

基盤化に基づく修復発話の理解

船越孝太郎 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 計算工学専攻

{koh,take,tanaka}@cl.cs.titech.ac.jp

1 はじめに

指令者が追従者の行動を言語で制御する対話を行動制御対話とよぶ。本稿では、行動制御対話において、指令者が発する(1-3)のような発話を、機械の追従者が理解するための枠組みを示す¹。(1-3)のように誤りを修復する機能を有する言語表現を含み、かつその修復対象がそれと同一ターンの中に存在しない発話を、修復を目的として発せられた発話という意味で修復発話とよぶ。

- (1-1) A: その本を右の棚において
(1-2) B: はい <(1-1) を解釈して行動する >
(1-3) A: そっちじゃない

2 修復発話理解の難しさ

修復発話の理解は修復対象の特定と修復内容の解釈の二つからなる。修復対象の特定とは修復発話の発話者が何を修復しようとしてその発話を発したかを認識することである。修復内容の解釈とはその修復対象をどのような状態にしたいのかを認識することである。

行動制御対話における修復発話を理解することの難しさは、時として修復発話が断片的となり、その発話を見ただけでは修復対象が何で修復内容が何なのかを知ることができない点にある。修復内容の解釈は修復対象に強く依存するが、その修復対象を特定する情報が発話内に存在しないという事態がしばしば発生する。

例えば対話(1-1-1-3)の場合、修復発話(1-3)は、「そっち」という指示詞を否定することで、Bの(1-1)に対する解釈の中で物体あるいは方向に関する選択が誤りであることを提示している。しかし、この発話単体からはBが「本」の解釈を誤ったのか「右の棚」の解釈を誤ったのか分からない。

(1-3)の解釈はそれがいつ発話されたかによって変わる。つまり(1-3)のような修復発話の解釈はその発せられた状況に依存する。(1-3)がBが本を手を取ろうとしたときに発せられたのであれば、それはBによる本の選択を修復していると解釈するのが自然であるし、Bが既に本を手を取った後にそれをまさに棚に置こうとしているときに発せられたのであれば、それはBによる棚の選択を修復していると解釈するのが自然である。

それでは、Aが(1-3)を発したのが、Bが手に持っている本を棚に置こうとしているときであったとき、Bはなぜ「本」の選択ではなく「棚」の選択が誤っていると判断できるのだろうか。本稿では、この問題を基盤化

[1, 2, 4] の概念によって説明する。基盤化とは複数の主体が対話を通じて相互信念を形成する過程である。

行動制御対話では対話の中にそれぞれの対話者の発話に加えて追従者の行動が含まれる。この追従者の行動を基盤化行為(4.1節参照)として捉え、基盤化の状態を細かく把握することで、修復対象の特定を行い修復発話を解釈する。Bが手に取った本を棚に置こうとしている時、Bによる本の選択は本を手を取るという行動を通じて既に基盤化されていると考えれば、(1-3)による修復の対象が既に基盤化された本ではなく、まだ基盤化されていない棚に対するものであると説明できる。

前述のように修復内容の解釈は修復対象に依存する。例えば、「逆」という修復発話は、修復対象が物体の向きであればそれを逆さまにすることを要求しているのであろうし、修復対象が移動の方向であれば180度反対の方向に移動することを求めているのであろう。そこで本稿では、修復発話の言語表現を解析してそこから修復対象と修復内容を抽出するのではなく、修復表現から得られる手がかりと基盤化の状態を基に談話の中から修復対象の候補を見つけ出し、特定した対象に応じて修復内容を解釈するというアプローチを取る。

3 行動を伴う対話の扱い

行動制御対話では、発話と同様に追従者の行動が対話を織りなす要素として重要な役割を担う。本稿では、従来発話だけで構成される対話を中心に議論されてきた談話の理論を、行動制御対話のように発話と同程度もしくはそれ以上に行動を含む対話に適用する。だが、行動を伴う対話に従来の談話理論を適用するには、次の二つの問題に対処しなければならない。

