

会話における省略とその補完

渡邊大 榊井文人 河合敦夫 椎野努

三重大学 大学院 工学研究科

〒514-8507 津市上浜町 1515

E-mail : {watamasa, masui, kawai, shiino}@shiino.info.mie-u.ac.jp

あらし 人間どうしの会話の場合、適切な情報を迅速に相手に伝えようとするため、共通の知識があれば、それらを省略して発話する。機械を相手にした場合も、人間は無意識のうちに省略を行なうものであり、またそれが許されないと煩わしく感じるものである。機械側は人間が省略を伴って発話した時に、その省略を補完して理解しなければならない。本稿では、質問と応答を含む2対話対における省略補完を取り上げる。まず、どの部分で省略が起こる可能性があるのか分類を行なう。次に、その分類に基づいて、補完ルールや、関連述語辞書を用いて省略を補完する方法を述べる。最後に今後の課題と展望をまとめる。

キーワード 会話 省略 補完 対話対

Conversational ellipsis and its complement

Masaru Watanabe Fumito Masui Atsuo Kawai Tsutomu Shiino

Faculty of Engineering, Mie University

1515 Kamihama-cho, Tsu City, Mie 514-8507 Japan

E-mail : {watamasa, masui, kawai, shiino}@shiino.info.mie-u.ac.jp

Abstract In conversation we omit common knowledge and utter only the suitable information in order to communicate rapidly. We omit unconsciously, even when conversing with machines. Also, we feel troublesome if ellipsis isn't admitted. When people utter with ellipsis, machines have to complement the ellipsis to understand its context. In this paper, we propose the method to complement the ellipsis in two sets of utterance pairs of a question and answer. First, we classify the two sets of utterance pairs into 5 types, depending on whether ellipsis occurs or not. Next, we explain the method of the complement by the use of the rules and the dictionary associated with predicate.

key words conversation ellipsis complement utterance pair

1 はじめに

本稿では、人間と機械の会話における、省略文の認識とその補完方法について述べる。

これまで、省略に関する研究はさまざまな観点からなされている。従来は論文や日本語テキストといった「書き言葉」における省略が主に取り扱われてきた。しかし、音声処理技術の向上に伴い、「話し言葉」における省略も研究対象に挙げられることが多くなってきている。藤崎 [1] は、対話における省略を補完処理の観点から次のパターンに分類している。

対話当事者に関する省略 (a) 「(主語が) (対象を) お送り致します。」のように、叙述表現や尊敬語、謙譲語の使用に伴う省略。

(b) 「思います」の主語は「私」であるように、特定の動詞の特性による省略。

(c) ある動詞を「お願いします。」で置き換える、代用表現による省略。

文脈省略 補完されるべきものが、対話中の別の箇所で言及されている省略。

共有知識による省略 (a) 「明日、(目的地に) 行く?」「うん、(目的地に) 行く。」といったような、対話当事者間のみには分からない共有知識に基づく省略。

(b) 「羽田から (飛行機で) 発ちます。」のように、一般的な常識から容易に推測できる事項の省略。

山本 [2] は決定木学習による外界省略の人称と数の補完や、文脈省略の認知について述べている。また田村 [3] では、センター理論によって省略補完を行なう手法が紹介されている。これらは、主語や目的語といった格要素の補完について述べている。しかし、格要素の省略のほかに、述語の省略といった文末が不完全な発話も日本語会話では多く見られる。しかも、主語や目的語の補完には述語の持つ意味情報が大きく影響する。そのため、まず述語の補完を行い、それを基に格要素を補完していく必要がある。

そこで本稿では、述語も含めた省略について、その補完方法を考える。構成は、2章においてここで扱う対象を明確にする。次に3章では、認識した表層意図から対話対を考慮した、会話の流れの中での省略の起こる位置を分類する。4章では本会話シス

テムの構成と、3章で行なった分類に基づいて省略を補完していく手法について述べる。最後に5章で今後の課題と展望を述べる。

2 補完対象

ここでは、1章での省略のパターン分けのうち、本稿でどの省略を補完対象として取り扱うのか明確にする。

● 文脈省略

人間どうしが会話を行なったとき、先に挙げた省略のうち、発生頻度が高いと考えられる省略は、ゼロ代名詞などの省略と、文脈省略である。機械と人間が会話することを想定した場合も、同じ事が言えるのではないかと考える。そのうち、ゼロ代名詞の省略補完はかなりの高精度な補完結果が報告されている (中岩 [4])。しかし、文脈省略の補完結果は、ゼロ代名詞の補完率ほどは高くない (田村 [3])。そこで、本稿では文脈省略の補完を試みた。

