

タスク文法による N-gram 確率の部分強化を用いた 認識アルゴリズムの評価

鶴身 玲典 李 晃伸 猿渡 洋 鹿野 清宏

あらまし 音声対話のための音声認識システムにおいては、言語モデルとして N-gram モデルとネットワーク文法が用いられる。統計的言語モデルである N-gram モデルでは、タスクで想定されていない発声もある程度柔軟に認識することができ、発話のバリエーションに対する頑健性が高いが、学習には大規模なコーパスが必要であり、また、文レベルの制約を表現することができない。一方、人手で記述されたネットワーク文法では、文法上受理可能なタスク内の発声は高精度で認識できるが、文法で受理できないタスク外の発声に対しては認識できない。本研究では、N-gram モデルを用いた認識において、ネットワーク文法の 2 単語間の接続制約を用いて N-gram 確率を動的に強化する認識アルゴリズムを提案する。本手法により、タスク外の発声に対する頑健性を保ちつつ、タスクに特有の定型表現の認識精度を向上させ、トータルでユーザー発話の変化に頑健で、かつ高精度な認識を実現する。受付案内対話タスクにおいて提案法を評価した結果、タスク内の発声に対して文法のみを用いる場合に近い高い認識率が得られ、トータルの単語正解精度は 82.4 % から 88.0 % に向上した。また、単語の生起確率のみを強化する場合や、サブタスクごとの強化の評価、および静的な確率値変更についても述べる。

キーワード 対話音声認識, デコーダ, Julius, 記述文法

Speech Recognition Algorithm Strengthening N-gram Probability with Task Grammar

Ryosuke TSURUMI Akinobu LEE Hiroshi SARUWATARI Kiyohiro SHIKANO

Abstract

In speech dialogue systems, both word N-gram model and written network grammar are mainly used as language models. Since a word N-gram model is a statistical language model, it is practically feasible for various speech expressions and even for an out-of-task utterance. However, it needs a large amount of training text and also cannot express sentence-level constraints. The written grammar, on the other hand, is a deterministic expression of sentence patterns and thus can result in accurate recognition if the content of an utterance is literally acceptable to the grammar. However, it easily fails when the utterance is out of grammar. In this paper, we propose a speech recognition algorithm using both word N-gram and network grammar, which strengthens N-gram probabilities dynamically by using word-pair constraint. By enhancing the connection of N-gram words that are allowed to be connected in the grammatical constraint, in-task utterances that are expected by the system can be recognized more accurately, while keeping the feasibility of N-gram against unexpected utterances. We evaluated this method in a robot receptionist task, and improvements of word accuracy from 82.4 % to 88.0 % was achieved. Comparison with enhancement without word context, enhancement per sub-topic task, effects for unknown word entries, and static modification of probabilities are also investigated.

Key words Dialogue Speech Recognition, Decoder, Julius, Grammar

1 はじめに

近年、音声インタフェースや対話音声認識システムが実用化されつつある [1]。音声認識システムでは、統計的言語モデルである N-gram モデルや、単語の接続をネットワークで記述したネットワーク文法が主に用いられている。N-gram モデルは、近傍の単語接続の統計的な言語モデルであり、コーパスから単語接続の出現確率を学習している。そのため、タスクで想定されない文についても柔軟に認識できるという利点があるが、学習には大規模なコーパスが必要であり、文レベルの制約まで表現することができないといった欠点がある。一方で、ネットワーク文法は、単語の接続を手手で記述しており、あらかじめ想定されている文に対しては、一般に N-gram モデルよりも高精度で認識できるという利点がある。しかし、システムが想定していない発声に対しては認識できず、あらかじめ想定される文パターンを全て記述しておくことは困難である。

そこで、本研究では N-gram モデルを用いた認識においてネットワーク文法の単語対制約を組み込む認識アルゴリズムを提案する。タスク外の発声にも柔軟に対応でき、タスクに特有の定型表現の認識精度を向上させることを目的とする。