発話と行動の重なり Clarkが指摘するように対話中の行動は一般に話者交代のルールに従わない[1]。しかし、行動がある程度明確な話者交代を示さない限り従来談話の理論を適用することは難しい。そこで本稿では、追従者が(2)、(3)に従って振る舞うことを前提とし、追従者の行動が指令者の発話と重ならないよう制限する。

- (2) 話者交代が許されるまで行動を開始しない。
(3) 行動中に発話で割り込まれたら直ちに停止する。

行為の階層性 一つの行為は複数の下位行為から構成され、全体として階層構造をなす。行動を発話と同様に談話の理論で分析するためには、分析の対象となる単位を設定する必要がある。本稿では、行為の階層構造の中に、人間が共通して行為の単位として認識できるレベルがあると仮定する。そのような行動はおおよそ、取る、歩

¹本稿についてのより詳しい議論は[3]を参照されたい。

く、見る、といった多くの言語に普遍的に含まれる基本動詞に対応すると考える。このような行動を基本行動とよぶ。

4 基盤化と修復発話の理解

修復内容の解釈はその修復対象に依存するため、修復対象の特定が修復発話理解の出発点となる。本節では、Traumの基盤化状態の計算理論 [4] を基に修復対象を特定する手法を提案する。

4.1 基盤化状態の計算理論

Traum[4] は有限状態遷移モデルを用いて対話の各時点における基盤化の状態を各主体の立場で動的に計算するモデルを提案した。そのモデルでは、談話中の発話²を基盤化を進めるための行為(基盤化行為)として捉える。

基盤化の過程は談話単位という理論的構築物を用いてモデル化される。談話単位は関連する基盤化行為の集合である。対話は談話単位の連なりとして捉える。談話単位を開き基盤化を開始した主体を開始者とよび、その他の主体を応答者とよぶ。

談話単位は表 1 に示す七つの基盤化状態のいずれかを取り、表 2 に示す七つの基盤化行為が入力される度、表 3 に示す遷移表に従って次の状態に遷移する(入力基盤化行為の末尾の I と R は開始者と応答者を表す)。基盤化行為は各発話に対して付与される。各談話単位の内容が基盤化されるためにはその談話単位が終了状態 F に到達している必要があるが、終了状態 F に到達してもその談話単位に関する対話が続くこともある。

表 4 に談話単位を用いた基盤化過程の分析例を示す。話者 A の第一発話を明確に理解することが出来なかった話者 B は次の自己の発話で問題点に対する修復を要求し(ReqRep-R)、それに対して A は第三ターンで修復を行っている(Repair-I)。最後に B が、A の第一発話を理解した証拠を示したところで、談話単位が終了状態に至る。

4.2 基盤化状態を用いた修復対象の特定

4.2.1 基盤化の度合いと証拠強度の導入

Traum 自身 [4] も指摘するように、ある命題が基盤化されている度合いを、基盤化されている/されていないの二通りだけで表すのは、過度の単純化である。またそのような度合いの差を考慮しないと、本稿で問題とする修復対象の特定が行えない。そこで、証拠の強さを四つの段階で定義する。基盤化の度合いは証拠の強度と同一視する。ここでは、紙面の都合上、最小強度の強度 1 と最大強度の強度 4 だけを記す。

強度 1: 応答者が発話の解釈に成功したらしいということ を裏付ける。例えば、「右に行って」という要求に対して「はい」と答えることは、応答者が開始者

²正確には発話単位 [4]。

S	初期状態。
1	談話単位への命題の追加や修復を続行している状態。
2	応答者が修復を要求している状態。
3	応答者からの修復を受けた状態。
4	開始者が修復を要求している状態。
F	終了状態。基盤化されたとみなせる。
D	中止状態。該当の談話単位は放棄される。

表 1: 基盤化状態

開始	Initiate	新しい談話単位を開始する
続行	Continue	内容を追加する
認定	Ack	理解の証拠を示す
修復	Repair	誤解を修復する
修復要求	ReqRepair	修復を求める
認定要求	ReqAck	認定を求める
取り消し	Cancel	談話単位の処理を放棄する

