

談話構造を利用した漸次的発話生成

堂坂 浩二 島津 明

NTT 基礎研究所

1 はじめに

対話は実時間の活動であり、時間制限によって制約される。さらに、対話従事者は制限された推論能力しかもたない [9]。時間と推論能力に関する制限のために、話者は発話を漸次的に生成せざるを得ない。漸次的発話生成とは、「話者が、発話すべき内容を考えながら、同時に発話を生成し、発話を行いながら、次に発話すべき内容を決定していく」ことをいう。

漸次的発話生成戦略のおかげで、話者は時間制限を守ることができ、黙っている間に他の対話従事者に発話権を奪われなくて済む。さらに、この戦略に基づく発話は、推論能力に限界のある聴者にとって理解が容易である。

本稿では、タスク指向型対話における漸次的発話生成のモデルを示す。このモデルは、与えられた問題を解きながら、同時に問題の解を提案するために発話を生成する。このとき、発話の間に一定の長さ以上のポーズを置かないという時間制限を遵守する。

このモデルは、一連の発話を一塊の談話として生成する。本稿では、タスク指向型対話のコーパスに現れる談話構造の分析結果を示し、頻繁に現れる談話構造が漸次的発話生成に貢献することを論じる。モデルは、そういった談話構造を利用することによって、発話を漸次的に生成する。また、発話の漸次的生成戦略は、発話内容が完全に決まる前に発話を開始するため、不適切な談話を生成しがちである。談話の適切さを保証するために、言語運用制約と文脈モデルを用いる。モデルの有効性は、発話シミュレーション実験によって確認されている。

2 関連研究

Levelt は、認知心理学的な見地から発話内容決定過程と発話形成過程が同時に進行することによって、発話が生成されるという考え方を示している [5]。しかし、認知的モデルは、人の発話生成の計算機シミュレーションや対話システム構築という目的のためには不十分なものである。本稿では、漸次的発話生成の計算モデルを示す。

POPEL は、漸次的発話生成の実働システムである [7]。POPEL は、POPEL-WHAT と POPEL-HOW の 2 つのモジュールから構成される。POPEL-WHAT は発話内容を決定し、段階的に POPEL-HOW に送る。それを受けて、POPEL-HOW は、段階的に文を形成する。しかし、ここでは、発話内容を決定しながら、談話構造に

基づいて一連の発話を漸次的に生成するための方策が示されていない。第 3 節で示すように、実時間対話では、話者は伝達すべき内容を複数の発話に分配することにより漸次的に発話を生成し、しかもそれらの発話が全体で一つの談話を構成することが頻繁に起きる。本稿で示すモデルは、談話構造に基づいて漸次的に談話を生成する。

言い淀みや言い直しを説明するために、時間制限を考慮した発話生成のアーキテクチャの構想が示されている [1]。しかし、そこでは、具体的な計算モデルは示されていない。また、談話構造も考慮されていない。

3 談話構造分析

タスク指向型対話のコーパスの中に現れる談話構造を分析した。話者が発話によって情報を伝達しようとするとき、伝達すべき情報全体を一塊として発話を生成するのではなく、情報を複数の塊に分配して、その塊ごとに発話を生成するものと考えられる。そのような情報の塊を情報単位と呼び、一つの情報単位を伝達する発話を発話単位と呼ぶ。発話単位は談話の最小構成要素である。話者は各発話単位を関係付けて、発話単位の列を一つの談話として生成する。そういった発話単位の間の関係を談話関係と呼ぶ。談話構造とは、談話関係によって互いに関係付けられた発話単位の階層構造である。

対話コーパスは次の会話実験によって収集した。実験には、90 人の被験者が参加した。各会話において、二人の被験者 (N と E) は、電話で会話することにより、「 N がある場所から別の場所へ行く」という問題の解を見つけることが要請された。被験者 N と E は、 E は問題を解くために十分な知識をもっており、 N はもっていないように選んだ。80 対話が収集され、書き起されている。15 対話を分析のために選んだ。

3.1 発話単位の分析

発話単位は、必ずしも表層の語彙・統語的な範疇に対応するわけではないが、第一近似として発話単位は次のように分析できると仮定した。

- 節は発話単位である。
- 間投詞的表現は発話単位である。
- つなぎ語は発話単位の区切りである。

節	929	NP と PP の列	41
間投的表現	665	副詞類	10
NP または PP	279	その他	4
接続詞類	84	計	2012

表 1: 発話単位の統語範疇の頻度分布

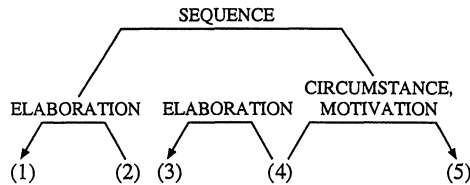


図 1: 対話 (d1) における談話関係

次の対話 (d1) は書き起こされた対話の一部である。“/” は発話単位の区切りを示す。

対話 (d1)

- E:(1) 愛甲石田までですね / (2) 行って / <はい>
 (3) そこでバスなんです / (4) 森の里青山行き
 という<はい>バスがあるので / <はい>
 (5) それに乗ってもらって / <はい> . . .

