

発話予測のための発話の動機に関する考察

Modeling Utterance Motivation for Predicting the User's Next Utterance

寺谷俊寛[†]

野村康雄[†]

山下洋一^{††}

溝口理一郎^{††}

Toshihiro TERATANI[†], Yasuo NOMURA[†], Yoichi YAMASHITA^{††}, and Riichiro MIZOGUCHI^{††}

[†]関西大学工学部

^{††}大阪大学産業科学研究所

[†]Faculty of Engineering, Kansai Univ.

^{††}I.S.I.R., Osaka Univ.

Abstract:

Prediction of user's next utterance is an important technique for understanding spoken dialog. Various kinds of dialog knowledge, such as utterance pairs, topic, and so on, are indispensable to the utterance prediction. We propose a new model of utterance motivation in order to enumerate possible topics in the next utterance. In our model, the utterance is motivated by two levels of motivation, the communication level and the problem solving level. The state of each piece of information transferred gives the motivation at the communication level. The purpose in the problem solving process invokes the motivation at the problem solving. We investigated the frequency of combination of two kinds of motivations using some simulated dialogs. A mechanism for predicting topics in the next utterance is described according to each utterance motivation.

1 はじめに

音声による対話は、人にとって最も親しみやすく、かつ自然なコミュニケーションの手段である。人と人の間だけでなく人と計算機などの機械との間でも、音声によるコミュニケーションを実現することを目指して、音声対話システムの研究が広く行なわれている。計算機と音声によってコミュニケーションを行なうには、まず人が発声した音声（以下、発話と呼ぶ）を計算機が正しく理解しなければならぬ。対話音声の認識・理解には、音声波形の持つ音響的な情報だけでなく、言語情報や対話情報などさまざまな高次知識が必要となる。特に、中規模／大規模語彙のタスクにおいては、対話知識を用いて発話における意味内容を制限する、いわゆる発話予測の技術が重要になってくる。

このような発話予測を行なう一つのアプローチとして、話題情報の利用が挙げられる。我々はこれまでに、個々の問題解決に依存しない汎用の対話管理システム MASCOTS(MAnagement System for COnversation using Twin-stacks and Sr-plan) について考察し、対話知識を利用した発話予測手法を提案し、その有効性を検証した [1]。この枠組では、発話対の種類ごとに用意された発話解析のテンプレートが、TPN(Topic Packet Network) と呼ばれる話題遷移モデル [2] に基づいて予測された話題によって具体化され、ユーザ発話の解析に用いられる。また川端は、話題を統計モデルにお

ける語彙の偏りとして表現する手法を提案し、確率文法におけるパープレキシティを削減し、探索空間の絞り込みが行なえることを示している [6]。

一般に対話における発話は、要求と応答の二種類に分類することができる。ユーザ発話予測の観点からみると、協調的な対話では応答発話の意味内容はそれに呼応し先行する要求発話から推測することができ、比較的容易に予測できる。また、応答発話における話題も要求発話と同一であると考えることができる。しかし、ユーザの要求発話ではこのような発話の対からは意味的制約を課すことが困難である。このため、我々は話題の遷移を表現するモデルとして TPN を導入した。TPN では、関連する話題を一つのバケット(話題バケット) に集め、さらにそれを相互に接続することによって、話題バケットを単位として話題の遷移可能範囲を設定していた。このため、遷移パターンが固定されており、対話の進行における情報の授受や問題解決の状態の変化に応じた柔軟な話題の予測ができない。

目的指向の対話は何らかの情報交換を行なうために進められ、やりとりを繰り返すことによってある目的が順に達成されていく。そのような対話では、発話を行なうのに「なぜ発話するのか」といった明確な動機を考慮することができ、それが発話の意味内容あるいは話題と密接に関連する。このような観点から、本報告では、次発話の話題としてありそうな話題を数え挙げ、すなわち「話題を予測する」ために、発話の動機

に注目したメカニズムについて考察する。

2 発話予測のための話題

2.1 話題とは

自然言語処理や音声理解/合成など様々な分野で話題を扱った研究が行なわれているが、「話題」が発話あるいは文におけるどのような概念を表すかは、必ずしも統一されていないように思われる。そこでまず、対話音声認識のための発話予測という観点から、「話題」について考察する。

