

## 課題解決のための対話における照応・省略の先行詞の認定

渡辺祐樹 吉谷祐作 渥美清隆\* 増山繁

{watanabe,yositani,atsumi}@smlab.tutkie.tut.ac.jp, masuyama@tutkie.tut.ac.jp

豊橋技術科学大学 知識情報工学系

### 1 はじめに

自然言語における照応・省略の先行詞を認定することは、計算機が人間と対話を行う際に解決しなければならない問題である。本研究では、計算機と人間が簡単な課題解決を行うための対話を対象にして、照応・省略の先行詞認定問題を解決することを目的とする。ここで簡単な課題解決のための対話とは、命題を一つずつ質問して答えを得ることで課題を解決していく対話とする。

この分野の研究として、センタリング理論 [1] がある。これは、前の文の焦点が次の文の主題となり省略されるという原理を用いて、照応・省略を解決するものである。この方法は、主題・焦点情報をどうやって特定するかという枠組みは述べていないが、極めて汎用的な照応・省略の先行詞の認定方法である。

これに対して、本研究では、簡単な課題解決のための対話での次のような特徴を利用している。

- 回答文の述語がそれと対になる質問文の述語と同じ場合、回答文の助詞とそれにかかる名詞句が、質問文のそれと同じなら省略される。

センタリング理論が、一つの文について一つの先行詞だけしか認定できないのに対し、本研究では一つの文における複数の照応・省略に対応できる。

この特徴について課題解決のための対話である DiaLeague 中の対話を用いて検証した。その結果、この手法がある程度有効であることが確認できた。

### 2 諸定義

#### 2.1 対象となる対話

本研究では、橋田らによって提案された DiaLeague コンテスト [2] における対話を対象とする。ここで行なわれる対話は、対話を行う人間とコンピュータの両者に課題が示され、その課題を両者が協調して達成するために行なわれる、日本語文による情報交換である。課題は、

\* 日本学術振興会特別研究員 (DC)

両者に異なる路線図を示し、その共通経路を発見することである。

人間と対話システムには、それぞれ、図1のような架空の路線図が示される。それぞれの地図には、名前が未知の駅と、一方の地図にしか存在しない路線がある。駅名が '?1' や '?2' などの駅が、名前が未知の駅である。自分の地図で未知の駅名は、相手の地図では既知である。一方の地図にしか存在しない路線は、利用できない路線である。例えば、図1の先手の地図で、新発田と名古屋はつながっているが、後手の地図ではつながっていないので、この路線は利用できない。これらの情報を互いに対話を通して交換し、自分と相手の地図の間で共通かつ唯一な出発駅から目的駅までの経路を求めることが課題である。

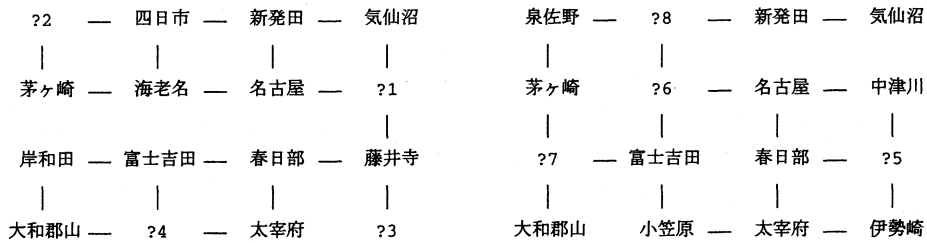
#### 2.2 発話の表現

本研究では、日本語で受話した文は全てフレーム [3] に変換した後、受話の内容に応じた処理を行なうものとする。また、発話の時には、発話内容をフレームで作成した後、対応した日本語の文に変換する。フレームは述語、属性、属性値で表現する。また、属性値にはデフォルト値が用意できる。

フレームは以下のように定義する。これらのフレームで表現できない文は、本研究では扱わない。

```
(行く (始点 [場所]='現在駅')
  ((終点 [場所])|
  (方向 [方向])(数 [数]='1')|
  (経路 [経路])|
  (經由 [場所]))
  (態 [態]))
(だ (主体 [駅])(客体 [駅])(態 [態]))
(だ (主体 [経路])(客体 [経路])(態 [態]))
(戻る (終点 [場所])(態 [態]))
(分かる (対象 [事柄]||[場所])(態 [態]))
```

'|' で区切られて括弧でくくられている複数の属性は、その内のいくつかの属性に値が入っていれば、その他の



先手の地図

後手の地図

図 1: ダイアリーグで実際に使用された地図

属性は空でも省略とみなさないとする。各属性値には「|」と「?」でくくられている意味素性により、属性のリスト、または、次の属性値が入る。

[場所]: 「豊橋」, 「二川」などの駅名

[経路]: 「豊橋, 二川, ...」のような具体的な経路

[方向]: 「東」 「北」などの方向

[数]: 1, 2, 3, ...

