

プランニングの知識を用いた照応・省略解決

志賀 聡子 徳永 健伸 田中 穂積

東京工業大学大学院 情報理工学研究所
〒152-8552 東京都目黒区大岡山 2-12-1

E-mail: {satoko,take,tanaka}@cl.cs.titech.ac.jp

これまで、話者の意図理解をするために、プランやゴールといった概念が用いられてきた。本稿では、この概念が対話中に現れる照応・省略表現の解決にも役立つことを示し、プランニングの知識を用いた新しい照応・省略解決手法を提案する。提案手法では、発話が指示する行動の効果と前提条件の連鎖から、同一のゴールを目指している複数の発話間の関連性を捉えることによって照応・省略の解決を行う。これにより、仮想世界上のロボットを音声対話で操作するシステムにおいて、従来のような文の表層に現れる情報に基づく手法では解決できなかった、関連性のある行動間での非明示的な焦点の移動に追従することが可能になった。

Plan-based Anaphora and Ellipsis Resolution

SHIGA Satoko, TOKUNAGA Takenobu, and TANAKA Hozumi

Department of Computer Science
Graduate School of Information Science and Engineering
Tokyo Institute of Technology
Ō-okayama Meguro-ku Tokyo 152-8552 JAPAN

E-mail: {satoko,take,tanaka}@cl.cs.titech.ac.jp

The concepts of the plan and the goal have been used for intention understanding of a speaker. In this paper, we show that these concepts are also useful to resolve the anaphora and ellipsis in speech dialog, and propose a new method to resolve them by using a plan library. The proposed method consists of 3 steps: (1) identifies a corresponding plan behind each utterance. (2) makes a group of utterances sharing the same goal by using effect-precondition relations. (3) finds the antecedent within this utterance group. The method was evaluated with a small dialog corpus and showed its effectiveness.

1 はじめに

人間とコンピュータのインターフェースの手段として、最も望ましいのは自然言語を用いることである。しかし、コンピュータに自然言語を理解させるには多くの困難な問題があり、その中の一つに照応・省略解決がある。

本稿では、これまで話者の意図理解のために扱われ

てきたプランやゴールといった概念 [6] が、対話中に現れる照応・省略の解決にも役立つことを示し、プランニングの知識を用いた新しい照応・省略解決手法を提案する。

自然言語理解の研究の一つの枠組みとして、仮想世界内のロボットを自然言語で操作する対話システムがある。「積木の世界」で知られる Winograd の SHRDLU が代表的なシステムである [8]。また、近年の音声認識

技術の発展に伴い、新山は音声による指令を理解するシステム傀儡を開発した(図1)[7]。提案手法は、このような対話システムにおける照応・省略解決を想定している。



図1: 傀儡のスクリーンショット

2 関連研究

従来の照応・省略解決の研究は、書き言葉を対象としたものが多く、主に代名詞などが出現した際に、それと先行文脈中の名詞句と結び付けることを目的としている。GroszらのCentering理論[4]が代表的な手法であり、この理論をもとに多くの研究がなされている。たとえば、亀山はCentering理論を日本語の省略解決に適用した[5]。これらの手法は、文法情報など文の表層に現れる情報を解析して先行詞を同定している。

近年では、話し言葉を対象とした研究もみられるようになってきており、Byronらによって、複数話者間での対話にCentering理論を適用した研究がなされた[2]。しかし、話し言葉は書き言葉に比べて、省略が多いこと、断片的で整った文法構造をしていないこと[3]、また複数話者間でそれぞれ異なる焦点を持つ可能性があることなど問題点が多い。

一方、SHRDLUや傀儡など、ユーザの指令に対して行動する対話システムでは、ロボットの行動手続きを生成するために、「誰が何をどうする」という情報をすべて理解することが必要で、省略の検出が必須になる。また、ユーザは仮想世界の状況を示すグラフィックディスプレイを見ながら指示をするため、先行詞がそれまでの発話文中に現れるとは限らないという問題点がある。さらに、先行詞は世界内の実体と一意に結び付ける必要がある。たとえば「球」という表現に対し、世界内に球が複数あれば、どの球かを同定しなければならない。

そこで、傀儡における照応・省略解決アルゴリズムは、まず発話を格フレームに変換し、その後、照応・省略解決済みの履歴発話を参照して先行詞を同定する。このために、発話スレッドと呼ばれる履歴データベース