表 2: 基盤化行為

入力基盤化行為	S	1	2	3	4	F	D
Initiate-I	1						
Continue-I		1			4		
Continue-R			2	3			
Repair-I		1	1	1	4	1	
Repair-R		3	2	3	3	3	
ReqRepair-I			4	4	4	4	
ReqRepair-R		2	2	2	2	2	
Ack-I				F	1*	F	
Ack-R		F	F*			F	
ReqAck-I		1				1	
ReqAck-R				3		3	
Cancel-I		D	D	D	D	D	
Cancel-R			1	1		D	

* 修復要求が無視された

表 3: 基盤化の有限状態遷移モデル

発話	談話単位 1	
A: can I speak to Jim Johnstone please?	Init-I	1
B: senior?	ReqRep-R	2
A: yes	Repair-I	1
B: yes	Ack-R	F

表 4: 談話単位を用いた対話の分析

の発話を解釈できたことを示す。しかし、その解釈が本当に正しいかどうかは分からない。

強度 4: 応答者が開始者の意図の理解に成功したことを示す。例えば、「右に行って」という要求に対して実際に行動して答える場合がこれにあたる。

4.2.2 修復対象の特定

基盤化の枠組みの上で修復対象を特定する仕組みを説明する。以下の議論は追従者の視点で行う。

対話(4-1-4-5)を考える。Cが指令者で、Fが追従者である。(4-2)のAck₁-R:Fという記述は、(4-2)が談話単位1に対する応答者(R)からの認定(ACK)の基盤化行為で、それによって談話単位1の状態がFに遷移したことを表している。

- (4-1) C: 前の赤い玉を右に置いて Init₁-I:1
- (4-2) F: はい Ack₁-R:F
- (4-3) F: < 玉を掴む > Init₂-I:1
- (4-4) F: < 玉を動かす > Cont₂-I:1
- (4-5) C: そっちじゃない Repair₂-R:3

追従者Fは、修復発話(4-5)を理解するために「そっち」が指し示す修復対象を特定しなければならない。(4-5)は、Fの行動(4-3)、(4-4)に対して発せられており表層的な照応解決では扱えない。ただし修復対象を示す「そっち」が方向、場所、物体の内のいずれかを差していることは、代名詞の持つ語彙的な制約からわかる。この場合、(4-5)の修復対象は(4-4)でFが玉を動かした場所であり、(4-1)でCによって「右」と参照された(とFが解釈した)場所である。

(4-1)の伝達内容は、(5)のように命題の集合として記述できる。

- (5) $\alpha = \text{Request}(C, F, \text{Put}(\#Agt1, \#Obj1, \#Dest1))$
 - (a) 発話行為 $\text{speechActType}(\alpha) = \text{Request}$
 - (b) 提示者 $\text{presenter}(\alpha) = C$
 - (c) 受信者 $\text{addressee}(\alpha) = F$
 - (d) 行動 $\text{actionType}(\text{content}(\alpha)) = \text{Put}$
 - (e) 動作主 $\text{agent}(\text{content}(\alpha)) = \#Agt1$,
 $\text{referent}(\#Agt1) = F$
 - (f) 対象 $\text{object}(\text{content}(\alpha)) = \#Obj1$,
 $\text{referent}(\#Obj1) = \text{Ball1}$
 - (g) 目的地 $\text{dest}(\text{content}(\alpha)) = \#Dest1$,
 $\text{referent}(\#Dest1) = \text{Pos1}$

α は伝達内容全体を表す。speechActTypeのように、小文字で始まる表象は述語である。例えば、(5a)は伝達内容 α の発話行為を表し、その値はRequestであることを表している。大文字で始まる表象はシンボル定数である。RequestとMoveはそれぞれ発話行為と基本行動の名前である。CとFは前述のように対話者を表し、Ball1は世界の中のある一つの玉を表す。Pos1は世界の中のある一つの場所を表す。#で始まる表象は談話中に導入される観念的な存在で、談話参照物とよぶ。談話