[事柄]: フレーム

[態]: 肯定, 否定, 不定のどれか

場所および、経路の属性値では属性リストとなる場合がある。その場合は次のように表現する。ここで、現在駅は探索している経路の末尾にある駅名である。

(場所 (基準 [場所]=「現在駅」)(方向 [方向])(数 [数]=「1」))  
 (経路 (始点 [場所]=「出発地」)(終点 [場所]=「目的地」))  
 特別なフレームとして、「はい」「いいえ」は次のように表現する

「はい」 (不定 (態「肯定」))

「いいえ」 (不定 (態「否定」))

ここで、不定の述語は任意の動詞と一致する。また、「何」「どこ」なども属性値が不定となることで表現される。

### 2.3 照応と省略

本研究における照応・省略の定義を示す。

- 省略: フレームの属性が空の属性値であること、または、「|」で区切られている属性がすべて空の属性値であること

- 照応: フレームの属性値が指示詞であること

本研究で取り扱う指示詞は「ここ」「そこ」「その」「それ」「これ」である。また、「これで」「それから」「それなら」「それで」などは接続詞として扱い、指示詞から除外した。

### 3 照応・省略の先行詞の認定方法

本論文で提案する先行詞の認定方法を 3.1 節で述べる。まず、その方法で先行詞の認定を試みる。失敗した場合には、意味素性による一致を取る方法を適用する。3.4 節で先行詞認定の全体の手続きを説明する。

#### 3.1 直前の相手の質問文と発話文との一致を取る

ここで、質問文とは、その発話文の態の属性値が不定になっているものである。直前の相手の質問文のフレームと、発話文のフレームの述語が一致する場合、その二つの属性の一致を取ることで照応・省略を解決する。ただし、相手の直前の質問文がすでに別の文との一致が取れている場合は、相手の直前の質問文よりも後になる相手の直前の発話文との一致を取る。ここで、発話文と一致を取る質問文は、既に照応・省略の先行詞が認定されているものとする。

属性同士の一致の取りかたを示す。各属性について、1~2を行う。すべての属性で一致がとれたら、成功。どれかの属性で一致が取れなかった場合、失敗である。

1. 質問文のある属性の属性値が不定の場合、回答文の同じ属性の属性値を埋める

2. 回答文のある属性の属性値が空の場合、質問文の同じ属性の属性値を埋める
3. 回答文のある属性の属性値に指示詞が入っている場合、質問文の同じ属性の属性値に、その指示詞の指すことのできる単語または属性リストが入っているなら、その値を埋める

ここで、指示詞と一致を取れる単語または属性リストを示す。

- ここ、そこ：「豊橋」など駅名、場所の属性リスト
- それ、これ：フレーム、経路の属性リスト

例 1 で示すと、

B2:(行く (終点 そこ)(態 ‘肯定’))  
 B2の直前の相手の質問:(行く (始点 ‘豊橋’)(終点 ‘二川’)(態 ‘肯定’))  
 それぞれの属性の一致を取ると、  
 B2:(行く (始点 ‘豊橋’)(終点 ‘二川’)(態 ‘肯定’))

また、一方のフレームが「はい」「いいえ」のフレームの場合、もう一方のフレームの述語も一致をとる。例を示す。

A: 豊橋から二川へ行けますか  
 (行く (始点 ‘豊橋’)(終点 ‘二川’)(態 ‘不定’))  
 B: はい  
 (不定 (態 ‘肯定’))  
 一致をとると、  
 B:(行く (始点 ‘豊橋’)(終点 ‘二川’)(態 ‘肯定’))

### 3.2 指示詞の先行詞の認定

一致が取れなかった場合の、指示詞ごとの先行詞認定法を示す。

- そこ：それまでの対話で一番最後に出現した場所の属性リストを先行詞とする
- その：「その」を「そのの」と置きかえて先行詞を認定する
- それ、これ：直前の相手の発話文に経路の属性リストが含まれていれば、その経路の属性リスト、そうでなければ、直前の相手の発話文のフレームを先行詞とする

- ここ：つねに現在駅を指示対象とする

次のような場合、

A: 木更津の1駅下は何ですか  
 B: そこはわかりません

直前の文には「木更津」と「木更津の1駅下」という候補があるが、このような場合、一番大きい単位で取る。つまり「木更津の1駅下」を先行詞として認定する。

### 3.3 省略の先行詞の認定

指示詞の照応の先行詞を認定した後、まだ空いている属性があれば、空いている属性値に付されている意味素性に一致する単語を直前の文から取り出す。ここで、述語「分かる」の対象の属性値に入ることのできる値の意味素性が、「文」「場所」で2種類ある。この場合は、前の文に場所の名詞句が含まれていれば、その中で一番最後に出現したもの、含まれていなければ、直前の文の照応・省略を補完したものを先行詞とする。