以下、2章で音声対話における言語制約適用の問題点について述べ、3章で提案手法の説明を行う。提案手法の有効性を示すため、4章では奈良先端科学技術大学院大学の受付案内ロボット ASKA [2] との対話タスクにおいて認識実験を行った結果を示す。また、単語対の効果調べるために、単語履歴を用いずに単純に文法上の単語を出現しやすくする手法との比較を行う。5章では未知語に対する確率強化の効果を評価する。6章では、サブタスクごとに分割した文法を用いて、話題同定と組み合わせた時の性能について検討する。7章では N-gram 確率値をあらかじめ静的に変更しておく手法について検討する。

2 音声対話認識のための言語モデル

音声対話システムにおける言語モデルとしては、従来人手で記述したネットワーク文法が多く用いられてきた。ネットワーク文法では単語同士の接続は文法で記述されているもの以外は許されないという強い制約があり、比較的小語彙の限定されたタスクにおける認識に向いている。つまり、それだけユーザーの発話に対する文候補の数が限られるため、文法で受理可能な発話に対しては高精度で認識することが可能である。しかし、ユーザーが文法で受理不可能な発声を行った場合は、認識に失敗する。実際、ユーザーの自然な発話におい

表 1: 言語モデルの特徴

文法	人手で記述 想定文に関しては高精度で認識可能 × タスク外の発声に対しては認識に失敗 × 全ての文パターンの記述は困難
N-gram	コーパスから統計的に学習 タスク外の発声も柔軟に認識可能 × 学習には大規模なコーパスが必要 × 文レベルの制約を表現できない

ては、同一の内容に対しても様々な表現が存在し、それらを全てカバーすることは困難である。

一方、近年の大語彙連続音声認識では N-gram 統計言語モデルが盛んに用いられている。統計的言語モデルである N-gram 言語モデルでは、通常数万語の単語の出現確率を大規模なコーパスから学習し作成される。N-gram モデル中では単語 N 個までの接続に対する出現確率が割り当てられており、言語的な制約が確率的にモデル化されている。学習された N-gram モデルは学習コーパスの語彙や単語の接続の制約を反映したものとなる。固定的なモデル表現ではないため、文法と異なり、想定外の発声であってもある程度認識することが可能である。逆に、「はい」「いいえ」など、対話の状況やシステムの状態によって対話のパターンが限られる場合のように、特定の少数パターンの文のみを認識したいといった場合には不向きである。この文法と N-gram モデルの特徴を表 1 にまとめる。

音声対話システムではディクテーションと異なり、ある程度対象とするタスクが限定されているため、ユーザーの発声の多くはあらかじめ想定可能なフレーズである。その反面、より自由な対話を実現する対話システムでは、ユーザーの多種多様な言い回しに対しても認識できるようにする必要がある。つまり、音声対話システムのための言語制約としては、あらかじめ想定できるフレーズを誤りなく認識でき、かつ、想定外の発話に対してもある程度の精度で認識できることが要求される。

3 文法の単語対制約を用いた N-gram 確率の強化アルゴリズム

3.1 単語対制約による N-gram 確率の強化

単語 N-gram 言語モデルを用いた認識の探索過程で、別に作成した文法における 2 単語間の接続制約を用いて N-gram 確率の強化を行う手法を提案する。これにより、文法で受理可能な単語列が探索されやすくなるため、あらかじめ想定されている発話に対しては認識精度が向上すると考えられる。ここでは文法の 2 単語間の接続制約のみを