15 対話における発話単位の統語範疇の頻度を表 1 に示す。表中で、NP と PP は名詞句と後置詞句を意味する。節、NP、PP、または、NP と PP の列として実現された発話単位について、一つの発話単位の中に現れる名詞句数の平均は 1.01 であり、分散は 0.28 であった。

3.2 談話関係の分析

談話関係として、修辭構造理論 [6] で提供される関係を用いる。また、文献 [4] の提案に従って、談話関係を意味関係と意図関係に分けて考える。意味関係とは、発話によって言及される命題内容の間の関係であり、意図関係とは、聴者の意図や信念に影響を与えるための関係である。対話 (d1) における談話関係を図 1 に示す。表 2 は談話関係の頻度分布を示す。

頻繁に現れた談話関係として、ELABORATION、CIRCUMSTANCE、MOTIVATION が漸次的発話生成に果たす役割を論じる¹。ELABORATION は、対話ドメインの行為、状態、対象の内容の一部を記述する発話単位と、さらに詳細な内容を記述する発話単位との間の関係である。典型例は、対話 (d1) の (3) と (4) の間の関係である。この関係によって、話者は記述対象の内容を完全に決定する前に発話を開始できる。CIRCUMSTANCE は、ドメインの行為や状態を記述する発話単位と、行為や状態の環境を記述する発話単位との間の関係である。特に、行

¹SEQUENCE は漸次的発話生成と直接の関係は認められなかった。

ELABORATION	305
SEQUENCE	74
CIRCUMSTANCE	60
RESULT	25
CONDITION	25
PURPOSE	2
CONTRAST	1

(a) 意味関係

MOTIVATION	66
BACKGROUND	14
EVIDENCE	10
INTERPRETATION	6
CONCESSION	3
ENABLEMENT	1

(b) 意図関係

表 2: 談話関係の頻度分布

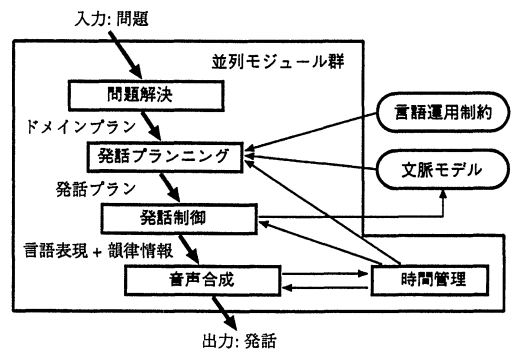


図 2: モデルの概観

為の提案の前に行為の前提を環境として記述する場合は、41 例 (68%) 見られた。その典型例は、(4) と (5) の間の関係であり、(4) で「バスが存在する」という前提が記述されている。行為の前提を記述する際には、行為の構成要素を言及することが普通である。したがって、この関係を使うことによって、話者は行為の内容を 2 つ以上の発話単位に分配することができ、行為の前提条件を記述するときに行為全体の内容を決定しておく必要はない。MOTIVATION は行為を提案する発話単位と、その行為の採用を聴者に動機付ける事実を記述する発話単位との間の関係である。典型例は、(4) と (5) の間の関係であり、行為の前提が成立していることをまず言ってから、行為の提案を行っている。このときは、CIRCUMSTANCE と MOTIVATION の 2 つの関係が同時に成立し、行為の内容を 2 つ以上の発話単位に分配できる。

4 時間制限下における漸次的発話生成モデル

モデルの概観を図 2 に示す。モデルは、問題解決、発話プランニング、発話制御、音声合成、時間管理を行うモジュールから構成される。問題解決モジュールは、与えられた問題を解くドメインプランを立案する。発話プランニングモジュールは、発話プランを立案する。発話プランは、ドメインプランを対話相手に提案するための

談話構造と個々の発話単位の内容を規定する。発話制御モジュールは、発話プランに基づいて、韻律情報を付加した言語表現を一定のタイミングで音声合成モジュールに送る。本稿では、言語表現の合成と韻律情報の付加に関しては触れない。音声合成モジュールは音声として発話を発する。時間管理モジュールは、音声合成モジュールを走査し、ポーズが一定の制限時間を越えたとき、発話プランニング及び発話制御モジュールにそのことを知らせる。