に発話を分類して保存しておき、先行詞の探索範囲を同一発話スレッド内の発話に制限する。このとき、発話の分類には、主語・述部の一致や手がかりとなる語といった表層に現れる特徴を用いる。そして、照応・省略表現を含む発話が入力された場合には、まず発話スレッドを同定し、発話スレッド内の履歴発話の格と、照応のある発話の格を比較することで、照応・省略を解決する。もし、参照すべき発話が見つからなければ、仮想世界内を探索する。

3 問題点

2節で述べたように、従来手法では文の表層に現れる情報を用いて照応・省略を解決している。しかし、実際には、照応詞と先行詞の対応関係を捉える手がかりが表層に現れないこともあり、それだけでは解決できない場合が存在する。

たとえば、次の対話例1を考える。

対話例1:

- #1-1 「ロボットAは赤い球を押して」
- #1-2 「青い球の前に行って」
- #1-3 「それを押して」

発話#1-3の目的格「それ」は「青い球」を指している。しかし、傀儡のアルゴリズムでは、表層上で述部「押す」が一致していることから、発話#1-1と発話#1-3を関連付けて、「それ」の先行詞を同じ目的格の「赤い球」としてしまふ。人間がこの照応を正しく捉えることができるのは、「発話#1-3で押すために、発話#1-2で球の前に移動した」という推論が背景にあるからである。

ただし、この対話例1の場合には、履歴発話の中で照応詞に最も近い名詞句を先行詞とするような単純な手法でも解決できる。だがここで次の対話例2を考えると、表層に現れている情報だけでは解決しにくいことが良くわかる。

対話例2:

- #2-1 「ロボットAは赤い球を持って」
- #2-2 「青い球の前に行って」
- #2-3 「それを置いて」

発話#2-3の目的格「それ」は発話#2-1中の「赤い球」を指しており、最も近い名詞句である発話#2-2中の「青い球」とは異なる。これは、「事前に持っていたからこそ置くことができる」という行動の連続性を考慮することによって解決できることであり、表層に現れる情報だけで発話#2-1と発話#2-3を関連付けることは難しい。

一般に、人間がロボットに指令を行う際、頭の中に大きなゴールがあり、図2のように予め人間自身がプランニングを行って、そのプランを分割して順次指示していくと考えられる。そして、その間で冗長な言い回しを避けるために照応・省略表現を使うので、このような照応・省略を解決するためには、システムもプランニングの知識に基づいた推論をする必要がある。

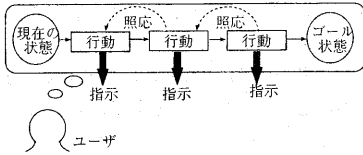


図2: ユーザのプランニングと指示

4 プランベースの照応・省略解決アルゴリズム

4.1 システムのアーキテクチャ

本稿における対話システムのアーキテクチャは図3のようにになっている。ユーザの一つの発話は一つの行動を指示しているものとし、各発話ごとに一連の処理が行われる。

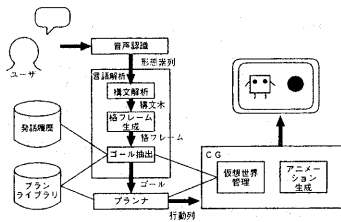


図3: 対話システムのアーキテクチャ

従来の傀儡と異なる点は、これまでプログラム中にハードコーディングされていたロボットの行動生成を行う、プランナがある点である。その際、プランニングに用いる知識がプランライブラリである。

システムは、ユーザの指令をプランナに渡すためのゴールに変換しなければならない。この言語解析部でも、プランナと共有のプランライブラリを用いる。ここで言うゴールとは、図2に示したようなプラン全体における大きなゴールではなく、部分プラン（プランオペレータ）である各行動が持っている小さなゴールである。

本稿の中心となるのはこのゴール抽出部であるので、この部分についてのみ述べる。ゴール抽出部は、さらに次の二段階にわかれている。

- プランオペレータへのマッピング
- 照応・省略解決

4.2項、4.3項でそれぞれについて説明する。

4.2 プランオペレータへのマッピング

まず、発話を格フレームに変換したものをもとに、対応するプランライブラリ中の行動を取り出す(図4)。今回は、ロボットのとれる行動が限定されているので、表層上の述部から一意に決定することができる。

