

確認対話の制御方式の効率と音声認識システムの性能との関係

新美 康永 西本 卓也 荒木 雅弘

京都工芸繊維大学 工芸学部 電子情報工学科

〒606-0962 京都市左京区松ヶ崎御所海道町

<http://www-vox.dj.kit.ac.jp> {niimi,nishi,araki}@dj.kit.ac.jp

あらまし 本稿では、確認対話の制御方式の効率と音声認識システムの性能との間の定量的な関係を報告する。対話の目標が、あるタスクを達成するために n 個の情報項目を獲得し、これを確認することであるとすると、対話の状態は3字組 (u, k, c) ($u+k+c=n$) によって表現される。但し、 u はシステムにとって未知の項目の数、 k は既知であるが確認されていない項目の数、 c は確認された項目の数である。このとき対話の状態変化は1つのマルコフ過程によって表現される。このマルコフ過程の状態空間は上述の3字組の集合であり、状態間の遷移確率は各情報項目の認識精度によって定まる。対話制御方式の効率はこの事実に基づいて計算される。

キーワード 音声対話システム、確認対話、対話制御方式の効率、音声認識システムの性能、情報項目

Relations between the Efficiency of Control Strategies for Confirmation Dialogue and the Performance of a Speech Recognizer

Yasuhisa NIIMI, Takuya NISHIMOTO and Masahiro ARAKI

Department of Electronics and Information Science, Kyoto Institute of Technology

Matsugasaki, Sakyo-ku, Kyoto, 606-0962 Japan

<http://www-vox.dj.kit.ac.jp> {niimi, nishi, araki}@dj.kit.ac.jp

Abstract This paper describes relations between the efficiency of dialogue control strategies and the performance of a speech recognizer. If the goal of a dialogue is to collect and confirm several items, a status of the dialogue is represented by a triplet of the three numbers, that is, numbers of unknown, known and confirmed items. The efficiencies of the strategies can be derived from the fact that a dialogue can be represented by a Markov process in which a state space is a set of such triplets and state transition probabilities are decided by a speech recognition rate.

Key words Spoken dialogue system, Confirmation dialogue, Efficiency of dialogue control strategies, Performance of a speech recognizer, Information item

1 はじめに

音声認識技術の成熟に伴い、マン・マシン・インタフェースの手段として音声対話システムの研究も多行われるようになった。しかし、自発的な発話では、発声が不明瞭であったり、書き言葉の文法から逸脱したような入力も多くなる。従って、音声認識の誤りを完全になくすることは困難であろう。これを克服するために、種々の対話制御方式が研究されているが [1,2]、いずれも実験的であり、理論的な解析は行われていない。

我々はこれまで、音声認識結果の信頼性が評価できるものと仮定して、信頼性の高い場合のみ認識結果を受理し、信頼性の低い場合は、確認を行ったり、再入力を求める対話制御方式の解析を行ってきた [3,4]。しかし、音声認識の結果の信頼性を発話内容に関わりなく一様に評価することは困難である。そこで、本稿では、認識結果は常に何らかの確認を行うものと仮定して、4種類の対話制御方式について、それらの効率と音声認識システムの性能との関係を解析した。

2. では、本稿で扱った4種類の確認対話の制御方式について概説し、3. では、対話制御方式の効率と音声認識システムの性能との関係を導いた。4. では、4種類の対話制御方式の比較を行った。

2 確認対話の制御戦略

この節では、本稿で考察の対象とした対話制御戦略を略述する。先ず対話制御戦略を記述するためのモデルを導入する。これは文献 [1] によって用いられたものである。いまあるタスク（あるいは部分タスク）を達成するために n 個の情報項目（例えば、データベースの検索に必要な属性の値など）が必要であると仮定する。 n 個の情報項目を取得することを目標とする対話の状態は、3つのパラメータの組 (u, k, c) によって記述することができる。ここで、 u は対話システムにとってまだ未知の項目の数、 k は既知であるがまだ確認されていない項目の数、 c は既知でしかも確認された項目の数である。これらの間には、 $u + k + c = n$ という関係があるので、独立なパラメータの数は2個である。この事実を用いると、3字組 (u, k, c) を $k - c$ 平面上の格子点に写像することができる。図1は $n = 4$ の場合を示したものである。