一般に、発話の抽象概念としての話題の捉え方には、発話中のどの概念をどの抽象レベルで捉えるかにより様々なものが考えられる。竹下は、情報検索支援などを目的とした談話理解の観点から、対話遂行者によって共有される談話指示物として「話題」を扱っている[4]。そこでは、話題を名詞句あるいは動詞句によって表される事象として捉え、インタラクティブな発話では話題は遷移しないと仮定し表層的な処理によって話題の決定を試みている。一方音声認識では、認識対象となる語彙空間を削減するために話題の利用が検討され、発話の意味内容を抽象的に表現するものとして「話題」が用いられる[3]。

我々は、ある問題を解決することを目的として協調的に行なわれる対話を対象とするため、要求に対する応答、確認のための繰り返し要求においては話題は遷移しないものと仮定し、要求文に対して話題を用いた意味内容の予測を図る。

ここで、目的指向対話は、目的とする情報の授受を行なうために関連する情報を授受するものと捉えることができる[5]。要求文に含まれる情報について考えると、例えば図1に示すように、ここでは話者が必要とする情報を直接表す概念『移動法』の他に、その欲する情報を特定するための概念『始点』『着点』を提示する必要がある。このような概念は要求している情報の持つ属性概念などにあたり、ある目的を実行するために必要な情報であると考えられる。どの概念を話題と捉えるかは、(1) 表層発話からの話題決定、(2) 文脈情報を利用した次話題の予測、(3) 話題からの表層発話の予測などの観点から有効なものを選択する必要がある。

図1の例のように、一般に要求文におけるその言い回しは要求する情報の概念と非常に密接な関連を持つ。例えば、『移動法』を要求する表現として、図1に示すような表現を容易に予測できるのに対し、『移動法』を提示して要求を行なう表現を考えると「電車な

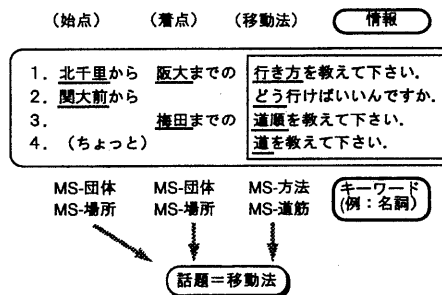


図1: 要求発話における話題

ら、どの駅で降りるのですか?」「JRで行くと学割は使えますか?」etc. など様々である。そこで、要求文における話題を話者の要求している情報の上位概念として捉えることにすれば、要求発話を構文との依存関係を重視した意味内容によって分類することが可能である。このような話題は、表層発話との関係付けに適している。

また、話者が要求文において直接必要としている概念は、話者のプランに基づいて展開されるため前発話までの話題から次話題を予測する際にも都合が良い。

このような点から、我々は要求文における話題を話者の要求している情報の上位概念として捉える。

2.2 対話における話題の展開

目的指向の対話では、そこに出現する話題の数はある程度限定され、話題を対話の領域に依存する知識として整理することが可能である。「ある問題解決を行なう」という大きな目的の下にそれを達成するためのいくつかの話題が存在し、さらにそれらの話題について詳細なやりとりを行なうための下位話題が存在するというように、対話全体を話題の階層的な展開として表現することができる。このような話題の展開をモデル化し、次発話の意味内容のトップダウンな予測を行なう。

3 発話の動機モデル

次発話をどのような動機で発話しようとしているのか知ることができれば、次発話においてどのような内容が発話されるのか予測することができる。ここでは、模擬対話の解析の結果から得た二つのレベルの動機について説明し、発話の動機をモデル化する。

3.1 模擬対話の解析

目的指向対話の発話文はどのような動機から発話されるのだろうか。地理案内およびスキーツアー予約の模擬対話を用いて発話の動機を解析した。前節で述べたように、応答発話での話題はそれと対をなす要求発話における話題と同一であると仮定していることから、対話中の要求発話のみを対象とした。

模擬対話における要求文は、

- 値が不明で、それが必要である
- 値が複数あり、それらを比較したい
- 値が複数あり、一つを選択したい
- ...

などの動機から発話されており、これらは、

- 授受された情報の状態に関する動機
- 授受された情報を使った思考内での情報操作に関する動機

の二つの動機が組合わさって一つの動機を形成していることがわかった。(図2)。ここで、授受された情報の状態によって決まる発話の動機を情報伝達レベルの動機と呼ぶ。また、授受された情報を使った思考内での情報操作によって決まる発話の動機を問題解決レベルの動機と呼ぶ。

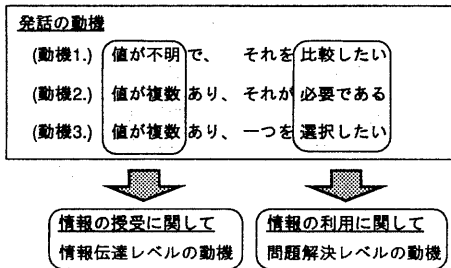


図2: 発話の動機の構成

なお、発話には情報を求める発話と情報を提示する発話の二つが存在するが、発話の動機に関しては情報の伝達方向を含めておらず、各動機について二つの情報伝達の方向を考えることができる。