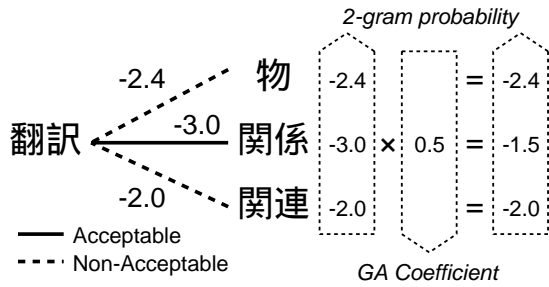


図 1: 単語対に対する 2-gram 確率の強化

用いて、部分的に文法にマッチする候補を押し上げる。

具体的には、図 1 のようにデコーダ内の探索過程で「翻訳-物」「翻訳-関係」「翻訳-関連」という「翻訳」に続く 3 種類の単語接続が現れたとき、「翻訳-物」と「翻訳-関連」は文法で受理不可能で、「翻訳-関係」のみが文法で受理可能であれば、この単語接続の対数尤度に係数をかけるという操作を行い、文法で受理不可能な単語接続に対しては操作を行わない。上記の例では、「翻訳-関係」という単語接続に対してのみ対数尤度に係数 0.5 がかけられ、他の候補よりも優先される。

N-gram モデルと文法の両方を用いた認識アルゴリズムの一例として、言語モデルごとに並列に認識を行った結果を選択する並列デコーディングが挙げられるが、この場合、候補の選択基準が曖昧である。また、N-gram での N-best 結果を文法で再評価する手法も考えられる [3] が、部分的な文法制約を表現できない。本手法では文法の単語対を用いて確率の強化を行うことにより、探索中に文法上許される接続を優先して、その他を候補から除外できるため効率がよい。また、文法と N-gram モデルのネットワークを結合する手法 [4] では固定的に文法を利用しているが、過不足のない文法を書くことは困難である。本手法では単語対を用いることにより、文の一部が文法に適合する場合などに対してもある程度有効であると考えられる。ただし、N-gram モデルの辞書と表記を統一するために、文法の単語も形態素単位で記述せねばならない。

3.2 単語対 GA 方式のアルゴリズム

単語対制約を用いた 2-gram 確率の強化アルゴリズムを具体的に述べる。単語対制約はあらかじめ文法から抽出しておく。単語 w_1, w_2 において、 w_1 が直前単語である時の w_2 の 2-gram 確率を $P(w_2|w_1)$ とする。 w_2 の N-gram 確率を計算する際に、 w_1 との単語接続が単語対制約で受理可能 (acceptable) ならば、0 より大きく 1 以下の係数 C をかけることにより、さらに対数尤度を上げるとい操作を行う。 w_1w_2 という単語接続が文法で受

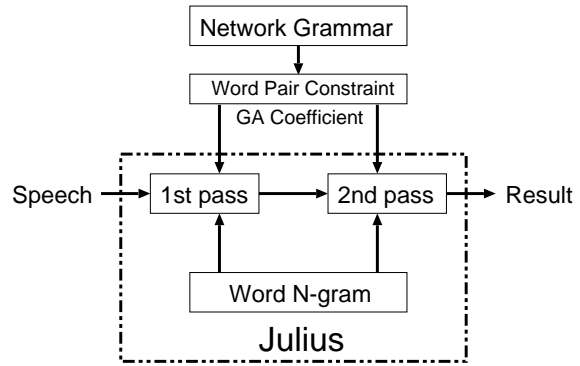


図 2: GA 方式の概念図

理不可能 (non-acceptable) ならば、この操作は行わない。つまり $P(w_2|w_1)$ の接続スコア $S(w_2|w_1)$ を求める式は次のようになる。

$$S(w_2|w_1) = \begin{cases} C \log P(w_2|w_1) & \text{acceptable} \\ \log P(w_2|w_1) & \text{non-acceptable} \end{cases} \quad (0 < C \leq 1)$$

なお、3-gram を用いる場合は 3 単語のうち直前単語のみを用いて接続の可否を判断する。以下、上式における係数 C を GA (Grammar Assisted) 係数、この手法を単語対 GA 方式と呼ぶ。

音声対話システムにおいては、特にキーワードを含むような発話の認識を高精度で行いたい場合が多い。単語対 GA 方式ではそれらの文法をあらかじめ記述しておけば確実に確率を強化することで、部分的なフレーズに対して認識精度を向上させることができる。文法と N-gram モデルのネットワークを結合させるアルゴリズム [4] とは異なり、提案手法では N-gram 頻度情報を用いておらず、また、デコーダ内で動的に確率値を変更しているため、バックオフを意識することなく確率値の変更を行うことができる。

