクラスの頻度が高く、客観的事実のクラスの頻度が低いのは、雑談では主に話者に関する内容が話題となるためだと考えられる。発話行為のクラスの頻度が低いのは、雑談対話が特定の目的を持たないためだと考えられる。

5. おわりに

本稿では、雑談対話を収録したテキストチャットデータにおける各発話に対して、複数の作業者に言外の情報を記述してもらい、収集されたデータを利用して、言外の情報の類型化を行った。その結果、言外の情報の類型化には対話理論の知見が有効活用できる一方で、対話理論以外の知見も必要となることがわかった。

今後の課題としては、収集した全ての言外の情報について、アノテーションを行いたいと考えている。今回、その他に属するクラスは、事例に基づいて作成されたため、対話理論との整合性を考え、より適切なクラスを検討する必要がある。また、数人でアノテーションを行うことにより、類型化の精度を高めたい。さらに、言外の情報を理解することのできる対話システムに向けて、言外の情報を推定するための手法を確立していきたい。具体的には、言外の情報の候補の自動生成やそのランキング、また、動的に推定する対象を選定する必要がある。その後は、推定のための知識の獲得などにも取り組まなければならない。また、言外の情報を推定するシステムを実装し、その有効性を検証していきたい。

参考文献

[Austin 62] Austin, J. L.: *How to do things with words*, Oxford University Press (1962), (坂本訳『言語と行為』大修館書店, 1978)

[Grice 75] Grice, H. P.: Logic and conversation, in Cole, P. and Morgan, J. eds., *Syntax and semantics*, Vol. 3: Speech acts, pp. 41–58, Academic Press, New York (1975)

[Higashinaka 14] Higashinaka, R., Imamura, K., Meguro, T., Miyazaki, C., Kobayashi, N., Sugiyama, H., Hirano, T., Makino, T., and Matsuo, Y.: Towards an open domain conversational system fully based on natural language processing, in *Proceedings of the 25th International Conference on Computational Linguistics*, pp. 928–939 (2014)

[Higashinaka 15] Higashinaka, R., Mizukami, M., Funakoshi, K., Araki, M., Tsukahara, H., and Kobayashi, Y.: Fatal or not? Finding errors that lead to dialogue breakdowns in chat-oriented dialogue systems, in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 2243–2248 (2015)

[Litman 87] Litman, D. J. and Allen, J. F.: A Plan Recognition Model for Subdialogues in Conversations, *Cognitive Science*, Vol. 11, pp. 163–200 (1987)

[Ritter 11] Ritter, A., Cherry, C., and Dolan, W. B.: Data-Driven Response Generation in Social Media, in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pp. 583–593 (2011)

[Takabatake 15] Takabatake, Y., Morita, H., Kawahara, D., Kurohashi, S., Higashinaka, R., and Matsuo, Y.: Classification and Acquisition of Contradictory Event Pairs using Crowdsourcing, in *Proceedings of the The 3rd Workshop on EVENTS: Definition, Detection, Coreference, and Representation*, pp. 99–107 (2015)

[Zhang 09] Zhang, C. and Chai, J. Y.: What do We Know about Conversation Participants: Experiments on Conversation Entailment, in *Proceedings of the 10th Annual Meeting of the Special Interest Group in Discourse and Dialogue*, pp. 206–215 (2009)

[Zhang 10] Zhang, C. and Chai, J. Y.: Towards Conversation Entailment: An Empirical Investigation, in *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pp. 756–766 (2010)

[Zhang 14] Zhang, M., Qin, B., Liu, T., and Zheng, M.: Triple based Background Knowledge Ranking for Document Enrichment, in *Proceedings of the 25th International Conference on Computational Linguistics*, pp. 917–927 (2014)

[Zhang 15a] Zhang, M., Feng, V. W., Qin, B., Hirst, G., Liu, T., and Huang, J.: Encoding World Knowledge in the Evaluation of Local Coherence, in *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (2015)

[Zhang 15b] Zhang, M., Qin, B., Zheng, M., Hirst, G., and Liu, T.: Encoding Distributional Semantics into Triple-Based Knowledge Ranking for Document Enrichment, in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, pp. 524–533 (2015)

[小磯 15] 小磯 花絵, 石本 祐一, 菊池 英明, 坊屋 真弓, 坂井田 瑠衣, 渡部 涼子, 田中 弥生, 伝 康晴: 大規模日常会話コーパスの構築に向けた取り組み—会話収録法を中心に—, 第 74 回言語・音声理解と対話処理研究会, 第 5 巻, pp. 37–42 (2015)