4.1 時間制限の遵守

時間制限を守りながら漸次的に発話を生成するために、以下の手法を用いる。すべてのモジュールは並列に動作する。問題解決と発話プランニングは、階層的プランニング手法 [8, Chap.12] に基づいて段階的に進める。ドメインプランは、段階的に詳細化され、発話プランは部分的に決定されたドメインプランに基づいて立案される。ポーズが制限時間を越えたらいつでも、その時点で立案されている発話プランに基づいて、発話は実行される。発話を実行している間にもドメインプランと発話プランの立案は進められる。ドメインプランがさらに具体化されたときには、発話プランを再立案する。制限時間内に如何なる発話プランも立案されていないなら、発話制御モジュールはつなぎ語を生成する。

4.2 発話の適切さの保証

生成される一連の発話が談話として適切であることを保証するために、言語運用制約と文脈モデルを用いる。ここで用いる言語運用制約は、以下のものである。

- (c1) 注視対象を代名詞化する。
- (c2) 伝達済みの情報は繰り返さない。
- (c3) 注視状態に対して適切な発話を行う。

文脈モデルは、伝達済みの情報を管理し、注視状態 [3] を追跡する。例えば、場所 S が注視されている状態で、「場所 S から路線 L で場所 D まで行く」という発話で移動行為を提案したとする。この発話を実行された時点で、注視されているドメインの対象は S から D に移行する。また、文脈に導入済みの対象が提題化されると、その対象は注視される。発話の注視状態に対する適切さに関しては、例えば、「場所 S から路線 L で場所 D まで行く」という発話は、場所 S が注視されているなら適切であると考えられる。

5 問題解決

ここでは、問題解決モジュールの細部には立ち入らない。図3に示す路線図において、武蔵野センタから厚木センタまで最短時間で行くという問題を例にとると、まず、「武蔵野センタからバスで最寄駅まで行って、そこから厚木センタの最寄り駅に行って、そこからバスで厚

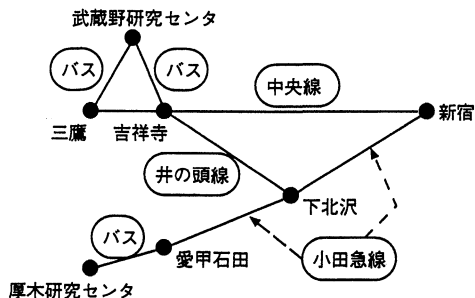


図 3: 路線図

木センタまで行く」という抽象的なドメインプランが立案される。さらに、「武蔵野センタから吉祥寺まで行って、そこから井の頭線で下北沢まで行って、...」というより具体的なドメインプランが立案される。

6 発話プランニング

発話プランとは情報伝達行為の列である。情報伝達行為として、「ドメインプランを提案する」、「ドメインプランを構成するドメイン行為を提案する」、「ドメイン行為の格要素を伝達する」、「言語表現を発話する」といった情報伝達行為をプランニング技法 [8, Chap.11] におけるオペレータとして定義する。また、第3.2節で述べた談話関係を利用した「行為の構成要素の存在に言及してからその行為を提案する」といった情報伝達行為も定義する [2]。

ドメインプランが与えられたとき、最初に立案される発話プランは、「そのドメインプランを提案する」という一つの情報伝達行為だけから成るプランである。この発話プランは、階層的プランニングの手法を用いて、ドメインプランを構成する各ドメイン行為を別々に提案するという発話プランに展開される。次に、ドメイン行為の格要素のそれぞれとドメイン行為のタイプを別々に伝達するという発話プランが立案される。ここまで情報単位が分解された時点で、発話プランに含まれる各情報伝達行為に対応する言語表現を合成し、その言語表現を発話するという情報伝達行為を立案し、発話制御モジュールに送る。この手法は、第3.1節で述べた発話単位の分析を反映している。

以上の手法によって、ドメインプランが最終的に決定する前に、ドメインプランの確定した部分から小さな発話単位を使って発話し、発話しながらドメインプランの内容を具体化することができる。結果として、時間制限を守りながら漸次的に発話を生成することが可能となる [2]。

発話の適切さを保証するために、言語運用制約を用いる。制約 (c1), (c3) は、言語表現の合成を行うときに適

用する。具体的には、(c1)に従って、名詞句の代名詞化、(c3)に従って、名詞句の提題化を行う。制約(c2)は、ドメインプランの具体化にもなって発話プランを再立案し、再立案された発話プランに基づいて発話を継続するときに適用する。具体的な適用例は、第7節で示す。