3.2 授受される情報

何らかの問題を解決しようとする目的指向の対話では、その問題解決に必要な情報が授受され、各発話によって授受された情報は、その発話における話題と考えることができる。

各発話の意味内容は、問題解決の状態とその領域での情報の関連性から決定される。従って、対話管理部

において情報伝達の状態を把握し、次発話を予測するためには、個々の対話の領域での情報の関連性を考え、その領域での知識構造をモデル化しておくことが重要である。そこで、このような関連性をもったまとまりを情報パケットと呼び、模擬対話の分析から次の二種類に分類する。

(1) 行為列

人間が行なう連続した行為によってある目的が達成される時、それらの行為をまとめて一つの行為列として捉える。即ち行為列は、目的を達成する上で必要となる行為の時系列として定義される。各行為には、その行為を実行する際に必要な情報が属性として持つ。

(2) オブジェクト

行為以外の“物”についての情報は一つのオブジェクトとして捉える。オブジェクトには、その“物”を規定する情報を属性として持つ。

行為列、オブジェクトの情報パケットの例をそれぞれ図3に示す。

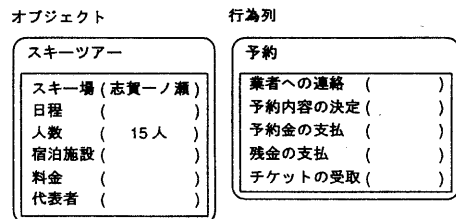


図3: 情報パケット

各パケットの属性には、対話において授受された情報を記述していくスロットあり、実際に授受された情報を把握することができる。

3.3 情報伝達レベルの動機

対話を情報の授受と考えると、授受された情報の状態という視点から発話の動機を考えることができる。情報伝達レベルの動機は、話者の思考内での情報の状態によって決まる。

例えば、同じ「人数」を授受する発話であっても(すなわち話題が人数)、話者の思考内の情報の状態によって、発話の動機は、1) 人数に関する情報がない、2) 人数に関する情報「多く」は持っているが、情報が曖昧だから、など、いくつかの動機が考えられる。

以下に情報伝達レベルの動機を示す。

- a. 値が未知だから
属性値の授受がまだ行なわれていない

- b. 値が曖昧だから
授受された属性値に曖昧さが含まれている
- c. 値が複数だから
複数の情報が授受された
- d. 属性/行為が未知だから
どのような属性値を授受すればいいかわからない
- e. 属性/行為が網羅されたから
あるパケットに関して情報の授受が完了した
- f. オブジェクト/行為列が曖昧だから
複数のパケットについての情報の授受が行なわれ、現在どのパケットについて情報の授受を行なっているかわからない
- g. 確認したい
- h. オブジェクト/行為列が未知だから
どのパケットについて情報の授受を行なうのかわからない

3.4 問題解決レベルの動機

情報伝達レベルの動機と同様に、目的指向対話を情報の授受と考えると、何をするために発話を行なうかという視点から発話の動機を考えることができる。問題解決レベルの動機は、話者の思考内でどのように情報の利用を行うかによって決まる。

例えば、何かの事柄を決めるために、その事柄を決定するために必要条件やその事柄に関連する情報を収集したい、また、いくつかある情報の差異を知りたいなどが問題解決レベルの動機である。

以下に問題解決レベルの動機を示す。

- A. 情報を比較したい／情報を比較してほしい
ある事柄に関するいくつかの情報ある場合、その情報の差異をとりたい
- B. 選択したい／選択してほしい
複数ある情報の中からいくつかを選びたい
- C. 順序を決めたい／授受を決めてほしい
複数ある情報に優先順位をつけたい
- D. 理由を知りたい／理由を伝えたい
何かをする場合その理由を知りたい
- E. 必要条件を知りたい／必要条件を伝えたい
何かをする上での必要な条件を知りたい
- F. 関連情報を知りたい／関連情報を伝えたい
何かをする上での関連する情報を知りたい
- G. 訂正したい／訂正して欲しい
間違った情報の授受を正しく直したい
- H. 制約を満たしたい／制約を満たしてほしい
制約条件を満たす範囲内に情報をおさめたい

