

# 試行型対話戦略による音声対話システムの確認発話削減

堂坂浩二 安田宜仁 相川清明

日本電信電話株式会社, NTT コミュニケーション科学基礎研究所

## 1 はじめに

音声対話システムは、ユーザとの音声対話を通して、情報案内、各種の予約といった与えられたタスクを遂行するシステムである。本稿では、ユーザはシステムが保有するデータベース内容についての問い合わせを行い、システムはユーザが必要とする情報を伝達するという音声対話システムを考える。音声対話システムは、音声認識誤りに対処するために、ユーザ問い合わせ内容を認識した結果をユーザに対して確認し、ユーザからの承認発話（例：「はい」）を待って、問い合わせ内容を確定するという一連のやり取り（確認対話）を実施することが普通である。確認対話は、音声認識誤りの下でタスクを遂行するために有効な手段であるが、過度な確認は対話の円滑さを損ねる。したがって、 unnecessary 確認は避けることが望ましい。

システムが確認によって確定した問い合わせ内容だけを正しいとして、ユーザに情報を伝達するという対話戦略を確定型の対話戦略と呼び、確定済みの内容だけでなく、未確定の問い合わせ内容も正しいと仮定して、情報伝達を行うという対話戦略を試行型の対話戦略と呼ぶこととする。試行型の対話戦略をとる場合には、システムは誤認識のため誤った情報伝達を行う危険がある。しかし、ユーザがシステムの誤認識に気づき、適切な情報を手に入れるまで対話を継続すると仮定するならば、いったんは誤った応答が生成されたとしても、何回か対話を繰り返しているうちに、ユーザが必要な情報を得る確率は向上する。音声認識精度が高い状況においては、試行型の対話戦略が確定型の対話戦略よりも平均的に短い対話でユーザに必要な情報を伝達できる場合があるはずである。本稿では、ユーザが必要とする情報をできるだけ短い対話で伝達するという目的の下で、試行型の対話戦略と確定型の対話戦略を適切に使い分けるための対話制御に注目する。

従来法として、音声認識結果の信頼度が閾値を超えているならば、確認を行わないことを主旨とする対話制御法が提案されている [4]。この従来法は、理解誤りを最小にするという目的の下で確認数を削減するための対話制御法と見なすことができる。一方、本稿では、対話の長さを最短にするという目的の下での対話

制御に注目している。対話の長さを最短にするための対話制御法としてデュアルコスト法を提案してきた [2]。デュアルコスト法は、確認対話の長さ（確認コスト）と確認対話後にユーザに情報を伝達するために生成されるシステム応答の長さ（情報伝達コスト）の和を対話コストととらえ、対話コストを最小化するように対話を制御する。確認コストは音声認識精度に依存し、情報伝達コストは、システムのデータベースの内容に依存する。従来の対話制御法 [1, 5, 6] は、確認対話の長さのみを削減することに注目しており、ユーザに必要な情報を伝達するためのシステム応答を含む対話全体の長さを削減することは考慮されていない。これらの従来法が避けることができない無駄な確認をデュアルコスト法は避けることができる。

デュアルコスト法は、確定型の対話戦略のみを考慮していたが、さらに、デュアルコスト法を発展させ、試行型の対話戦略も考慮する試行型デュアルコスト法を提案してきた [3]。しかし、そこでは、システムが未確定の問い合わせ内容に基づいて応答を行い、応答が誤っていた場合には、ユーザからの拒否発話（例：「ちがいます」）があることを前提としていた。拒否発話があった場合には、確定型の対話戦略にしたがって対話をやり直す。しかし、応答が誤っていた場合にはユーザからの拒否発話があるという前提は、必ずしも自然なものではない。本稿では、ユーザからの拒否発話があるという前提を置かず、試行型の対話戦略も考慮して、できるだけ短い対話でユーザに必要な情報を伝達するための対話制御法を提案する。以下においては、本稿において提案する方法を単に試行型デュアルコスト法、従来提案してきた手法 [3] を旧試行型デュアルコスト法と呼ぶ。

## 2 音声対話システムにおける対話制御

システムとユーザの対話は、ユーザ問い合わせ確定フェーズとシステム情報伝達フェーズという2つのフェーズを移行しながら進行する。ユーザ問い合わせ確定フェーズでは、確認対話を通してユーザ問い合わせ内容が確定されていく。対話の各時点でシステムが認識している問い合わせ内容はシステム理解状態と呼ばれ、3つ組  $\langle$  属性, 値, 確定フラグ  $\rangle$  の集合とし