7 発話シミュレーション実験

モデルに基づいて実験システムを作成し、発話シミュレーション実験を行った。目的は、モデルの有効性を検証することである。実験は、言語運用制約を利用した場合と利用しなかった場合に分けて行った。いずれの場合も実験システムは、時間制限を守りながら発話を生成することができた。

次の発話例は、言語運用制約を利用した場合に実験システムが生成した発話を書き起こしたものである。

(発話例)

- (e1) 武蔵野センタからは バスです 最寄の駅に
- (e2) えーと 吉祥寺まで 出てください
- (e3) それから 井の頭線に乗って 下北沢まで行って
- (e4) それで 小田急線で 愛甲石田に行ってください
- (e5) それから バスなんです 森の里青山行きのバスがありますので それに乗って 通信研究所前まで 出ます

上の発話例の発話(e1)の時点で問題解決モジュールは、“武蔵野研究センタからバスで最寄りの駅まで行く”というドメインプランを立案している。発話(e1)の最中に問題解決が進行し、“武蔵野センタからバスで吉祥寺まで行く”というドメインプランが立案され、それにもなって発話プランが再立案され、結果として、(e2)が発話される。このようにドメインプランが部分的に決まった時点で発話を開始することにより、時間制限を守ることができる。(e2)が発話される時点で、言語運用制約(c2)が使われている。制約(c2)を使わないと、(e1)で発話された「武蔵野センタから」、「バスで」といった表現が(e2)でも繰り返されて、発話が不自然なほど冗長となる。この時点で吉祥寺から愛甲石田駅までの経路が決定されており、(e3)、(e4)が発話された。また、制約(c1)は、主として移動行為の源泉格をゼロ代名詞化することに用いられている。

発話(e2)が終了した時点で、吉祥寺駅と愛甲石田駅間の経路が決定されていない場合に、発話(e2)以降に次の発話(e3)’,(e4)’が生成された。

(発話例)’

- (e3)’ それで 小田急線の愛甲石田まで 行けばよいのですが
- (e4)’ えーと 吉祥寺からはですね 井の頭線で 下北沢まで 出てください ...

発話(e3)’は時間制限を守るために生成された。発話

(e3)’の終了時には愛甲石田駅が注視されている。ここで吉祥寺駅からの経路が詳細化され、発話(e4)’でその経路を詳細に説明し始めるが、制約(c3)により、吉祥寺駅からの経路を説明するためには、注視状態を愛甲石田駅から吉祥寺駅に戻す必要がある。したがって、吉祥寺駅を提題化した「吉祥寺駅からはですね」という表現がまず発話された。言語運用制約を適用しない場合には、冗長で注視状態に対して不適切な発話が生成された。

8 おわりに

本稿では、漸次的発話生成モデルを提案した。モデルは、時間制限の下で、発話内容を考えながら同時に発話することができる。さらに、対話コーパスに現れた談話構造を分析し、漸次的発話生成戦略に関与する談話構造を抽出した。そういった談話構造に基づいて、モデルは一連の発話を一つの談話として生成する。談話の適切さは、言語運用制約によって保証される。モデルの有効性は、発話シミュレーション実験によって確認されている。

謝辞: 日頃より御指導頂く石井健一郎情報科学研究部長、有益な示唆を頂く対話理解研究グループの諸氏に感謝致します。

参考文献

- [1] Carletta, J., Caley, R. and Isard, S. (1993). A system architecture for simulating time-constrained language production. Research Paper HCRC/RP-43, University of Edinburgh.
- [2] 堂坂浩二, 島津明 (1995). 実時間対話における漸次的発話のプランニングモデル. 信学技報 Vol.95, No.321, NLC95-40 (1995), 1-8.
- [3] Grosz, B. J. and Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, Vol. 12, 175-204.
- [4] Hovy, E. H. (1993). Automated discourse generation using discourse structure relations. *Artificial Intelligence*, Vol. 63, 341-385.
- [5] Levelt, W. J. M. (1989). *Speaking: From Intentions to Articulation*. The MIT Press.
- [6] Mann, W. C. and Thompson, S. A. (1988). Rhetorical structure theory: Towards a functional theory of text organization. *Text*, Vol. 8, 243-281.
- [7] Reithinger, N. (1991). POPEL - a parallel and incremental natural language generation. in: C. L. Paris et al. (eds.), *Natural Language Generation in Artificial Intelligence and Computational Linguistics*, Kluwer Academic Publishers, 179-199.
- [8] Russel, J. R. and Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Prentice Hall.
- [9] Walker, M. A. and Rambow, O. (1994). The role of cognitive modeling in achieving communicative intentions. in: *Proc. of the 7th International Conference on Natural Language Generation*.