- I. 目的を知りたい／目的を伝えたい
情報の授受の目的を知りたい
- J. 完了を知りたい／完了を伝えたい
情報の授受の完了を知りたい

3.5 動機の組合せ

我々は、情報伝達レベルの動機と問題解決レベルの動機が組合わさって一つの発話の動機を形成すると考える。実際の対話ではこれらの動機にどのような組合せが起こるのか、模擬対話を用いて検討を行なった。同時に、タスクによって発話の動機にどのような違いが生じるのかを調べるために、道案内タスクおよび予約タスクでそれぞれデータを収集し、検討した。

タスク	対話ID	要求発話数
地理案内	osa0008d.	16
	osa0009d.	27
	osa0010d.	84
スキー予約	tut0004d.	34
	tut0005d.	61
	tut0006d.	30
合計		252

表 1: 用いた対話例

		問題解決レベルの動機										
		A	B	C	D	E	F	G	H	I	J	nil
情報伝達レベルの動機	a					56	57		1			
	b	2	1		3	3	3					
	c	6	5	1								
	d	1				6	2					
	e										6	
	f					2						
	g											83
	h									7		
	nil						5	2				

(計 252発話)

表 2: 各発話の動機の頻度 (地理案内 + スキー予約)

表中各レベルの動機における nil は、発話にその動機が含まれないことを表す。確認の発話では、情報の授受自体が問題にされ、それが問題解決に使われることはないため、情報伝達レベルの動機 g. 「確認したい」の場合には、問題解決レベルの動機はない。逆に、訂正を行なう発話では、授受された情報の状態を参照するだけでは、その情報が間違った情報なのか正しい情報なのか判断できないため、情報伝達レベルの動機を nil としている。

地理案内、スキー予約を合わせた総合の結果を見ると、情報伝達レベルと問題解決レベルの組合せの全てが存在するのではないことがわかる。また、オブジェクト/行為列に関する情報伝達レベルの動機 f. と動機

		問題解決レベルの動機										
		A	B	C	D	E	F	G	H	I	J	nil
情報伝達レベルの動機	a					35	27					
	b			1		1	1					
	c	1	2	1								
	d	1				4						
	e											
	f					1						
	g											48
	h									4		
	nil											

(a) 地理案内 (計 127発話)

		問題解決レベルの動機										
		A	B	C	D	E	F	G	H	I	J	nil
情報伝達レベルの動機	a					21	30		1			
	b	2			3	2	2					
	c	5	3									
	d					2	2					
	e										6	
	f				1							
	g											35
	h									3		
	nil						5	2				

(b) スキー予約 (計 125発話)

表 3: 各発話の動機の頻度 (タスク別)

h. の発話が少ないのは、これらの模擬対話で参照される情報パケットの数が少ないためと考えられ、タスクの規模が大きくなると、このような動機の頻度も増加すると思われる。

タスク別の結果をみると、両タスク間で出現する動機の組合せに大きな違いがないことがわかる。

4 動機と話題の遷移

ここでは、表 2 に現れた動機の組合せのそれぞれに対して、それらの動機の組合せと話題の遷移の関係について検討し、発話の動機を用いた話題の予測手法について述べる。

4.1 遷移関係

次発話の動機がわかれば、次発話の話題がどんな話題に遷移するのがある程度予測できる。そこで発話の動機 (情報伝達レベルの動機と問題解決レベルの動機の組合せ) ごとに話題の予測パターンを考察した。

以下にいくつかの発話動機における話題の予測パターンを示す。

(c. 値が複数あるから
A. 情報を比較したい)

次発話の動機としてこの動機が予測されるには、直前の発話までにあるスロットの値が授受され、そのスロットに複数の値が埋められていなければならない。そして、この値に関して情報を比較したいのであるから、次発話ではこの情報を新たなオブジェクトまたは行為列としてとらえた場合の属性/行為に関しての情報の授受が行なわれる。すなわち、次発話の話題は複数の値が入ったスロットを名前として持つオブジェクトまたは行為列の属性/行為に遷移する。

(e. 行為/属性が網羅されたから
J. 完了を知りたい)

次発話の動機としてこの動機が予測されるには、現在の発話までにある情報パケットの全てのスロットが埋められていなければならない。そして、ある状況での情報の授受が完了しているため、次発話では新しい状況への情報の授受が行なわれる。すなわち、次発話の話題は、現在の情報パケットとは異なるパケット中の話題へと遷移する。

(h. オブジェクト/行為列が未知だから
I. 目的を知りたい)