● 質問と応答の対話対

対話では、どちらかの質問に対して他方が答えるという形で繰り返されることが多いが、途中で関連する別の会話が入れ子になる場合もある。そこで、本稿では、対話の中で出現度の高い質問文と応答文に補完対象を絞る。

● 不完全な述部

質問文や、応答文では文末表現が不完全な文¹も多く見受けられる。文末からは、省略された格要素の補完や、意図把握のために有益な情報が得られる。そこで、本稿では述部の省略に着目し、その補完方法を考える。

● 質問から予測できない応答

4.1で述べる予測辞書は、システムの質問に対してユーザ側の返答を予測する機能である。ユーザの応答文が、その予測と一致すれば、ユーザの意図は把握できるので、補完対象としない。例えば、次の会話例 (a) ~ (c) である。逆に (d)、(f) のように予測辞書と一致しないか、

¹文末が名詞や、接続助詞、格助詞、副助詞などで打ち切られている文などを指す。

(e) のように一致しても他に述部が存在する場合は補充対象として取り扱うこととする。

S² : A病院に行きますか?

< 「はい」、「いいえ」、「行く」を予測 >

U (a) : はい。(肯定、予測通り)

(b) : 行きます。(肯定、予測通り)

(c) : 行かない。(否定、予測通り)

(d) : 明日の午前中。(肯定、予測外)

(e) : 忙しいから行かない。
(否定、予測に部分一致)

(f) : 忙しい。(否定、予測外)

次に、会話の流れにおいて、どこで省略が生じるのかを検討した。

3 対話対の分類と省略の可能性

人間どうしの対話を理解するために、質問と応答の対話対を認識しようとする研究は高野 [5] によって行なわれている。人間対計算機の対話において、計算機側の表層意図は自明である。よって、ユーザ発話の表層意図の理解が問題である。ここでは、対話対の認識についてはシステム側の表層意図からの予測とユーザ発話の表層的特长から、内容に深く関与せず認識する。対話はどちらかの質問に対して他方が答えるという形が繰り返される場合が多い。そこで簡単のため対話参加者をシステムとユーザの2者、かつ2対話対までと限定して、単純な質問や依頼と応答という対話対や、入れ子状態になった対話を含む対話対について分類を行い、その中でユーザ発話のどの部分で省略が起こりうるか考察を行なった。

対話対の分類と省略の可能性を分析したものを図1に示す。

《分類1》 システムの質問に対して、ユーザが応答を続けるというシステム主導型の流れの例である。発話 S1・S2 で得られた情報が省略される可能性があるため、発話 U1・U2 で省略が起こりうる可能性がある。

< 会話例 >

²S は System を、U は User を表し、右側の数字が同じものは対話対であることを示す。

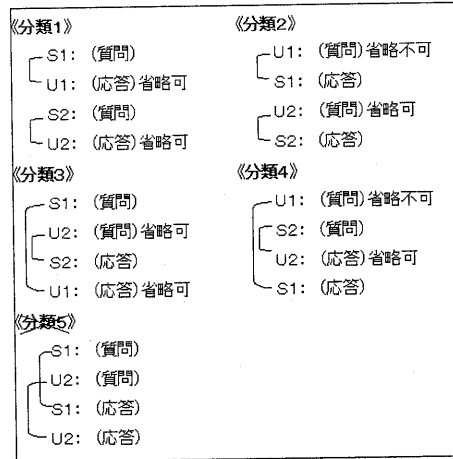


図 1: 対話対の分類と省略可能性

S1 : 病院に行きますか?

U1 : 明日の午前中 (病院に)³ (行きます)。

【ルール3, 4, 5により補充】⁴

S2 : どの病院に行きますか?

U2 : A病院。 【予測辞書と一致】

《分類2》 ユーザ側からの質問や依頼に対して、システムがそれに応答していくという、ユーザ主導型の流れの例である。ある話題がユーザにより始められたり、話題が変更された時は、通常、ユーザ側の発話に省略はない。したがって、省略が起こりうる可能性があるのは発話 U2 のみとする。

< 会話例 >

U1 : 明日、A病院はあいていますか?