て保持される。問い合わせタイプの種類と属性の全体集合はタスクごとに決まっている。確定フラグは、属性の値がユーザの承認発話により確定するまで「未」という値をとり、確定されると「済」という値をとる。

本稿では例として気象情報案内システムを想定する。警報問い合わせ、天気問い合わせ、気温問い合わせ、降水確率問い合わせの4種類の問い合わせタイプが可能である。場所、日、警報種別、情報種別の4つの属性がある。場所属性は特定の場所の名前を値としてとる。日属性は今日か明日という値をとる。警報問い合わせでは、日属性は無効な属性である。警報種別属性は、大雨、洪水といった値をとる。警報以外の問い合わせでは、警報種別属性は無効な属性である。情報種別属性は、問い合わせタイプに応じて、警報、天気、気温、降水確率という値をとる。

ユーザ問い合わせ確定フェーズでは、システムは確認行動、情報要求行動のいずれかの行動を行う。確認行動は、システム理解状態において値が与えられている属性について、値が正しいかどうかユーザに対して確認するための確認発話を行い、「はい」といったユーザの承認発話によって値が確定されるまで、その属性の値の確認を繰り返すという行動である。ユーザは、システムの確認内容が誤っている場合には、システム理解状態を訂正するための訂正発話を行うことができる。情報要求行動とは、システム理解状態において値が与えられていない属性について、ユーザから属性の値を引き出すための情報要求発話を行い、その後、その属性についての確認行動を実施するという行動である。システム情報伝達フェーズでは、システムはユーザが必要とする情報を伝達するためにシステム応答を生成するという情報伝達行動を行う。対話制御とは、対話の各時点において、短い対話でユーザの必要とする情報を伝達するという観点から最適なシステム行動を決定することである。

システムは、情報伝達行動を行なうとき、いくつかの属性値を正しいと仮定した上で応答文を生成する。それらの属性値を情報伝達行動や応答文の基礎となる属性値と呼ぶ。たとえば、本稿で想定している気象情報案内システムでは、神奈川県の子の天気伝達するシステム応答の基礎となる属性値とは、場所属性の値（神奈川県）、日属性の値（明日）、情報種別属性の値（天気）となる。情報伝達行動やシステム応答が確定型であるとは、その基礎となる属性値が確定済みの属性値ばかりであることを言う。情報伝達行動やシステム応答が試行型であるとは、その基礎となる属性値の中に未確定の値が含まれるときを言う。

### 3 試行型デュアルコスト法

本稿で提案する試行型デュアルコスト法は、デュアルコスト法と同様に、対話の各時点において以下のように動作する。

(Step1) 現時点のシステム理解状態において可能なユーザ問い合わせタイプを導き出し、ユーザ問い合わせタイプの確率分布を計算する。

(Step2) ユーザ問い合わせタイプごとに、可能な対話プランをすべて生成する。対話プランとは、対話手順を表したシステム行動の列である。各対話プランを実施したときの対話の長さの期待値を確認コストと情報伝達コストの総和として計算する。この総和を対話プランのコストと呼ぶ。

(Step3) 現時点で実行可能なシステム行動をすべて生成する。ユーザ問い合わせタイプごとに、システム行動を含む対話プランの中で最少のコストをもつ対話プランを探索する。そのコストをシステム行動のコストと呼ぶ。

(Step4) ユーザ問い合わせタイプの確率分布に基づいて、各システム行動のコストの期待値を計算し、コストの期待値が最小となるシステム行動を選択する。

(Step5) 選択されたシステム行動を実行する。

対話プランとしては、確定済みの問い合わせ内容のみが正しいとして、応答を生成する確定型の対話プランだけでなく、未確定の問い合わせ内容も正しいと仮定して、応答を生成する試行型の対話プランを考慮する。試行型対話プランには試行型の情報伝達行動が含まれている。

確定型の対話プランの対話コストは、デュアルコスト法と同様に計算すればよい。試行型の対話プランの対話コストについて説明する。試行型の対話プランであっても、確認行動、情報要求行動の確認コストは、確定型の対話プランと同様に計算できる。属性の集合が与えられるとき、すべての属性の値を確定するための確認行動の確認コストは、その属性をすべて正しく認識する属性認識精度を  $r$ 、確認すべき属性の数を  $m$  個とすると、次の式で定義される [2]。

$$\frac{2m}{r} - m + 1 \quad (1)$$

すべての属性の値を確定するための情報要求行動の確認コストは、次の式で計算される [2]。

$$\frac{2m}{r} + m + 1 \quad (2)$$

次に、試行型対話プランに含まれる試行型の情報伝達行動のコストについて説明する。デュアルコスト法では、確認対話の長さは確認対話の中でシステムとユーザの間でやり取りされる自立語の数、システム応答の長さは応答に含まれる自立語の数と定義される [2]。この定義に沿って、試行型の情報伝達行動のコストは、その行動によって引き起こされる対話の中に含まれる自立語の数の期待値であると定義する。