この3字組を用いると、対話の目標は対話の状態

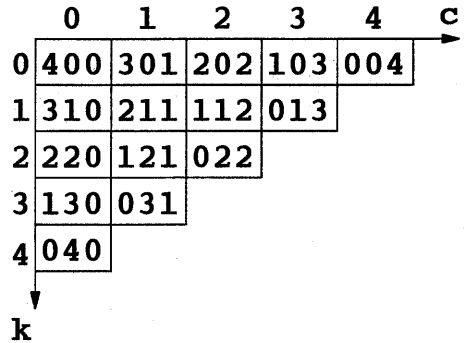


図1: $n = 4$ の場合の (k, c) -平面。

を $(n, 0, 0)$ から $(0, 0, n)$ へ変化させることである。対話制御戦略は、この状態遷移をどのような途中経過を経て行うかということで記述できる。また、対話制御戦略の効率は、上記の遷移を実現するのに必要な発話回数で測ることができる。

このような状態の遷移を引き起こさせるオペレータの例は、(a) 対話システムからの unknown な項目の質問に対するユーザの回答、(b) システムからの known な項目の確認に対するユーザの確認応答、等である。(a) の場合は、 $(-1, +1, 0)$ の状態変化をもたらし、(b) の場合は、ユーザの確認応答が 'yes' であれば $(0, -1, +1)$ の状態変化をもたらし、ユーザの確認応答が 'no' の場合は $(+1, -1, 0)$ の状態変化をもたらし、ユーザの確認応答が訂正応答であれば状態変化は起こらない。

従って、種々の対話制御戦略とユーザの行動は3字組 (u, k, c) の集合を状態空間とするマルコフ過程で記述することができる。後に詳述するように、状態間の遷移確率は、音声認識システムの性能、すなわち、1つの項目を正しく認識する確率によって決定される。本稿では、この事実に基づき、種々の対話制御戦略に対して、その効率と音声認識システムの性能との間の定量的な関係を導いた。

本稿では、次に述べる4つの対話制御戦略について詳しい解析を行った。対話の目的は、 n 個の情報項目をすべて取得し、かつ確認することである。

(1) 戦略1—個別入力、個別確認。一度に1つの項目を項目を質問し、それに対するユーザからの応答を直接確認する。ユーザから 'yes' の応答を得たら次に進む。これは1つの情報項目の状態を unknown

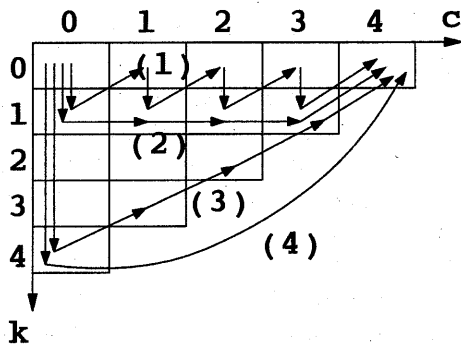


図 2: 四つの対話制御戦略の状態遷移.

→ known → confirmed としていくことを目標にしたもので、図 2 の矢印 (1) に沿って遷移を行わせようという戦略である。

(2) 戦略 2 — 個別入力、間接確認。最初 1 つの項目を質問して取得した後は、次の項目の質問の中に known な項目を埋め込む間接確認を用いる。間接確認の結果 known な項目が 2 つになったら、その 1 つについて 'yes' の応答が得られるまで、直接確認を行う。図 2 の矢印 (2) に沿って対話を制御しようとする戦略である。

(3) 戦略 3 — 一括入力、個別確認。一度に n 個の項目の質問をし、ユーザも一度に n 個の項目を答える。その後、1 項目ずつ戦略 1 と同様に直接確認を行う。図 2 の矢印 (3) に沿って、対話を制御しようとする戦略である。

(4) 戦略 4 — 一括入力、一括確認。一度に n 個の項目を取得し、その n 個を一度に確認する。確認の結果に従って、次の確認では訂正された項目だけの確認を行う。確認すべき項目の数を漸次減らしていく戦略で、図 2 の矢印 (4) に沿って対話を制御しようという戦略である。