3.3 Julius への実装

単語対 GA 方式を Julius [5] に組み込んだときの概念図を図 2 に示す。Julius では、第 1 パスでは 2-gram 確率、第 2 パスでは逆向き 3-gram 確率を用いて探索を行っているが、GA 方式は第 1 パス、第 2 パスともに独立の GA 係数を用いて処理を行う。また、factoring も全て GA 係数をかけた値で計算する。

さらに、N-gram 上のある単語が文法ではどの単語に相当するかを決定するため、N-gram 上の単語と文法上の単語のマッピングを行う必要がある。今回は N-gram 上と文法上で漢字かな交じりの表記が一致する単語同士についてマッピングを行った。N-gram 上の単語とマッピングを行うために、文法は N-gram モデルに合わせて形態素単位で記述し、1 形態素を 1 単語とした。

さらに、文法内では文字列表記は同じでも複数の単語が複数の文法カテゴリに属することがある。そのため、ある N-gram 上の単語 ID (Word ID や N-gram ID) に対して、その単語の文字列表記と等しいものを持つ文法カテゴリ (Category ID) を全てマッピングし、複数の Category ID を参照できるようにする。そして、それらのカテゴリ間でいずれかの接続が文法で受理可能ならば、その 2 単語は接続可能であると判断する。

4 認識実験

4.1 実験条件

受付案内ロボット ASKA[2] との対話タスクにおいて認識実験を行った。テストセットは奈良先端科学技術大学院大学の学生 58 人 (男性 51 人、女性 7 人) による ASKA への質問文 (1 人あたり 10 文) のうち、タスク内の 384 文を使用する。

言語モデルには Web から収集したコーパスを元に作成した受付案内ロボット対話タスクの単語 N-gram 言語モデル [2] と、形態素単位で記述した Julian 文法を用いた。単語 N-gram 言語モデルの単語数は 2 万語、テストセットに対する未知語率は 1.2 %、テストセットパープレキシティは 28.0 であった。文法の単語数は 433、文法語彙のテストセットに対する未知語率は 18.0 % であった。またテストセットのうち、文法で受理可能な文は 282 文、受理不可能な文は 102 文であった。

音響モデルには日本音響学会の JNAS データ [6] から作成された、3000 状態 64 混合の PTM tri-phone、性別依存モデルを使用した。

デコーダには 2 パス方式の Julius (Ver 3.2) に 3.3 章に従って GA 方式を組み込んだものを使用した。

4.2 実験結果

まず、単語対 GA 方式の有効性を確認するため、Julius の第 1 パスのみに単語対 GA 方式を適用した。GA 係数を 0.05 ~ 1.00 まで 0.05 刻みで変化させたときの GA 係数と単語正解精度の関係を図 3 に示す。

この図において、提案手法を用いない場合の結果は GA 係数が 1 の場合に相当する。文法で受理可能な文 (acceptable)、受理不可能な文 (non-acceptable) とともに GA 係数が 0.7 付近でピークを持っており、単語 N-gram のみを用いた場合に比べて単語正解精度の向上が確認できた。

第 2 パスまで実行した時の最終的な単語正解精度を表 2 に示す。「N-gram のみ」は N-gram モデルのみで認識を行った場合、「文法のみ」は文法のみで認識を行った場合、「単語対 GA」が提案手法である。なお、GA 係数は第 1、第 2 パスとも認識率が最大となる値を 0.05 刻みで探り、その最