### 3.4 照応・省略の先行詞の認定手続き

以下で、発話文中の照応・省略の先行詞を認定する方法を示す。

1. 指示詞「その」を「そのの」に変換。
2. フレームに変換。
3. 肯定文、否定文であれば、直前の相手の質問文とその発話文との一致を取る(3.1節参照)。取れれば5を行い終了。
4. 一致がとれなければ、指示詞の先行詞を認定する(3.2節参照)。
5. デフォルト値が付されている省略された属性があればデフォルト値を入れる
6. まだ省略されている属性があれば、前文から先行詞を認定する(3.3節参照)。

## 4 実験

対話ログ(DiaLeague'97第2回対話コンテスト 846対話)を手作業でチェックし、この方法でどの程度、照応・省略が解決できたかを調べる。

省略については16対話を調査して、次の値を求めた。

$$\text{再現率} = \frac{\text{正しく省略を補完した数}}{\text{すべての省略の数}}$$

$$\text{適合率} = \frac{\text{正しく省略を補完した数}}{\text{省略を補完した数}}$$

省略の先行詞認定については再現率は  $302/327=92\%$ 、適合率は  $302/346=87\%$  となった。省略を補完した手続きの内訳は、次のようになった。

手法	適用回数	誤認識数
デフォルト値	215	29
前文との一致	164	13
前文から先行詞を取得	23	12

指示詞の照応については、全対話を調査して、適合率は  $61\%$  となった。指示詞ごとの適合率は、「そこ」の先行詞の認定が  $5/6=83\%$ 、「それ」の先行詞の認定が  $4/8=50\%$ 、「その」の先行詞の認定が  $5/6=83\%$ 、「これ」の先行詞の認定が  $0/1=0\%$ 、「ここ」の先行詞の認定が  $2/3=66\%$  だった。ここで、照応は指示詞が出現したさいに必ず検出されるとしたので、再現率は適合率と等しくなる。また、指示詞の数が少ないのは、指示詞が使用された文で、フレームに変換可能な述語が使用されていたものが少なかったからである。

## 5 考察

省略の先行詞認定が成功した例の大部分が「豊橋へ行けますか」のような、「行く」のフレームの始点が省略されたものに、デフォルト値をあてはめたものであった。また、前文との一致による省略の補完も高い認識の成功率を出すことができた。省略の先行詞の認定に関しては高い再現率、適合率を出すことができた。

「分かる」の対象の属性値が省略された場合の先行詞の認定が失敗することが多かった。

A: 小笠原から左に行けますか

B: いいえ

A: 分からなくなりました

この場合、本手法では、「分かる」の対象の属性値に先行詞として、「小笠原から左に行けません」というフレームを入れる。しかし、この場合2回目のAの発話文は、「どうしていいか分からなくなった」という意味で使われていて、不適当な先行詞を認定したことになる。

A: 木更津から藤井寺に行けますか

B: わかりません

このような対話で本手法では、Bを「藤井寺が分かりません」と解釈するが、正解となる解釈は「どう答えたらいいか分からない」なのか、それとも「木更津が分からない」なのか、はログだけでは判別不可能であった。このような場合、Bの地図で藤井寺の場所が対話の始めの時点で未知だったなら、本手法の解釈で正解とした。

デフォルト値を入れることで誤って省略を補完してしまった例があった。

A: 下へ藤井寺まで行けますか

この文にデフォルト値を入れると「現在駅から1駅下へ藤井寺まで行けますか」という解釈になるが、藤井寺が現在駅の1駅下であるとは限らない。そこで、このような補完は間違いであるとした。この場合は、省略でない場合に余計な要素を補完してしまうことになる。

指示詞の先行詞の認定については指示詞「そこ」「その」は、すべて一番最近に現れた場所を先行詞とすることでほとんど正しく認定できた。しかし、「これ」「それ」は、場所でなく事柄を先行詞として認定することもあり、この場合に認定が失敗している。これは、前文のフレームを先行詞とするべき場合に、前文に場所の名詞句が含まれていることにより失敗している。前文に、複数の指示対象があった場合に、どの指示対象を選ぶかを調べる方法を考察する必要がある。

## 6 むすび

DiaLeagueの省略の先行詞の認定に関しては、高い認定の再現率、適合率を得ることができた。他の簡単な課題についての対話でも、この手法を用いれば高い認定の成功率を得られると推測している。

## 参考文献

- [1] M. kameyama: "A Property-Sharing Constraint in Centering", 24th Annual Meeting of the Association for Computational Linguistics, pp.200-206 (1986).
- [2] 橋田, 伝, 他: "DiaLeague — 自然言語処理システムの総合評価 —", 人工知能学会誌, vol. 12, no. 3 (1997).
- [3] 長尾 編: 自然言語処理, 岩波講座ソフトウェア科学, 15, 岩波書店 (1996).