3 対話制御戦略の解析

3.1 システムとユーザの行動

本稿で考えている対話システムにおいて、対話はシステムとユーザがとる次のような行動によって構成されるものとする。

システムの行動として次の 3 種類を考える。

(1) $Q(?item_1, ?item_2, \dots, ?item_n)$: n 個の情報項目を質問する。

(2) $DC(item_1, item_2, \dots, item_n)$: n 個の情報項目を一度に直接確認する。

(3) $IDC(item, ?newitem)$: 間接確認。1 個の known な項目 $item$ を質問文に埋め込み、1 つの unknown な項目 $newitem$ を質問する。

ユーザの行動として次の 3 種類を考える。

(1) $A(item_1, item_2, \dots, item_n)$: n 個の項目を同時に答える。

(2) *Yes*: システムの直接確認に対して、確認すべき項目がすべて正しいときのユーザの応答。

(3) *No*: システムの直接確認に対して、確認すべき項目がすべて誤っていたときのユーザの応答。

これらの行動を用いて、後に使用する subdialogue を定義しておく。

(1) *DC_subdialogue* — Direct Confirmation Subdialogue

- ・システムはある既知の項目 $item$ を *DC* によって確認する。

- ・項目が正しく認識されていれば、ユーザは *Yes* と応答する。この場合この subdialogue は終了する。

- ・項目が誤っていれば、ユーザは $A(item')$ によって訂正応答を行う。この場合は新しい項目 $item'$ について *DC_subdialogue* が再帰的に開始される。

以上の手続きを C-program 風に書くと次のようになる。

```
DC_subdialogue(item)
{ DC(item);
  if( item is correct ) Yes;
  else{ A(item'); DC_subdialogue(item'); }
  return;
}
```

この subdialogue に入るときにシステムにとって既知である $item$ が正しく認識されている確率を p とする。また、この subdialogue 内で、システムの確認 $DC(item)$ に対するユーザの訂正応答 $A(item')$ を認識したとき、 $item'$ が正しく認識される確率を q とする。また、*Yes* (及び *No*) はすべて正しく認識されるものと仮定する。

この *DC_subdialogue* の状態遷移図を描くと図 3 のようになる。ここで、 $*$ は任意の正の整数を、 $1(p)$ ($1(q)$) は 1 個の既知の項目が正しく認識されている確率が p (q) であることを示している。状態から状

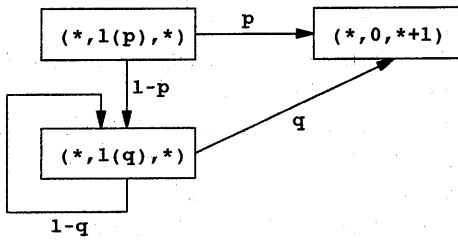


図 3: DC_subdialogue に対する状態遷移図。

態へ至る矢印には、システムとユーザの1つの発話の交換が対応している。また、矢印に付記された記号は状態遷移確率である。(2) IDC_subdialogue

— Indirect Confirmation Subdialogue

- システムは既知の項目 *item* を質問文に埋め込んで、未知の項目 *newitem* を質問する。

- item* が正しく認識されていれば、ユーザは質問された *newitem* のみを答える。この場合はこの subdialogue は終了する。

- item* が誤って認識されていれば、ユーザはその *item* に対する訂正情報と、質問された *newitem* に関する情報を1つの発話で返す。システムは、訂正された *item* について先に定義した DC_subdialogue を行い、これが完了した時点で IDC_subdialogue を終わる。

以上のやりとりを C-program 風に書くと次のようになる。

```

IDC_subdialogue(item, ?newitem)
{ IDC(item, ?newitem);
  if( item is correct ) A(newitem);
  else
  { A(item', newitem);
    DC_subdialogue(item'); }
  return;
}

```

この subdialogue による状態遷移図を描こう。ここでは簡単のため、ユーザの行動 $A(item)$ を認識したとき *item* が正しく認識される確率を p とする。また、 $A(item, item')$ に対して、*item*、*item'* の各々が正しく認識されている確率を q とする。この遷移図を対話戦略2の解析に用いる場合、この subdialogue に入ってくる時既知になっている項目が正しく認識されている確率は p である。このとき、IDC_subdialogue

