

テキストコーパスを用いた音声認識誤り訂正手法

石川 開

隅田 英一郎

E-mail: {ishikawa, sumita}@itl.ATR.co.jp

ATR 音声翻訳通信研究所

1. はじめに

近年の音声認識、および機械翻訳の性能向上により、これらの統合である音声翻訳システムの実現を目指した研究活動が活発に行われている。実用的な対話システムの実現を考えた場合、音声認識と機械翻訳それぞれ独立の性能の向上だけでなく、システム全体として情報を伝えるための頑健性の向上も重要である。このため、誤りを含んだ音声認識結果を適切に翻訳処理する枠組みが求められている。

このような枠組みとして、脇田らは認識結果のうち正しい部分のみを翻訳する部分翻訳を提案し、システム全体の頑健性を改善できることを示した[1]。しかし、誤り部分に本来あるべき情報が失われてしまうという問題がある。

我々はこの問題へのアプローチとして、誤り部分を用例文から推定し訂正する手法をベース中の用例を用いて誤りの訂正候補を作成し、音韻と意味の両方の観点からその妥当性を判断するもので、評価実験で有効性が確認された。しかし、用例と認識結果の解析木を照合する際、認識誤りのために構造の不一致が生じると、適切な句構造での訂正が妨げられるという問題があった。

本稿では、テキストコーパス中の用例文を、解析木ではなく形態素列で照合する訂正手法を提案する。本手法によれば、用例を構造に依存しない形態素列によって照合するので、用例のより柔軟な照合が可能である。本提案手法を音声翻訳システムに組み込んで、その有効性を翻訳率の観点から評価した。旅行会話337文に対して、認識結果、訂正結果の翻訳率を比較した結果、それぞれ64%、74%となり、その有効性が確認された。

2. 提案手法

2.1. 全体の処理の流れ

本節では、まず提案手法の全体的な処理の流れを説明する。処理手順は、次に示すように、認識結果の(1)意味的訂正判断、音韻的に近い用例を用いた(2)形態素列訂正、意味距離に基づく(3)訂正候補の妥当性判断、からなる(図1参照)。

(1) 意味的訂正判断

(1-1)構造解析と意味距離計算

(1-2)訂正必要性判断

(1-3)訂正箇所の特定

(2) 形態素列訂正

(2-1)テキストコーパスからの用例文検索

(2-2)訂正箇所との対応付け

(2-3)形態素列の置換

(2-4)訂正候補の音韻的選別

(3) 訂正候補の妥当性判断

(3-1) 構造解析と意味距離計算

(3-2) 訂正必要性判断

(1)意味的訂正判断では、まず入力された音声認識結果をCBパーザー[3]を用いて構造解析し、解析木構造を得る(1-1)。この際、構造の各ノード及び構造全体に対し、意味距離が計算される。得られた構造全体に対する意味距離が閾値 θ よりも小さい場合、誤り訂正は不要と判断する(1-2)。大きい場合は、閾値 θ を超える意味距離を持つすべての部分構造を抽出し、その部分構造を含む形態素列を訂正箇所とする(1-3)。

(2)形態素列訂正では、(1-3)で得られた訂正箇所の形態素列に対して、音韻的に近い形態素列をテキストコーパス中から取り出す。まず類似文字列検索によりテキストコーパス中から似た文字列を含む用例文を検索する(2-1)。この検索された用例文の形態素列から、訂正箇所の形態素列に対応する部分を抽出し(2-2)、この部分に関して認識結果の形態素列を置換することで訂正候補を得る(2-3)。この訂正候補の認識結果に対する音韻的距離を計算し、閾値 Δ より小さな候補のみを残す(2-4)。

(3)訂正候補の妥当性判断では、(2-4)で残った訂正候補に対して(1-1)と同様に構造解析を行い、訂正候補に対する解析木構造を得る(3-1)。訂正候補の中で、構造全体に対する意味距離の値が閾値 θ より小さいものが存在した場合、それら全ての候補は誤りを回復したものとみなし、音韻的距離が最小のものを最終的な訂正結果として終了する(3-2)。候補中に条件を満たすものが存在し