参照物は言語的に表現されたなにものかを表す。例えば(5f)の#Obj1は指令者Cによって「前の赤い玉」と表現されたなにものかを表す。この場合その「なにものか」はBall1である。我々は談話参照物と世界の中の実体を結びつけながら対話を行う。行動制御対話における多くの問題は、この実体の特定を誤ることから発生する。

(4-1)を解釈し(5)を得た追従者Fは、Put(#Agt1, #Obj1, #Dest1)を遂行するために行動計画(6)を立案し、左から順に実行する。

- (6) Grasp(#Agt1, #Obj1), Move(#Agt1, #Obj1, #Dest1), Release(#Agt1, #Obj1)

(4-1-4-5)を、(5)のように述語形式で記述したものを(7-1-7-5)に記す。Displayは追従者の行動の提示を表す。Repairは修復発話を表す。

- (7-1) C: Request(C, F, Put(#Agt1, #Obj1, #Dest1))
- (7-2) F: Accept(F, C, α)
- (7-3) F: Display(F, C, Grasp(#Agt1, #Obj1))
- (7-4) F: Display(F, C, Move(#Agt1, #Obj1, #Dest1))
- (7-5) C: Repair(C, F, #Dest1)

ところで、Traumのモデルでは発話を最小の粒度として基盤化過程を記述したが、修復対象の特定は命題レベルの粒度で行う必要がある。そこで(5)に示したような命題を基盤化の最小の粒度として考える。その上で修復対象の特定に関して以下の仮説をおく。

(8) 修復対象特定に関する仮説

- (a) 追従者の行動に割り込んで指令者が発する断片的な修復発話の修復対象は、その割り込みの時点で追従者が行っていた行動によって理解の証拠が提示されるものである。
- (b) 対話者は自己の対話を常に観察している。実際に行動してみせることは、強力な証拠(強度4)を提示するので、もしそこに誤りがあれば指令者はすぐに割り込んで修復する。
- (c) 逆に、追従者が実際に行動してみせても指令者が修復しないならば、追従者は自分の解釈は正しかったと判断できる。
- (d) 修復対象でありうる命題が複数ある時は、より強い証拠が提示されており、かつその証拠よりも小さい強度でしか基盤化されていないものが選択される。

(8)に則って(7-2)から(7-5)までの流れを追って見る。まず、(7-2)(つまり(4-2))によって α 全体が強度1で基盤化される。次に、(7-3)でFはGrasp(#Agt1, #Obj1)を実行する。この動作が完了してもCからは否定的な反応がないので、(8c)により、Fは#Agt1と#Obj1の解釈が正しかったことが確認されたと判断し#Agt1と#Obj1を証拠強度4で基盤化する。次にMove(#Agt1, #Obj1, #Dest1)を実行したところ、その途中で指令者から(7-5)による割り込みをうける。

「そっちじゃない」という断片的な修復発話(7-5)をFはRepair(C, F, x)と解釈するが、修復対象である

x が言語表現からは判断できないので文脈情報である談話状態から特定を試みる。(8a) から修復対象は割り込まれた行動 (7-4) によって提示された命題と考え、また「そっち」という語の制約によって修復対象の候補から行動や発話行為の種類 (actionType, speechActType) を除外し、談話参照物の #Agt1, #Obj1, #Dest1 のいずれかに絞り込む。ところで #Agt1 と #Obj1 は、(7-3) の完了により既に強度 4 で基盤化されている。そこで、(8d) により、F はまだ強度 1 でしか基盤化されていない #Dest1 を修復の対象と判断する。

まとめると、仮説 (8) に基づく修復対象の特定手法は以下ようになる。

(9) 基盤化の状態に基づいた修復対象の特定手法

- (a) 言語表現から修復対象の型を絞り込む
- (b) 直前の提示内容の中から (9a) と合致するもののリストを抽出する
- (c) (8d) に従って、(9b) で抽出した修復対象の候補を順位付けし、最尤のものを修復対象として選択する