内の $IDC(item, ?newitem)$ の後 $A(newitem)$ でこの subdialogue を終わる確率は p である。このとき状態は $(*, 1(p), *)$ から $(* - 1, 1(p), * + 1)$ 遷移する。すなわち、いままで既知であった項目が確認され、未知であった1つの項目が新たに *newitem* として既知になる。これは $A(newitem)$ によって既知になっているので、それが正しく認識されている確率はやはり p である。

また、 $IDC(item, ?newitem)$ の後確率 $1 - p$ で $A(item', newitem)$ が起こる。このときは、いままで既知であった項目 *item* が *item'* と訂正され、また未知であった項目が *newitem* として既知になる。これらの各々が正しく認識されている確率は、仮定により q である。このとき状態は $(*, 1(p), *)$ から $(* - 1, 2(qq), *)$ へ移る。この先は DC_subdialogue によって、最終的に状態は $(* - 1, 1(q), * + 1)$ へ遷移する。

従って、状態 $(*, 1(p), *)$ で IDC_subdialogue を実行すると図4に示すような2つの異なった状態 $(* - 1, 1(p), * + 1)$ 及び $(* - 1, 1(q), * + 1)$ に遷移することが分かる。

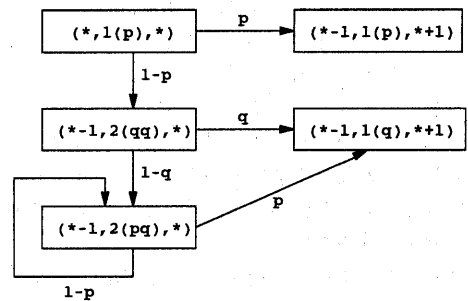


図 4: IDC_subdialogue に対する状態遷移図。

3.2 対話戦略1の解析

3.1で定義した記号を用いて、この対話戦略に従う対話の進行状況を C-program 風に記述すると、次のようになる。

```

for(k = n; k > 0; k --)
{ Q(?item_k); A(item_k);
  DC_sudialogue(item_k);
}

```

図3を参考にして、上記の対話の進行状況のうち、for-loopの内部を状態遷移図に表すと、図5のようにになる。システムの質問に対するユーザの応答とシステムの確認に対するユーザの訂正応答を比較すると、一般に後者の方が「いいえ、...でなくて...です」のようにやや複雑になり、認識が若干難しくなる可能性がある。従って、 $p \geq q$ である。この事を考慮して、ここでは両者における情報項目の認識率を区別した。しかし、実際にはこの差はそれほど大きくはならないと考えられるので、この区別は戦略1の解析のみで使用し、それ以外の場合は区別していない。また、直接確認の変形として、 $DC(item)$ によって確認された項目が誤って認識されていた場合、ユーザは訂正情報を与えるのではなく、単に'no'といい、これに対してシステムは、最初の項目に関する質問を繰り返すもの、を考える。この場合の状態遷移図を図6に示した。

図5の任意の状態 $s = (u, k, c)$ から対話の最終状態 $(0, 0, n)$ に至るまでに必要な平均 exchange の数を $N(u, k)$ と表現する。図5に対して次の式が成立する。

$$N(k, 0) = 1 + N(k-1, 1(p))$$

$$\begin{aligned} N(k-1, 1(p)) &= 1 + pN(k-1, 0) + (1-p)N(k-1, 1(q)) \\ N(k-1, 1(q)) &= 1 + qN(k-1, 0) + (1-q)N(k-1, 1(q)) \end{aligned}$$

これから $N(k-1, 1(p))$ と $N(k-1, 1(q))$ を消去すると、次の漸化式を得る。

$$N(k, 0) = \left(2 + \frac{1-p}{q}\right) + N(k-1, 0) \quad (1)$$

また、 $N(n, 0)$ を $N(n)$ と略記し、 $N(0, 0) = 0$ であることに注意すると、

$$N(n) = \left(2 + \frac{1-p}{q}\right)n \quad (2)$$

上式の右辺の第2項が、音声認識の誤りにより増加した exchange の数を表している。

図6について同様の解析を行うと、次式が成立する。

$$N(k, 0) = 1 + N(k-1, 1(p))$$

$$N(k-1, 1(p)) = 1 + pN(k-1, 0) + (1-p)N(k, 0)$$

両式から $N(k-1, 1(p))$ を消去すると、次の漸化式を得る。

$$N(k, 0) = \frac{2}{p} + N(k-1, 0)$$

これから、

$$N(n) = \frac{2n}{p} \quad (3)$$

となる。(2)及び(3)で定義される $N(n)$ をそれぞれ N_{corr} と N_{no} と書くと、簡単な計算により、 $q > p/2$ ならば $N_{no} > N_{corr}$ となるから、 $q > p/2$ である限り、システムの $DC(item)$ に対して、 $item$ が誤って認識されていた場合ユーザは訂正応答を与えた方が有利であることが分かる。