ない場合、(1-3)に戻り、それぞれの訂正候補から再帰的に訂正を試みる。

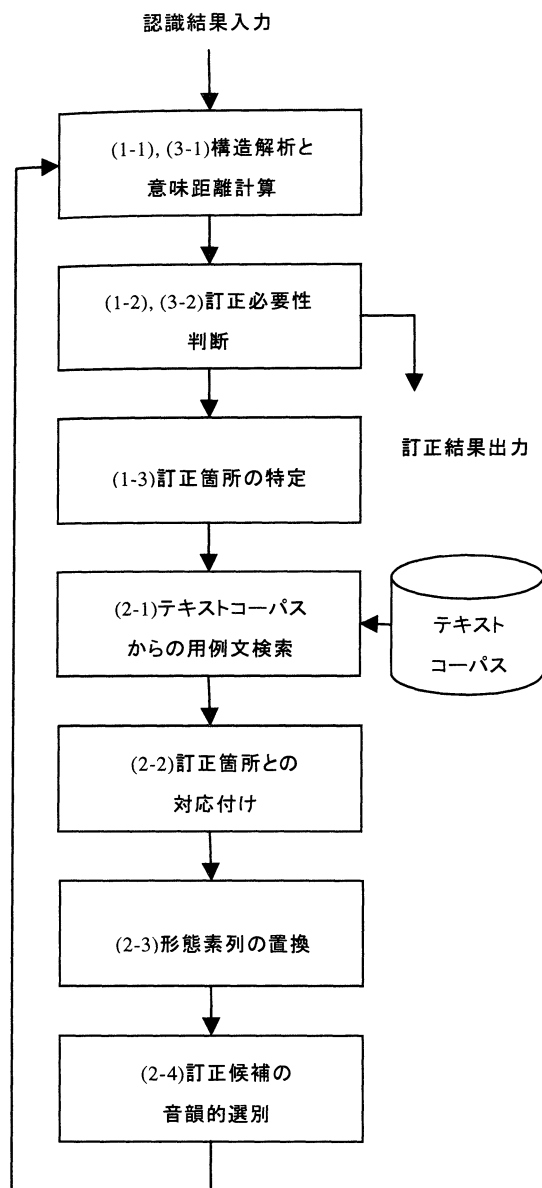


図1: 処理の流れ

2.2. 各処理の詳細

2.2.1. 構造解析と意味距離計算

(1-1)および(3-1)では、認識結果または訂正候補の構造解析を行う。本手法では、構造解析としてC Bパーサー[3]を用いる。C Bパーサーは、入力文に対して依存関係を適用し、ボトムアップに依存構造を作成する。同時に、依存構造の各係り受けに対して、意味距離が計算される。この意味距離は、用例と入力の意味的な距離であり、シソーラスの意味情報[4]に基づいて計算される。適切な係り受けである場合、この意味距離が小さな値をとる。このため、係り受けの妥当性の指標とすることができる。

2.2.2. 訂正箇所の特定

(1-2)では、構造解析の結果得られた依存構造と意味距離に基づいて、訂正箇所の特定を行う。依存構造の各ノードに対する意味距離は、値が小さければ依存関係が妥当であると信頼できる。それゆえ、訂正箇所を含む部分特定すればよい。まず、依存構造の全ノードのうち閾値 θ よりも大きな意味距離を持つノードを信頼性の低い依存関係と見なし、これらノードより上位のすべて依存構造を取り出す。そして、この取り出された構造それぞれが含む形態素列を訂正箇所とする。これにより、信頼性の低い依存関係を含むすべての依存構造を訂正箇所として考慮することになる。

2.2.3. テキストコーパスからの用例文検索

テキストコーパスは、あらかじめ形態素解析によって、品詞付きの形態素列に変換し蓄えられている。(2-1)と(2-2)では、この中から訂正箇所に音韻的に近い形態素列を抽出する。(2-1)では、このうち前段的な処理として、音韻的に近い形態素列を含む用例文を検索する。ここで、音韻的な近さを文字列の編集距離によって近似することにより、訂正箇所の文字列とテキストコーパスの文字列間の誤りを許した照合の問題に帰着する。本手法では、訂正箇所の文字列に対して、編集距離の比が閾値 δ 以下の文字列を含む用例文を、Y. LePageの方式[5]を用いて検索する。