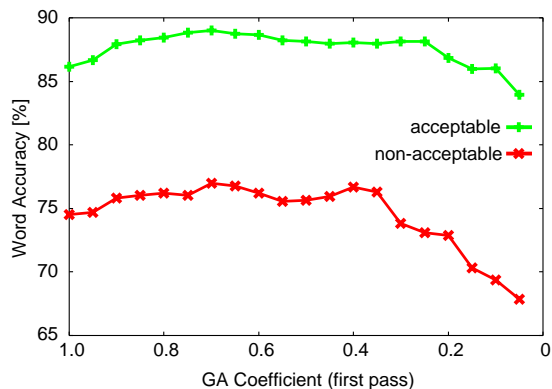


図 3: GA 係数と単語正解精度 (第 1 パス)

表 2: GA 方式の単語正解精度の比較

	受理可能 (282 文)	受理不可能 (102 文)	平均 (384 文)
N-gram のみ	86.1	74.5	82.4
文法のみ	95.9	52.3	81.8
N-best 方式	88.4	75.0	84.0
単語対 GA	92.3	79.1	88.0
単単語 GA	90.1	77.6	86.1

Acc(%)

適値 (それぞれ 0.55) を用いた。また比較のため、N-gram の N-best 結果から文法にマッチする候補があればそれを採用する手法 [3] での値を「N-best 方式」として示している。ここでは、上位 100 候補を利用した。単語対 GA 方式では、単語 N-gram のみを用いた場合よりも平均 5.6 % の向上が見られた。具体的には、単語 N-gram のみを用いた場合「学食は何時までやっていきますか」という発話に対して、「各色を何時までやっていきますか」というように、タスクの想定内であるが N-gram 上で確率が低いために「学食」という単語が出なかったのに対して、単語対 GA 方式では文法による確率強化の結果、正しく「学食」が出力されるという例が多く見られた。

N-gram モデルのみを用いた場合は、受理可能、不可能な文ともある程度の認識率が得られているが、文法のみを用いた場合は、受理可能な文に対しては高精度な認識ができていないものの受理不可能な文では N-gram モデルよりも大きく下回る結果となった。一方、提案手法である単語対 GA 方式では受理可能な文は文法に近い値になっており、受理不可能な文でも N-gram モデルのみを用いた場合よりも高精度で認識できている。また、単純な N-best 方式に比べ、受理可能な文だけでなく、受理不可能な文に対しても認識精度が向上したのは、文全体ではなく、部分的に文法制約が適用され、定型表現が精度よく認識されたためであると考えられる。

4.3 単単語の確率強化との比較

単語対 GA 方式では 2 単語間の接続制約を用いているが、この単語対による確率強化を評価するために、単単語に対する GA 方式の適用を試した。N-gram の探索過程において、直前単語に関わらず、文法語彙に含まれる単単語の出現確率を強化する。以下、この手法を単語対 GA 方式に対して単単語 GA 方式と呼ぶ。

単単語 GA 方式について同様の認識実験を行った。結果を表 2 の最下段に示す。GA 係数は第 1 パスが 0.70、第 2 パスが 0.75 である。単単語 GA 方式では、単語 N-gram のみを使用した場合に比べて平均 3.7 % の向上が見られたが、単語対 GA 方式よりも 1.9 % 低い結果であった。このことから単純に語彙のみを強調するよりも、2 単語間の繋がりを考慮した単語対 GA 方式の方が有効であることが示された。

5 未知語の追加に対する効果

実際のシステムでは、音声対話システムが要求するタスク特有の語彙が認識辞書に含まれない場合、それらは認識できない。そのため、単語辞書にそれらの未知語を追加することにより、タスク特有の単語の認識も可能にする必要がある。

N-gram モデルでは確率の再学習は困難であるため、UNK エントリを未知語に割り当てる。これは学習時の未知語全てを UNK という 1 つの単語とみなしてその出現確率を集約したものである。しかし、単純に UNK に単語を追加すると妥当な確率が割り当てられず、わき出しが起りやすい。