取り出した行動には、必須格が引数として付与されている。これらの引数を発話の格フレームと対応させて埋めるが、代名詞であったり、発話の中に存在しない格は変数として残る。また、表層から世界内の実体と一意に対応付けられないものも変数となる。図4に示す例では、主格はロボットAに決定しているが、代名詞であった目的格と省略されていた場所格が変数として残っている。こうして変数のまま残された部分が、次に照応・省略を解決すべき箇所であるとわかる。

また、各行動にはそれぞれ、その行動を行うために必要な前提条件と、その行動を行った結果得られる効果が付与されている。行動を取り出すことができれば、その行動の効果がユーザのゴールであると推測する。

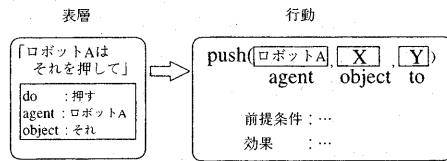


図4: プランオペレータへのマッピング

ここで、発話の述部が省略されている場合には、プランライブラリとの対応が取れないので、まず述部の省略を解決する。これには、従来の傀儡における省略解決手法と同様に、先行文脈を参照する手がかり語などを利用する。

4.3 照応・省略解決

4.3.1 プランベースの照応・省略解決

抽出したゴールに含まれている照応・省略を解決するために、まず世界内の実体に対し、表層に現れてい

る情報によって制約をかけて、先行詞の候補を絞り込む。たとえば、「それ」「その」という指示詞がある場合、候補はユーザから見えているものに限定される。ここで得られた候補に対し、本手法では次の二つの点にプランニングの知識を利用して、先行詞を同定する。

1. ユーザのプランを認識して参照先の発話を同定
ある行動の効果が、別の行動の前提条件となっていた場合、この二つの行動間には意味的な関連性があり、同一ゴールを目指した発話を分割したものであると考えられる。そこで、行動の効果と前提条件の連鎖から、現時点までのユーザのプランを認識する(図5)。照応・省略を含む発話が入力された場合に、その発話が属するプランがわかれば、先行詞を探す範囲を同一プラン内の発話に制限できる。

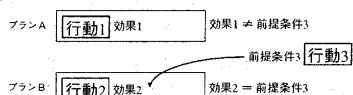


図 5: プランの同定

2. プラン内の情報を利用して先行詞を同定
プラン認識時に、参照先発話の効果と、照応・省略のある発話の前提条件が一致しているので、これらのマッチングを行うことで、発話中の参照すべき部分を同定する(図6)。対話例1の発話#1-2と発話#1-3の場合、照応詞「それ」と先行詞「青い球」の格は一致していないが、本手法ではこのような場合にも先行詞を同定できる。

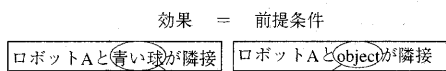


図 6: 先行詞の同定

4.3.2 効果と前提条件の部分マッチング

一般に、行動には複数の前提条件と効果があるので、マッチングの際にはそのうち一つでもマッチすればよいとする。これは、ユーザがプランニングした結果の行動を逐一指示するとは限らず、途中を省略してロボットに推論させる場合があることを考慮している。たとえば、次の対話例3に示す場合である。

対話例3:

#3-1「赤い球の前に行って」

#3-2「それを右に押して」

発話#3-2が指示された時点から赤い球を右に押すためには、その前にまず球の左に移動して、それから右を向くという行動をしなければならない。図7に示すように、ユーザはまずロボットに対象を同定させるためにプラン中の行動1「球の前に行く」を指示し、次に行動4「右に押す」を指示しており、途中の行動2「球の左に移動」と行動3「右を向く」の指示は省略している。この場合には、発話#3-2の時点で行動4を実行するのに必要な前提条件を全ては満たしておらず、赤い球とロボットが隣接しているという条件だけが成立している。こういった場合にも行動1と行動4の関係を捉えられるように、部分マッチングを行う。