S1 : 休みです

U2 : じゃあ、(明日) B病院は (あいていますか)? 【ルール1により補充】

S2 : 明日は10時から17時まであいています。

《分類3》 システムの質問に対して、ユーザが別の質問を返してきた、いわゆる「入れ子」状態の流れの例である。<分類1>同様、発話 S1・S2 を受け発話 U1・U2 で省略が起こる可能性がある。

³ () 内の情報を補充する。

⁴ [] 内に補充方法を述べる。詳しくは4.2節を参照。

<会話例>

- S 1 : いつA病院に行きますか?
U 2 : (A病院は) 土曜日は診察していますか?
【ルール5により補完】
S 2 : 休みです。
U 1 : じゃあ、月曜日に (A病院に) (行きます)。
【ルール2により補完】

《分類4》 ユーザ側からの質問や依頼で会話が始まっているか、ユーザの意図により話題が変換されたパターンである。この場合、最初のユーザ発話 U 1 において、応答に必要な情報が明確に示されていないか曖昧なために、情報不足となり、システム側がその不足部分を聞き返さなければならない。このパターンで省略の起こりうる可能性があるのは発話 U 2 のみとなる。

<会話例>

- U 1 : A病院へはバスでどれくらいかかりますか?
S 2 : 時間ですか、料金ですか?
U 2 : 料金です。 【予測辞書と一致】
S 1 : 津駅から 350 円です。

また、<分類5>のように自分の質問に対して自分で応答することはないので、対話対が交差することはありえない。

ここでは、実際の会話で起こりうる<分類1>から<分類4>までの会話パターンを取り扱うこととする。

4 システム構成と補完手順

ここでは、3章の分類をもとに、省略部分を補完するためのシステムと補完手順について述べる。

4.1 システム構成

補完処理も組み込まれたシステムの構成を図2に示す。

各部の概略は次のようなものである。

- **会話文理解** 音声認識部から出力された候補文から、予測辞書、意味解析により入力文の解析を行い、フレームに格納する。状況メモリ、ユー

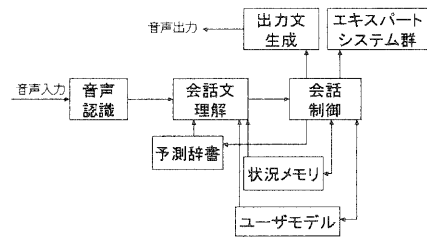


図 2: システムの概略図

ザモデルと相互的に情報のやり取りを行いながら、省略補完はここで行なわれる。

- **予測辞書** 入力文の認識率を高めるため、会話において次に入力される可能性の高い単語の集合を用意し、解析に際して優先的に参照する。その集合を予測辞書という。
- **会話制御** 会話文理解部で把握されたユーザの発話内容に対し、意図理解を行い、また、その内容を当該する分野のエキスパートシステム群に渡す。また、これらの部分が処理した出力を受け取り、応答文生成部に渡す。
- **エキスパートシステム (ES) 群** 分野別に用意され、その分野の話題が進む間は ES モジュールが対応する。ユーザの発話内容により、推論、判断、指示、質問など、次にシステムから発話する内容を生成したり、指示する画像、図表の選択を行なう。
- **ユーザモデル** 個別のユーザについての固定的な情報 (年齢、住所、病歴 etc...) や、過去の会話によって得られた動的情報を保持する。
- **状況メモリ** 会話によって現在及び、それまでに得られた情報をフレームの形で保持するもの。次節で示す補完ルールや関連述語辞書と照合を行ない、この状況メモリの内容と一致したものが補完される。

4.2 補完手順

ここでは、どのように省略を認知し、補完しているのかについて述べる。

省略がなされているかどうかの判断については、会話文処理部で予測辞書と一致が無かった場合とする。

補完手順のフローチャートを次の図3に示す。

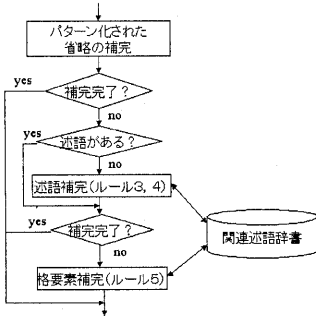


図3: 補完手順

まず、ユーザ発話文の中には形式が類似した省略発話が見られる。それらをパターンと呼ぶ。パターン化された省略に対しては、補完ルールにより補完を行なう。パターン化できる発話として、《分類2》のU2や、《分類3》のU1が挙げられる。《分類2》のU2は先の質問(U1)のある格要素だけを同じ属性の名詞で置き換えて発話したものである。このときは次のような補完ルールが働く。

<ルール1>

If 前のユーザ発話(U1)のある格要素だけを同じ属性の名詞で入れ替えた疑問文
Then 前のユーザ発話(U1)の内、条件部で一致した格要素以外の述語、格要素を補完する

《分類3》のU1は入れ子会話が続いてから、応答を受け取っていない最初の質問(S1)に対する答えを発話した場合である。このときは次のようなルールが働く。

<ルール2>

If 未解決のシステム発話(S1)が残っている
And ユーザ発話がその質問文に対する応答文になりうる
Then その質問文の述語、格要素を補完する

パターン化できない省略に対しては、関連述語辞書を用いる。関連述語辞書は、取り得る格データ⁵が類似している述語ほど関連性が深いとしてグループ化を行い、知識として構築したものである。研究目的である「高齢者に対する医療診断、交通案内」に即した会話を収録し、作成した。