試行型の情報伝達行動の基礎となる属性値の中には、未確定の属性値が含まれている。それらの未確定の属性値が正しい確率を  $p$  とする。システムが確率  $p$  で正しい応答を行なう場合、ユーザは必要な情報を得ることになる。試行型の情報伝達行動によって生成されるシステム応答に含まれる自立語の数の期待値を  $Inf_1$  とする。

システムが確率  $1-p$  で誤った応答を行なう場合、システムは  $Inf_1$  個の自立語を含む誤った応答を行うことになる。この後、ユーザは必要な情報を得るまで、現在と同じ内容の問い合わせを繰り返すと仮定する。現在考えている問い合わせタイプの対話について、確認対話の長さの平均値は  $Conf_2$ 、システム応答の長さの平均値は  $Inf_2$ 、情報伝達行動の基礎となる属性値が正しい確率は  $q$  であるとする。これらの値は、システム-ユーザの対話を繰り返すことによって獲得するものとする。ユーザが必要な情報を得るまで、同じ問い合わせタイプの対話が  $\frac{1}{q}$  回繰り返され、各対話の長さの平均は  $Conf_2 + Inf_2$  となる。

結局、試行型の対話プランにおいて、情報伝達コストが引き起こす対話の長さの期待値は次の式で与えられる。

$$pInf_1 + (1-p)(Inf_1 + \frac{1}{q}(Conf_2 + Inf_2)) = Inf_1 + \frac{1-p}{q}(Conf_2 + Inf_2) \quad (3)$$

試行型対話プランのコストは、プランに含まれる確認行動のコスト (1)、情報要求行動のコスト (2)、試行型の情報伝達行動のコスト (3) の和となる。

#### 4 シミュレーション対話実験

システムと模擬ユーザとの間のシミュレーション対話実験 [2] によって、本稿で提案した試行型デュアルコスト法の評価を行った。模擬ユーザとは、実ユーザの

振舞をシミュレートするエージェントである。システムと模擬ユーザは、音声で対話するのではなく、発話内容を属性と値の対のリストとして表現した上で、属性と値の対のリストをやり取りすることによって対話を行う。ユーザの発話内容をシステムに送るときには、属性認識精度に応じて属性値に誤りが含まれるように、システムの音声認識誤りをシミュレートした。また、未確定の属性の値が正しい確率は属性認識精度と等しいとした。

実験で用いたタスクは気象情報案内タスクである。天気、気温、降水確率、警報の4つの情報種別をもち、情報種別に対応して4つの問い合わせタイプがある。場所は50個の都市、日は今日か明日の2通りである。警報種別としては、洪水、大雨など10個の種別がある。システムは、データベースの中に、各都市の今日、明日の天気、最高気温と最低気温、6時間ごとの降水確率のデータを保持している。また、警報については、現在どこにも警報は発表されていないということデータを保持している。

各対話の冒頭で模擬ユーザは問い合わせ内容が与えられる。対話が始されると、まず最初に、模擬ユーザは、与えられた問い合わせ内容の全て或いは一部を無作為に選んで、システムに伝達する。このとき、擬似的な音声認識誤りが発生し、システムは、ユーザが伝達した内容の一部を誤って受け取る。この後、システムとユーザは確認対話を行い、システムがユーザの問い合わせにしたがって応答が可能であると判断したとき、システム応答が生成される [2]。

試行型デュアルコスト法と比較するために、旧試行型デュアルコスト法 [3]、デュアルコスト法 [2]、できるだけ多くの属性を一度に確定しようとする一括確認法、属性を一つずつ確定する個別確認法、全く確認を行わない無確認法を用いた。デュアルコスト法、一括確認法、個別確認法は、確定型の対話プランのみを考慮する。一括確認法と個別確認法は、音声認識精度、データベース内容にかかわらず、すべての属性を確認する。

試行型デュアルコスト法を動作させるためには、式 (3) に示したように、誤ったシステム応答に後続する対話について、対話の長さの平均値  $Conf_2 + Inf_2$ 、システム応答の基礎となる属性値が正しい確率  $q$  の値が必要となる。これらの値としては、試行型デュアルコスト法のシミュレーション対話実験を事前に何度か繰り返し、収束した値を用いた。