次発話の動機としてこの動機が予測されるには、動機 (e. J.) 同様に、現在の発話までにある情報パケットの全てのスロットが埋められているか、または対話の始めて全ての情報パケットのスロットが空でなければならない。話題も、現在の情報パケットとは異なるパケット中の話題へと遷移する。さらに、この発話の後では、この発話によって導入された情報パケットに関して情報の授受が行なわれる。

(g. 確認したい
nil)

この動機は、他の動機と異なり、情報伝達レベルの動機のみから構成される。この動機は、全ての情報の授受の後に考えることができる。確認の発話は、ある情報の授受の直後にその情報について、直前の発話と同じ発話が繰り返される。すなわち、次発話の話題は、直前の授受された情報と同じである。

4.2 予測手法

発話の動機に基づく話題の予測手法は、情報パケットのスロットの状態、および発話の動機（動機の組合せ）ごとに決めた話題の予測パターンを参照することにより次発話の話題を発話ごとに動的に予測していく。話題予測の手順は以下のとおりである。

1. 授受された情報の状態から情報伝達レベルの動機を決定する。（情報の状態は情報パケットのスロットの状態を把握）
2. 問題解決の状態から問題解決レベルの動機を決定する。
3. 次発話の動機を情報伝達レベルの動機と問題解決レベルの動機の組合せより求める。
4. 求められた発話の動機における話題の予測パターンを用いて、次発話で可能な話題を動的に決定する。

4.3 予測例

図4に示した2.の発話が行なわれた後で、スキーツアーオブジェクトのスロットが、[スキー場(志賀-ノ瀬)、日程(), 人数(15人)、宿泊施設(一井ホテル、ホテル山楽)、料金(), 代表者()]である場合を考える。

情報伝達レベルの動機を考えると、日程、料金、代表者の各スロットが空であることから「(動機 a.) 値が未知だから」、宿泊施設のスロットに値が複数入っているために「(動機 c.) 値が複数あるから」、また全ての発話の後で考えられる「(動機 g.) 確認したい」の3つが次発話動機として考えられる。

話者が今のような問題解決の状態にいるのかわかり、問題解決レベルの動機として「A. 情報を比較したい」が考えられた場合、動機の組合せから次発話の動機は、『(c. A.) 値が複数あるから、情報を比較したい』となる。

この時、動機(c. A.)における話題の予測パターンより、次発話の話題としては、複数の値が入っているスロットを名前としてつオブジェクト、宿泊施設オブジェクトの各属性、「収容人数、グレンデまでの距離、場所」などが予測される。

5 おわりに

本稿では、目的指向対話における発話の動機について考察した。発話の動機は情報伝達レベルの動機と問題解決レベルの動機から構成される。情報伝達レベルの動機は授受された情報の状態により決まり、問題解決レベルの動機は話者が問題解決のどのような状態に

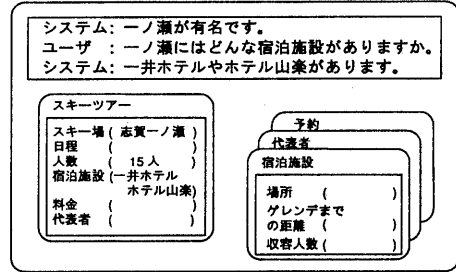


図 4: 予測例

いるかにより決まる。また、発話の動機を発話予測に利用する予測メカニズムを提案した。

今後は、発話の動機を用いた発話予測メカニズムについて、有効性を検証していく予定である。

謝辞

本研究の一部は、科研費重点領域研究「音声対話」(No.05241105)の支援を受けた。また、本研究では、日本音響学会の研究用連続音声データベースを利用した。

参考文献

- [1] Y. Yamashita, et al: "Next Utterance Prediction Based on Two Kinds of Dialog Models", Proc. of Eurospeech '93, Berlin, pp.1161-1164 (1993).
- [2] Y. Yamashita, et al: "MASCOTS II: A Dialog Manager in General Interface for Speech Input and Output", IEICE Trans., 76-D, 1, pp.74-83 (1993).
- [3] 堀他: "音声理解システムのための連想パーザ ASP-認知科学的検討に基づくフレームワークの設計-", 電子情報通信学会論文誌, J71-D, 5, pp.774-781 (1988).
- [4] 竹下 敦: "音声対話のインタラクション構造を用いた話題の認識", 情処学会研究会資料, NL-87, pp.75-82 (1992).
- [5] 柏岡他: "情報の授受に着目した対話の構造解析", 情処学会研究会資料, NL-76, pp.1-8 (1990).
- [6] 川端豪: "確率文法と話題マルコフモデルに基づく音声認識のための話題制御", 信学技報 SP93-115, NLC93-55, pp.7-13 (1993).