この未知語の追加に対する単語対 GA 方式の有効性を調べるために認識実験を行った。単語 N-gram のみ、単語対 GA 方式、単単語 GA 方式のそれぞれに対して UNK エントリに未知語を追加して認識を行う。当テストセットにおいて、N-gram 語彙には含まれない単語の数は 31 単語で、これらの単語が含まれるテスト文は 37 文であった。これらの単語を文法に組み入れ、N-gram 上では UNK の確率を割り当てる。Julius の第 1 パスでは、UNK に登録した単語中のどの単語であるかを特定できないため、UNK に対しては対数尤度の変更は行わない。その他の実験条件は 4.1 章に等しい。

認識実験の結果を表 3 に示す。GA 係数は、4 章と同じ値を用いた。単単語 GA 方式では、UNK に単語を追加することによって認識率が低下する傾向が見られた。それに対して、単語対 GA 方式では受理可能、不可能な文ともに認識率の低下は見られず、平均の単語正解精度は 0.8 % 向上した。

単語 N-gram のみを用いた場合、UNK エントリの単語が含まれる 37 文のうち、それらの単語が正しく出力された文が 32 文であったことから、

表 3: 未知語を追加した場合の認識率の比較

言語モデル	受理可能 (282 文)	受理不可能 (102 文)	平均 (384 文)
N-gram のみ	86.1	74.5	82.4
N-gram のみ+UNK	86.9	75.1	83.1
単語対 GA	92.3	79.1	88.0
単語対 GA+UNK	93.1	79.9	88.8
単単語 GA	90.1	77.6	86.1
単単語 GA+UNK	90.1	77.4	86.2

Acc(%)

表 4: サブタスクの分類

サブタスク番号	文の数	内容
1	21	ASKA 自身に関する質問
2	150	学内及び周辺の施設・設備
3	54	先生方の居室と内線番号
4	18	研究内容に関わる講座名
5	59	電車やバスの発車時刻
6	42	天気予報
7	40	今日のニュース

単純に UNK エントリに未知語を追加するだけで大部分は認識可能となった。ただし、未知語の追加によるわき出し誤りが見られた。一方、単語対 GA 方式では UNK エントリの単語に対しても単語間の接続を考慮した確率の付与がなされ、わき出し誤りが抑制された。

6 サブタスク単位の確率強化

対話音声認識システムでは、様々なタスクの発話に対する認識を行う場合が考えられるが、単一の汎用 N-gram モデルだけでは十分な認識率が得られにくい。しかし、タスクごとに異なる文法を適用すると、システムがタスクの選択を誤った際には容易に認識に失敗する。また、話題ごとに N-gram モデルを用意するのも煩雑である。そこで、話題に応じて文法を切り替えて GA 方式を適用することを考える。

文法の切り替えによる影響を調べるため、今回は受付案内ロボット対話タスクの文法を表 4 に示す 7 つのサブタスクに細分化し、話題にマッチした文法で N-gram 確率を強化した場合と、話題に対してミスマッチな文法が選択された場合の認識結果を比較した。実験条件は 4.1 章と同じである。

実験結果を表 5 に示す。話題にマッチしたサブタスク文法を用いたとき、平均の単語正解精度が 88.8 % であり、タスク全体を表す単一の文法を用いる場合に比べて 0.9 % 向上した。一方、話題にミスマッチなサブタスク文法が選択された場合は、サブタスクによっては認識精度の大幅な低下が見られたものの、N-gram モデルのみを用いた場合とおおよそ等しい結果になっている。つまり、話題に応じて文法を切り替え、単語対 GA 方式を適

表 5: サブタスクごとの認識結果

タスク\文法	N-gramのみ	分割前の文法	1	2	3	4	5	6	7
1	71.8	87.2	88.9	69.2	67.5	70.9	73.5	70.9	68.4
2	86.7	92.1	87.0	92.9	87.1	86.0	86.4	86.8	87.2
3	83.7	87.1	85.3	87.8	87.8	85.7	83.5	83.1	83.5
4	84.8	89.1	84.8	88.1	86.7	87.6	86.2	84.3	85.2
5	78.1	85.8	80.2	77.1	79.4	77.7	86.1	79.3	77.0
6	77.6	83.9	74.8	74.1	78.0	75.9	75.2	86.0	76.9
7	80.1	82.0	80.4	76.9	79.4	79.4	78.5	79.8	84.2