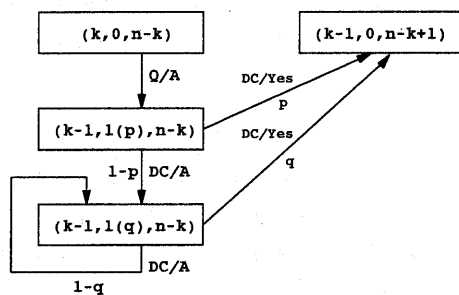


図5: 対話制御戦略1の状態遷移図。

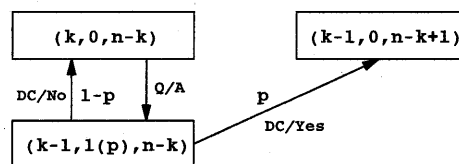


図6: 対話制御戦略1の変形に対する状態遷移図。

3.3 対話制御戦略2の解析

この対話制御戦略に従う対話の進行を C-program 風に表すと、次のようになる。

```
Q(?itemn); A(itemn);
for(k = n; k > 1; k--)
    IDC_subdialogue(itemk, ?itemk-1);
DC_subdialogue(item1);
```

上記のプログラムから

$$N(n) = 1 + N(n-1, 1(p)) \quad (4)$$

である。状態 $(n-1, 1(p), 0)$ から先は、図4に示した遷移図に従って遷移する。状態 $(*, 1(q), *)$ からの遷移は、状態 $(*, 1(p), *)$ からの遷移と同様で、確率 q で状態 $(*-1, 1(p), *+1)$ へ、確率 $1-q$ で状態 $(*-1, 1(q), *+1)$ へ遷移する。従って、 $N(k, 1(p))$ と $N(k, 1(q))$ (以下ではそれぞれ a_k 、 b_k と表す) に関して、次の漸化式を得る。

$$a_k = 1 + pa_{k-1} + (1-p)N(k-1, 2(qq)) \quad (5)$$

$$b_k = 1 + qa_{k-1} + (1-q)N(k-1, 2(qq)) \quad (6)$$

$N(k-1, 2(qq))$ については、対話制御戦略1の解析における $N(k-1, 1(p))$ と同様な関係が成立するから、

$$N(k-1, 2(qq)) = b_{k-1} + 1 + \frac{1-q}{p} \quad (7)$$

一方、図4に従う遷移は状態 $(0, 1(p), n-1)$ または状態 $(0, 1(q), n-1)$ で終わり、そこから先は図3に示した遷移と同様の経過をたどり状態 $(0, 0, n)$ に移る。(但し、図3と図4では p と q の役割の異なることに注意)。従って、上記の連立漸化式の初期条件として、次の関係を得る。

$$a_0 = \frac{1}{p} \quad (8)$$

$$b_0 = 1 + \frac{1-q}{p} \quad (9)$$

(5) ~ (9) をまとめると、

$$a_k = pa_{k-1} + (1-p)b_{k-1} + A \quad (10)$$

$$b_k = qa_{k-1} + (1-q)b_{k-1} + B \quad (11)$$

$$A = 1 + (1-p)b_0 \quad (12)$$

$$B = 1 + (1-q)b_0 \quad (13)$$

となる。ここで、

$$c_k = a_k - b_k$$

$$d_k = qa_k + (1-p)b_k$$

とおくと、(10) 及び (11) は、 c_k 及び d_k の単独の漸化式に帰着できる。この事実を用いると、 a_k に関して次式を得る。

$$a_k = \frac{k+1}{p} + \frac{2(1-p)\delta p}{p(1-\delta p)} \left[k - \frac{\delta p(1-\delta p^k)}{1-\delta p} \right] \quad (14)$$