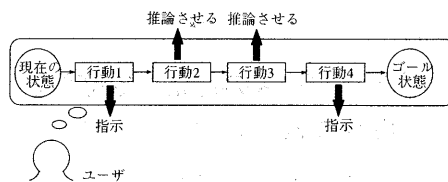


図 7: ユーザのプランニングと指示の省略

4.3.3 効果と仮想世界の状態

これまで述べてきたように、行動の効果と前提条件のマッチングを行うことで、照応・省略を解決することができる。しかし、ある行動による効果は、他の行動によって失われない限り持続する。そのため、同一プラン内の直前の行動の効果だけではなく、現在持続しているすべての効果とのマッチングをとるべきである。現在の仮想世界の状態には、先行するすべての行動の効果が反映されて、持続している効果のみが残っている。したがって、現在持続しているすべての効果を捨てるためには、図8のように前提条件と現在の仮想世界の状態とのマッチングを取ればよい。これにより、文脈に応じて起こる、プランライブラリに明示されていない効果も捉えることができる。

また、直前の行動の効果によるものではなく、もともと前提条件が満たされていることもある。その場合には、いきなり「それ」という指示を与えても、同様の推測が可能である。行動の前提条件と現在の仮想世界の状態のマッチングをすることで、このような直示表現にも対応できる。

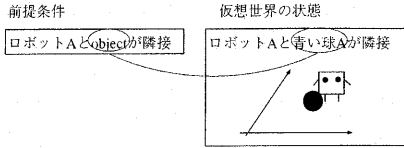


図 8: 仮想世界と前提条件のマッチングによる解決

4.3.4 注目度スコアによる先行詞の決定

先行詞の候補を絞り込んだ結果、候補が複数残ったり、一つも残らない場合が出てくる。そこで、候補が複数残った場合は、その各候補にスコアを与える。候補が一つも残らなかった場合は、表層から選出した各候補にスコアを与える。そして、最もスコアの高いものを先行詞とする。本手法では、次の式でスコアを計算する。

$$score = \alpha S1 + \beta S2 + \gamma S3$$

α, β, γ は人手で付けた重みであり、 $S1, S2, S3$ は次の要素である。

$S1$: 先行文脈からの注目度

$S2$: 実行するロボットの視線による注目度

$S3$: ユーザ（カメラ）の視線による注目度

$S1$ は、その候補について言及された頻度と、どれだけ最近言及されたかの履歴をもとに計算する。また、 $S2, S3$ は、現在の仮想世界の状態を見て、距離の近さと、視界の中心線との角度をもとに計算する。

5 結果と考察

提案手法を対話コーパス中の対話例に人手で適用し、評価した。このコーパスは、人間が傀儡と同様のシステムを模擬し指令を受けて行動する様子を収録して、人手で格解析したものである [1]。その結果、本手法は従来の手法で解決できなかった例のうち、主格と目的格の照応・省略に特に有効であった。

しかし、本手法はまだ次の点で不十分である。

位置表現の照応・省略

「〇〇にある」という主格や目的格の参照先を示す位置表現や、場所格の中の照応・省略解決には、まだ不十分な点が多く残っている。位置に関する照応・省略（特に省略）は、発話の間にプランの関係があって起こる場合よりも、空間表現自体の曖昧性によるものが多い。たとえば、次の対話例 4 の場合、ユーザから見て

右という解釈も、ロボットから見て右という解釈もありえる。

対話例 4:

#4-1 「ロボット A は右にある赤い球を押して」

この曖昧性の解決は、照応・省略の問題の前提となるので重要である。こういった空間表現の曖昧性の一部を解決するには、本手法のスコアで取り入れた、注目度が利用できると思われる。たとえば、主格として注

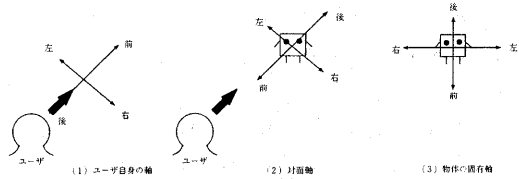


図 9: 方向軸

目しているか、目的格として注目しているかで、位置表現のための方向軸（図 9）の変化を追う方法も考えられる。

様態を表す副詞的な表現に関する照応・省略

コーパス中に、「〇〇のように」という様態を表す副詞表現を用いた指示が見られる。このような表現は格解析しにくく、現状では扱えていないため、もっと分析する必要がある。いろいろな表現があり、現段階では次の二種類が発見されている。

- 結果の様態
- プロセスの様態

結果の様態は、ゴールを直接指示するものであるから、プランニングの概念で扱いやすい。対話例 5 がその例である。

対話例 5:

#5-1 「カメラは赤い球が中心に見えるように（映して）」

ゴールを直接指示された場合、述部が省略されることがあるが、その省略は指定されたゴールを効果に持つ行動をプランライブラリ中から検索することで、解決できるようになる。一方、プロセスの様態は、対話例 6 のようなものである。

対話例 6:

#6-1 「ロボット A が画面からはみ出すことのないように（映して）」

これに対応するには、まず状態を継続する行動を扱え

るように、動作の途中の状態を監視する動的なプランニングを導入する必要がある。そして、照応・省略を解決するために、さらに制約などの情報も考慮することが必要である。

また、対話例7のように様態表現内に照応・省略がある場合には、その表現をさらに主格や目的格に分解できれば、提案手法が適用できると思われる。

対話例7:

#7-1「ロボットAは赤い球の前に行って」

#7-2「それが青い球の隣に来るように押して」

プランオペレータへのマッピング

本手法では、対話例8のような、手がかり語を用いずにそれ以前のプランとの並列性を持っている例も解決できない。まず最初に述部の省略「行って」のみを補った時点で、ロボットBが今の位置から左に移動するという解釈が可能になり、「赤い球」という情報が欠如する。これは、述部を単純に動詞のみで扱うのではなく、その他の格も考慮してプランオペレータにマッピングすべきであることを示している。

対話例8:

#8-1「ロボットAは赤い球の右に行って」

#8-2「ロボットBは(赤い球の)左(に行って)」

実体以外を指示する照応・省略

今回は世界内の実体を指示する照応・省略のみを対象としたが、対話例9のような、イベントなど実体以外のものを指示する照応・省略もあり得る。この場合、プランの概念でイベントのまとまりを扱うことが考えられる。

対話例9:実体以外を指示する場合

#9-1「それ、もう一度やりなおして」

6 おわりに

本稿では、ロボットを自然言語で操作するタスクにおける、プランニングの知識を用いた照応・省略解決手法を提案した。本手法では、行動の持つ前提条件と、以前の行動の効果が反映された仮想世界の状態のマッピングを行うことにより、先行詞を同定する。これにより、従来のような表層に現れる情報に基づく手法では解決できなかった、関連性のある行動間での非明示的な焦点の移動に追従することが可能になった。

今回は、限定された単純な行動のみを対象としているが、今後、複数のロボットでの協調作業や、動的なプランニングが必要な行動など、複雑な行動を扱えるように拡張することが考えられる。提案手法では、照

応・省略解決に用いる知識をロボットが行動するためのプランニングに用いる知識と共有しているため、拡張性に優れている。しかし、プランオペレータへのマッピング時の曖昧性など、新たなアルゴリズムの拡張が必要となることも考えられる。

また、人間は言葉で指示を与えながら、同時に視線や指差しなどのジェスチャーによって対象を指し示しており、照応・省略の解決を行う際に、パラ言語学的な情報も重要な要素となる。たとえば、カメラをさきかえて視線が移動したときにも非明示的な焦点の移動が起こることから、現状のようにカメラの向きでユーザの視線を考慮するのではなく、視界内で具体的に何を見ているか、より詳細な情報を取得できることが望まれる。

参考文献

- [1] <http://tanaka-www.cs.titech.ac.jp/pub/qdc/>.
- [2] Donna Byron and Amand Stent. A Preliminary Model of Centering in Dialog. In *proceedings of ACL '98*, pp. 1475-1477, 1998.
- [3] Philip R. Cohen. The Pragmatics of Referring and the Modality of Communication. *Computational Linguistics*, Vol. 10, No. 2, pp. 97-146, 1984.
- [4] Barbara J. Grosz, Arvind K. Joshi, and Scott Weinstein. Providing a Unified Account of Definite Noun Phrases in Discourse. In *proceedings of ACL*, pp. 44-49, 1983.
- [5] Megumi Kameyama. A Property-Sharing Constraint in Centering. In *proceedings of ACL*, pp. 200-206, 1986.
- [6] Roger C. Schank and Robert P. Abelson. *Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*. L. Erlbaum, 1977.
- [7] 新山祐介, 徳永健伸, 田中穂積. 自然言語を理解するソフトウェアロボット: 傀儡. 情報処理学会論文誌, Vol. 42, No. 6, pp. 1359-1367, 2001.
- [8] T. Winograd. *Understanding Natural Language*. Academic Press, 1972.