Acc(%)

表 6: モデル GA 方式との認識率の比較

	受理可能 (282 文)	受理不可能 (102 文)	平均 (384 文)
N-gramのみ	86.1	74.5	82.4
単語対 GA	92.3	79.1	88.0
モデル GA	91.2	77.0	86.6

Acc(%)

用することにより、話題同定の失敗に対するリスクを押さえつつ、成功した場合にはさらに高精度な認識が可能となる。

7 モデルに対する静的な N-gram 確率の強化

単語対 GA 方式は、デコーダに処理を組み入れることにより、探索途中で動的に N-gram 確率の強化を行っている。これに対して、言語モデル中の N-gram 確率をあらかじめ静的に強化しておくことも可能である。この手法では、デコーダを改変する必要がないため、内部処理の追加による認識時間の増加を防ぐことができる。

ただし、本手法ではモデルに対して静的に確率値を変更するため、1-gram 確率に対しては直前の単語を特定できず、単語対制約を組み込めない。よって、ここではその単語が文法語彙に含まれれば確率の強化を行う。つまり、バックオフ処理では文法で受理不可能な単語接続に対しても強化される可能性があり、正しい確率値の変更が行えない。また、Julius の第 1 パスと第 2 パスで異なる GA 係数を用いることができないという制限がある。

単語対 GA 方式との比較を行うため、受付案内ロボット対話タスクにおいて認識実験を行った。実験条件は 4 章と同一である。認識結果を表 6 に示す。「モデル GA」が静的に N-gram 確率を強化した時の値である。この結果より、モデル GA 方式はデコーダに組み込む単語対 GA 方式と同等の効果は得られないものの、N-gram モデルのみを用いた場合よりも平均 4.2 % の単語正解精度が向上した。このことから、既存の音声対話システムにおいても静的に N-gram 確率を強化したモデルを

用いることは有効であると言える。

8 まとめ

音声認識エンジン Julius における探索過程で、文法の単語対制約を用いて部分的に言語確率を強化する認識アルゴリズムを提案した。音声認識実験により、単語 N-gram 言語モデルのみを使用した場合よりも 5.6 % の単語正解精度が向上し、提案手法の有効性が示された。本手法では、文法で受理可能な文、不可能な文にかかわらず認識精度が向上することを示した。これは文全体ではなく部分的に文法制約が適用されたためであると考えられる。さらに、提案手法は未知語の追加に対して頑健であり、認識率の低下が押さえられていることが分かった。文法のミスマッチに対しても N-gram のみを用いた場合とほぼ等しい認識率が得られた。言語モデルに対して確率を強化する手法では平均で 4.2 % の単語正解精度が向上し、GA 方式を静的に言語モデルに適用した場合にも効果があることを示した。

今後の課題として、話者ごとの文法切り替えや、別タスクでの評価、単語対 GA 方式を組み込んだ Julius の高速化が挙げられる。

参考文献

- [1] 西村 他: 生駒市コミュニティセンター音声情報案内システムの開発と運用. 情処学研報, 2003-SLP-45, 2003
- [2] R. Nishimura et al.: ASKA: Receptionist Robot with Speech Dialogue System. IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS2002), 2002
- [3] 鶴身 他: 単語 N-gram とネットワーク文法を併用した音声認識アルゴリズムの検討. 音講論集, pp145-146, 2002.9
- [4] 秋葉 他: 部分 N-gram 頻度情報を利用した質問応答定型表現への言語モデル適応. 情処学研報, 2002-SLP-42-6, pp.31-38, 2002
- [5] A. Lee et al.: Julius - An Open Source Real-Time Large Vocabulary Recognition Engine. Proc. EUROSPEECH2001, pp.1691-1694, 2001
- [6] 河原 他: 連続音声認識コンソーシアム 2000 年度版ソフトウェアの概要と評価. 情処学研報, 2001-SLP-38-6, pp.37-42, 2001