ただし、 $\delta p = p - q$ である。

特に $p = q$ のときは、(14) の第2項が消えるので、

$$N(n) = 1 + a_{n-1} = 1 + \frac{n}{p} \quad (15)$$

となる。

3.4 対話制御戦略3の解析

この対話制御戦略のもとでは、対話は次のように進行する。

$Q(?item_1, ?item_2, \dots, ?item_n);$

$A(item_1, item_2, \dots, item_n);$

$for(k = n; k > 0; k --)$

$DC_subdialogue(item_k);$

この対話制御戦略の効率の解析を簡単にするために、次のことを仮定する。すなわち、 n 個の項目を一度に認識したとき、各項目の置換誤りのみが起こり、脱落及び挿入誤りは起こらない。また、誤りは各項目ごとに独立に起こるものとする。1つの項目が正しく認識される確率を q (n の値に依存することに注意) とする。この場合、状態 $(0, k(q^k), n-k)$ から状態 $(0, k-1(q^{k-1}), n-k+1)$ への遷移は、上記のように直接確認によって引き起こされるので、その状態遷移は図3と同じものとなる。但し、遷移確率の値は異なる。従って、3.2節の解析結果を参照して、

$$N(n) = 1 + N(0, n(q^n)) = 1 + (1 + \frac{1-q}{p})n \quad (16)$$

となる。

3.5 対話制御戦略4の解析

この対話制御戦略のもとでは、対話は次のように進行する。

$Q(?item_1, ?item_2, \dots, ?item_n);$

$A(item_1, item_2, \dots, item_n);$

$DC(item_1, item_2, \dots, item_n);$

$A(item_{11}, item_{12}, \dots);$

.....

$DC(item_{k1}, item_{k2}, \dots);$

$Yes;$

ここで、音声認識の誤りに関して前節で行ったのと同様の仮定をおく。また、 $A(item_1, \dots, item_m)$ を

認識した際に、1つの項目が正しく認識される確率を q_m とする。

いま、対話の状態が $(0, m, n - m)$ にあるとする。上記の定義により m 個の known items のうち、 k 個が誤っている確率は ${}_m C_k q_m^{m-k} (1 - q_m)^k$ ($0 \leq k \leq m$) である。これが、システムが m 個の known items の確認をしたとき、ユーザが k 個の項目の訂正入力をして、対話の状態が $(0, k, n - k)$ に移る確率となる。従って、 $a_m = N(0, m)$ とおくと、 a_m に関して次の漸化式が成立する。

$$a_m = 1 + \sum_{k=0}^{m-1} {}_m C_k q_m^{m-k} (1 - q_m)^k a_k \quad (1 \leq m \leq n) \quad (17)$$

$a_0 = 0$ だから、

$$a_m = \frac{1 + \sum_{k=1}^{m-1} {}_m C_k q_m^{m-k} (1 - q_m)^k a_k}{1 - (1 - q_m)^m} \quad (1 \leq m \leq n) \quad (18)$$

となる。従って、求める $N(n)$ は、

$$N(n) = a_n + 1 \quad (19)$$

となる。

4 対話制御戦略の比較

この節では、前節の解析結果を用いて、各対話制御戦略の効率を比較する。まず理想的な場合、すなわち、音声認識に全く誤りがない場合、各対話制御戦略の $N(n)$ を表1に示した。この結果は図1に示した各制御戦略の状態遷移から直感的に明らかである。

次に音声認識に誤りのある場合を考える。各対話制御戦略とも複数個のパラメータを含むので、簡単のためまず、対話制御戦略1、2、3では、 $q = p$ 、対話制御戦略4では、 $q_m = p$ ($1 \leq m \leq n$) とした $N(n)$ の下限（最も効率の良い場合）について考える。対話制御戦略1、2、3については、表1に示すように $N(n)$ の表現が簡単になる。また、対話制御戦略4も含めて $N(n)$ と p との関係を図示すると、図7のようになる。

表1から分かるように、対話制御戦略1、2、3では、 $N(n)$ は n と共に直線的に変化し、その係数は1より大である。しかし、対話制御戦略4では数値計算の例を示すと、 $p = 0.9$ のとき、 $N(1) = 2.11$ 、