2.2.4. 用例形態素列の置換

(2-2)では、(2-1)で検索された各用例文から、認識結果の訂正箇所に対応する形態素部分列を抽出する。ここでは、用例文の可能な形態素部分列のうち、訂正箇所に対して形態素部分列の編集距離が最小となるものをDPマッチングによって求める。このようにして得られた用例文の形態素部分列を認

識結果の訂正部分と置換することにより、訂正候補の形態素列が得られる。

2.2.5. 音韻的距離の計算

(2-4)では、訂正前の認識結果の音素列に対する訂正候補の音素列の編集距離の比を音韻的距離と定義し、訂正候補の認識結果に対する音韻的距離が閾値Δより小さな候補のみを残す。これにより、訂正候補の中で元の認識結果に音韻的に近いものだけが得られる。

2.2.6. 再帰的訂正処理

(2-4)において、音韻的に認識結果に近い訂正候補が得られても、(3-2)において、意味距離の観点で誤りが回復したと判断されなければ訂正結果としては認めない。このような訂正候補の中には、単に訂正が間違っているものの他に、訂正は正しく部分的には誤りが回復したが、訂正箇所以外にさらに誤りが存在するために完全には回復していないものも含まれる。このような後者の候補に対しては、さらに誤り訂正を繰り返すことで完全な回復が見込まれる。このため、訂正候補中に意味距離の条件を満たすものが存在しない場合、(1-3)に戻り、それぞれの訂正候補から再帰的に訂正を試みる。

3. 評価実験

本提案手法の有効性を確認するために、提案手法を音声翻訳システムに組み込んで翻訳実験を行った。入力には日英翻訳システムで学習済みの旅行会話データのうち日本語337文を用いた。テキストコーパスは、同様に旅行会話データのうち入力文を含んだ異なり15,264文を用いた。ここで意味距離に対する閾値θは1.0を、音韻的距離に対する閾値Δは0.3を、類似文字列検索に対する閾値δは0.3を用いた。

まず、訂正必要性判断と認識誤りの関係を表1に示す。(1-2)訂正必要性判断において、音声認識結果が正しいものに対しては、91%の精度で訂正が不要と判断し、認識誤りのある入力に対しては、59%の精度で訂正が必要と判断している。さらに、訂正が必要と判断されたものに対して、訂正結果が出力されたのは認識誤りのある入力のうち27%であり、正しいにも関わらず訂正が必要と判断された入力に対しては訂正は得られていない。

表1: 訂正必要性判断と認識誤りの関係 (全337入力文)

	誤った入力 224	正しい入力 93
訂正必要 153	145 (59%)	8 (9%)

(訂正出力 66)	66(27%)	0(0%)
訂正不要 184	99 (41%)	85 (91%)

次に、翻訳結果の評価について述べる。翻訳結果に対する評価基準として、以下に示すような、情報伝達に着目した4段階の基準を用いた。

表 2: 翻訳結果の評価基準

評価 A (完全)	原言語正解分と完全に同じ意味にとれる
評価 B (十分)	原言語正解分とはほぼ同じ意味にとれる
評価 C (部分)	原言語正解文の主要情報が部分的には伝わる
評価 D (不適)	原言語正解文の主要情報が伝わらない／誤解が生じる

入力337文の音声認識結果に対して、直接翻訳処理した結果と、提案手法を用いた結果を評価した結果を表3に示す。

表 3: 提案手法による翻訳結果 (全337入力文)

	訂正なし	訂正あり
翻訳率	64%(217)	74%(249)
評価 A	41% (139)	54%(185)
評価 B	11% (38)	10% (34)
評価 C	12% (40)	9% (30)
評価 D	25% (85)	18% (61)
NIL	10% (35)	8% (27)

訂正によって評価Aへ回復したものが多く、翻訳率（評価A B Cの合計）も64%から74%に向上していることが分かる。さらに訂正の出力が得られた66文のみに関して、比較してみると、表4のようになる。

表 4: 訂正の得られた66入力文に関する比較

	訂正なし	訂正あり
翻訳率	45% (30)	94% (62)
評価 A	5% (3)	74% (49)
評価 B	20% (13)	14% (9)
評価 C	21% (14)	6% (4)
評価 D	42% (28)	6% (4)
NIL	12% (8)	0% (0)