シミュレーション実験では、各属性の属性認識精度を等しく 0.7 から 1.0 まで 0.05 刻みで変化させていった。各認識精度において 1000 回のシミュレーション対話が実施された。各認識精度において問い合わせ

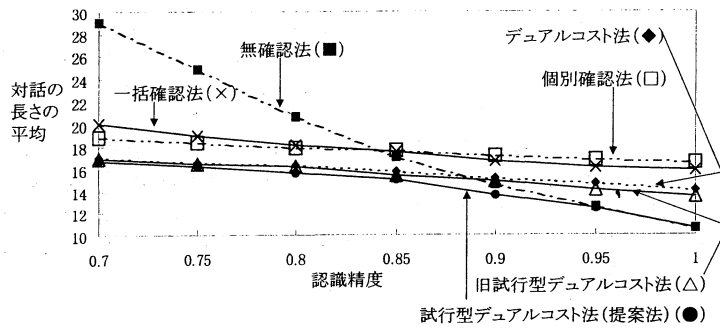


図 1: ユーザ問い合わせタイプを無作為に発生させたときの属性認識率に応じた対話の長さの平均

タイプの出現頻度が均等になるように、各タイプの問い合わせを無作為に発生させ、模擬ユーザに与えた。各方法の評価は、システムから模擬ユーザにユーザ問い合わせに合致した情報が伝達されるまでの対話の長さによって行った。対話の長さはシステムと模擬ユーザがやり取りする自立語の数とした。

評価結果を図 1 に示す。横軸は属性認識精度、縦軸はユーザ問い合わせに合致した情報が伝達されるまでの対話の長さの平均値である。提案法の試行型デュアルコスト法は、いずれの認識精度においても、他の従来法に比べて短い対話でユーザに必要な情報を伝達できている。認識精度が 1 に近いときは、試行型デュアルコスト法は結果として確認を行わない戦略をとることになるので、対話の長さは試行型デュアルコスト法と無確認法とでほぼ変わらない。しかし、無確認法は認識精度が低くなると、対話の長さが極端に長くなるが、試行型デュアルコスト法は、認識精度が低くなくても、常に他の方法よりも短い対話で必要な情報を伝達できることが分かる。

旧試行型デュアルコスト法 [3] は、認識精度が 1 に近いとき、確定済みの問い合わせ内容のみにしたがって応答するデュアルコスト法よりも、短い対話で問い合わせに応えることができるが、無確認法に及ばない。旧試行型デュアルコスト法は、システムの誤った応答に対してユーザからの拒否発話があることを前提としており、拒否発話を受け取った後は、確定済みの問い合わせ内容のみにしたがって応答する。本稿で提案した試行型デュアルコストは、この前提を取り除くことによって、旧試行型デュアルコスト法、無確認法に優る効果を得ることが分かる。

## 5 おわりに

本稿では、デュアルコスト法 [2] を発展させ、ユーザに対する確認によって確定済みとなった問い合わせ内容

だけでなく、確定済みでない問い合わせ内容も正しいと仮定した上で情報を伝達することを可能とする試行型デュアルコスト法を提案した。シミュレーション対話実験により、試行型デュアルコスト法が従来法よりも短い対話でユーザに情報を伝達できることを示した。

謝辞: 日頃よりご指導いただくメディア情報研究部 村瀬洋部長、熱心に討論して下さるマルチモーダル対話研究グループの諸氏に感謝致します。

## 参考文献

- [1] Jennifer Chu-Carroll. Mimic: an adaptive mixed initiative spoken dialogue system for information queries. In *Proc. of ANLP-2000*, pp. 97-104, 2000.
- [2] 堂坂浩二, 安田宜仁, 相川清明. システム知識制限下での効率的音声対話制御, 自然言語処理, Vol.9, No.1, pp.43-63 (2002).
- [3] 堂坂浩二, 安田宜仁, 相川清明. 情報伝達効率最大化に基づく音声対話制御法, 言語処理学会第 8 回年次大会発表論文集, pp.260-263 (2002).
- [4] 駒谷和範, 河原達也. 混合主導対話における音声認識誤りに対処するための対話管理, 言語処理学会第 6 回年次大会発表論文集, 336-339 (2000).
- [5] Yasuhisa Niimi and Yutaka Kobayashi. Dialog control strategy based on the reliability of speech recognition. In: *Proc. ICSLP*, 1996.
- [6] Satinder Singh, Michael S. Kearns, Diane J. Litman, and Marilyn A. Walker. Empirical evaluation of a reinforcement learning spoken dialogue system. In *Proc. AAAI-2000*, pp. 645-651, 2000.