ここで、(9) に示した手順は基盤化の観点から修復対象候補の尤度を順位づけるだけであることに注意しなければならない。手順 (9) で特定した修復対象に対して修復発話の修復内容を解釈しようとしても妥当な解釈が見つからないことは十分起こりうるが、そのような事態に応答者がどのような対応を取るべきかは、今後の研究課題である。

4.2.3 修復内容の解釈

(9) によって修復対象を特定したあとは、その修復対象の特性に応じた修復内容の解釈を行う。現状では、修復内容の解釈を一般的に記述することは難しく、対象ごとに解釈のバリエーションを定義するしかない。例えば、修復発話が「逆」というものであったとき、修復対象が物体の向きであればそれを逆さまにすることが求められているし、修復対象が移動の方向であれば 180 度異なる方向に移動することが求められている。また、「机の右側を通る」か「机の左側を通る」とかという背反的な経路の選択に関することであれば、選択した経路ではない方の経路を選択する必要がある。

4.3 行動の中での修復発話の理解

4.2 節では、基本行動を発話単位と同様に扱って行動制御対話を分析し、修復の対象を特定する手法を提案した。この手法では、各基本行動を時間軸上の一点での現象として単純化して扱っている。しかし実際に連続した時間の中で行動に対する修復を扱う際には、行動とその時間的な幅に起因する問題に対処する必要がある。これらの問題に対する詳しい議論は [3] を参照されたい。

5 まとめ

本稿では、行動制御対話における修復発話を理解するための枠組みと手法を提案した。修復発話の理解は修復対象の特定と修復内容の解釈からなり、修復内容の解釈が修復対象の特定に依存していることを指摘した上で、修復発話を理解するための枠組みを提示した。その枠組みでは、発話が修復発話と判定された場合、まず修復対象を特定することから始め、その後候補として挙げられた修復対象毎に修復内容の解釈を試みる。修復対象の特定については、基盤化という観点からそれを説明した。本稿では特に断片的な修復発話の理解について議論したが、例えば修復発話が断片的でない場合でも、修復対象を特定することは解釈の探索空間を限定するために必要であり、本稿で提案したアプローチが有効である。

本稿で提案した修復対象特定の手法は、修復対象の絞り込みに表層の情報 (言語表現) を用いるが、特定そのものには利用しない。このことは、対話システムの頑健性に大きく貢献でき、特に音声認識器の誤認識に対して有効である。

表層の情報は一般に修復対象を特定するには十分ではない。例えば、発話 (10) に応答して行動している追従者に対して、指令者から発話 (11) による修復がかかったとする。

- (10) その赤い玉を机の右に置いて
- (11) ごめん青だった

このとき、表層的な情報から修復対象を特定しようとするならば、語彙的な類似性などから最も可能性が高いと思われるのは「赤い玉」であろう。しかし、この方法は (12) のような発話になるとすぐに破綻する。しかし、提案手法のように基盤化の状態に目を向ければ、発話のタイミングによって「赤い玉」と「緑の机」のどちらが修復されている可能性が高いかを判断することができる。

- (12) その赤い玉を緑の机の右に置いて

また (12) が (13) のように「赤い」が欠落して認識された場合でも、基盤化の状態から「玉」と「緑の机」のどちらが修復されている可能性が高いかを判断することができる。

- (13) その玉を緑の机の右に置いて

参考文献

- [1] Herbert H. Clark. *Using Language*. Cambridge University Press, 1996.
- [2] 石崎雅人, 伝康晴. 談話と対話. 東京大学出版会, 2001.
- [3] 船越孝太郎. 行動制御対話における修復表現の理解に関する研究. PhD thesis, 東京工業大学 大学院情報理工学研究科, 2005.
- [4] David Traum. *Toward a Computational Theory of Grounding*. PhD thesis, University of Rochester, 1994.