$N(2) = 2.21$ 、 $N(3) = 2.30$ 、 $N(4) = 2.39$ 、 $N(5) = 2.46$ と変化し、 n に依存するところは極めて少ない。

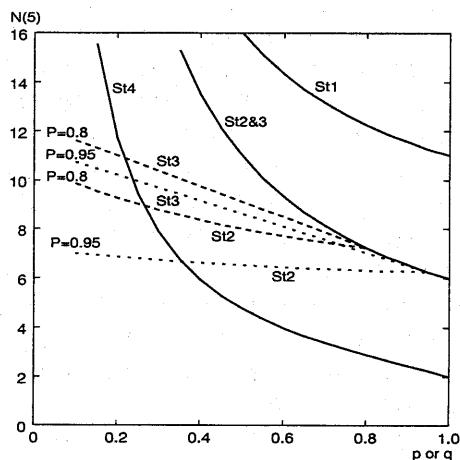


図7: 四つの対話制御戦略に対する p と $N(5)$ の関係。

また、各対話制御戦略の効率の p への依存性についてみると、対話制御戦略1、2、3では $1/p$ で変化する。ところが、対話制御戦略4では図7の $N(n)$ から分かるように、 p の指数関数的に変化しているように見える。また、図7の各曲線は $N(5)$ の下限を与えるものであるが、これらを用いて $N(5)$ の範囲を推定することも可能である。例えば、対話制御戦略4において $q_1 = 0.95$ 及び $q_5 = 0.7$ であれば、両者の $N(5)$ の値を求めると 2.24 と 3.39 となるので、 $N(5)$ は $2.24 \leq N(5) \leq 3.39$ の範囲に収まらうと推定できる。

次に、対話制御戦略2、3の効率の第2パラメータ q への依存性を数値計算によって調べよう。 $p = 0.95$ 及び $p = 0.8$ のときの q 変化に対する $N(5)$ の変化を図7に示した。これにより実用的な確率値の範囲では、第2パラメータ q に対する $N(n)$ の依存度は小さいといえる。また、対話制御戦略2と3では対話制御戦略2の方がやや効率が勝っているといえる。

表1 対話制御戦略の効率

対話制御戦略		1	2	3	4
N(n)	ideal	2n	n+1	n+1	2
	p=q	n+n/p	1+n/p	1+n/p	—

5 あとがき

本稿では、確認対話の制御方式の効率と対話システムに使用される音声認識システムの性能との間の定量的な関係を導いた。対話の目的が一定数の情報項目を取得することであると仮定して、目的達成に必要なシステムとユーザの間の発話の交換回数、すなわち、システムとユーザの各1発話が対になっているとして、その対の数で、対話制御戦略の効率を表現した。また、音声認識システムの性能としては、1つの発話に含まれている情報項目の1つが正しく認識される確率を用いた。対話制御方式としては次の4種を検討した。(1) 情報項目の個別入力、個別(直接)確認、(2) 1項目の間接確認を含む個別入力、(3) 必要な情報項目の一括入力、個別確認、(4) 一括入力、一括確認。

この研究により、音声認識システムの性能と、タスクを達成するために必要な情報項目数が与えられたとき、最適な制御戦略を選択するための、あるいは、フィールドの環境により音声認識システムの性能が動的に変化するとき、対話制御戦略を動的に適させるための理論的な基礎が与えられたといえよう。

しかし、本稿では音声認識システムの誤りとして、置換のみが起こると仮定しており、挿入や脱落誤りを考慮していないので、これらについては今後の検討が必要である。

参考文献

- [1] Cozannet, A. and Siroux, J., "Strategies for oral dialogue control," Proc. of ICSLP'94, pp.963-966 (1994).
- [2] Waterworth, J. A., "Interactive strategies for conversational computer system," in Structure of Multimodal Dialogue, ed by Taylor, M. M., et al., pp.331-340, North Holland (1988).
- [3] Niimi, Y., Kobayashi, Y., "Modeling dialogue control strategies to relieve speech recognition errors," Proc. of EUROSPEECH'95, pp.1177-1180 (1995).
- [4] Niimi, Y., Nishimoto, T., and Kobayashi, Y., "Analysis of interactive strategy to recover from misrecognition of utterances includ-

ing multiple information items," Proc. of EUROSPEECH'97, pp.2251-2254 (1997).