関連述語辞書により補完を行なうため、次のルールを用いる。

<ルール3>

If ユーザ発話の述語が省略
Then ユーザ発話中の格データが関連述語辞書の格データと一致する全ての述語を取り出す
And ルール4を発火する

<ルール4>

If 取り出した述語が直前のシステム発話中に存在する
Then その述語をユーザ発話の述語とする
つぎに、格要素補完には次のルールを用いる。

<ルール5>

If 元からあった、または補完された述語に対し関連述語辞書を参照し、該当する格データを抽出する
And 抽出した格データのうち、直前の発話で用いられている名詞属性がある
Then それを補完する

たとえば、《分類1》のU1はまず、「明日の午前中」から『時間格』 = {【未来】}の格データを得る。ルール3において、関連述語辞書内の『時間格』 = {【未来】}という格データを参照して、「行く」、「診断する」、「予約する」、「出る」といった述語を取り出す。そして、ルール4が発火され、直前のシステム発話に用いられていた述語「行く」が補完される。次に、ルール5に移り、関連述語辞書の「行く」から格データとして『場所格』 = {【場所】、【組織】}、『動作始点格』 = {【場所】、【組織】}、『動作終点格』 = {【場所】、【組織】}、『時間格』 = {【未来】、【日付】}、『乗り物格』 = {【乗り物】}を抽出する。それらの名詞属性のうち直前のシステム発話で用いられている「A病院」が『動作終点格』に補完される。

⁵ 『格要素』と『名詞属性』の集合が対になったもの。

《分類3》のU2の『行為者格』である「A病院」も、ルール5により補完される。述語「診察する」から、格データ『行為者格』= {【病院】}、『時間格』= {【未来】}が抽出される。そのうち、直前の発話で用いられていた名詞属性【病院】である「A病院」が『行為者格』に補完される。

5 まとめ

本稿では、会話中に含まれる文脈省略を補完するための手法を提案した。まずは2対話対に限定して、会話の分類を行い、ユーザ発話のどの部分で省略が起こる可能性があるのか考察した。次に、ルールと関連述語辞書を用いて補完を行なう方法を述べた。今後の課題として、次の2つが挙げられる。

- 深い推論を要する省略補完 2章の会話例(f)のように「病院に行きますか?」という質問に対して「忙しい(から、行かない。)」と補完するのが困難となっている。
- 会話パターンの分類とルールの抽象化 今後、話題が増えていくにしたがって、適切なパターン分けと、ルールの増加を防ぐための抽象化が不可欠になってくる。

参考文献

- [1] 堂下修司, 新美康永, 白井克彦, 田中穂積, 溝口理一郎共編: 音声による人間と機械の対話 5編2章, オーム社出版局
- [2] 山本和英, 隅田英一郎: 決定木学習による日本語対話文の格要素省略補完, 情報処理学会, Vol.6, No.1, pp.3-28(1999)
- [3] 田村浩二, 奥村学: センター理論による日本語談話の省略解析, 自然言語処理 107-12, pp.91-96(1995)
- [4] 中岩浩巳: 日英対訳コーパス中のゼロ代名詞とその指示対象の自動認識, 情報処理学会研究報告, NL-123-5 (1998)
- [5] 高野敦子, 柏岡秀紀, 平井誠, 北橋忠宏: 発話対と認識手法について ~質問とその応答~, 自然言語処理 94-3, pp.17-24 (1993)
- [6] 望主雅子, 酒寄哲也, 小島裕一, 佐藤奈穂子: ナビゲーション対話における省略文の分析, 音声言語情報処理 13-5, pp.25-30 (1996)
- [7] 巖寺俊哲, 石崎雅人, 森元暹: 表層パターンを用いた対話構造の認識, 情報処理学会論文誌, Vol.39 No.8, pp.2452-2465 (1998)
- [8] 西澤信一郎, 中川裕志: 日本語の自由会話における談話構造の推定~因果関係の場合について~, 自然言語処理 113-14, pp.89-94 (1996)
- [9] 鳥居肖史: 曖昧な入力を可能とする自然言語インターフェース, 自然言語処理 100-7, pp.49-56 (1994)
- [10] 梶井孝洋, 河合敦夫, 椎野努他: 高齢者のための知的ヒューマンインターフェースシステム, 電気関連学会東海連合大会, pp.395 (1999)
- [11] 渡邊大, 梶井文人, 河合敦夫, 椎野努: 会話システムにおける省略とその補完, 電気関連学会東海連合大会, pp.388 (2000)