

AdaBoost を用いたシステムへの問い合わせと雑談の判別

佐古 淳[†] 滝口 哲也^{††} 有木 康雄^{††}

[†] 神戸大学大学院自然科学研究科 〒657-8501 兵庫県神戸市灘区六甲台町 1-1

^{††} 神戸大学工学部 〒657-8501 兵庫県神戸市灘区六甲台町 1-1

E-mail: †sakoats@me.cs.scitec.kobe-u.ac.jp, ††{takigu,ariki}@kobe-u.ac.jp

あらまし 音声インターフェイスとして用いる際、認識結果中の単純なキーワードを基にシステムへの要求を推定することには問題がある。これだけでは、キーワードを例に使い方を説明するような発話でも動作してしまったり、関係のない雑談からキーワードがわき出ただけでも動作してしまう。一方で、固定されたネットワーク文法から要求を推定することにも問題がある。柔軟な表現による要求が困難になり、システムに慣れないと思いつ通りに操作することができなくなる。本研究では、AdaBoost を用いて、音声認識による言語情報からシステムへの要求か否かの判別を行う。認識誤りを考慮して Confusion Network を構築し、深い認識仮説まで用いて判別を行う。単語正解精度 42.1% のタスクにおいて、システム要求判別の適合率 93.8%、再現率 93.8% を実現した。

キーワード AdaBoost, システム要求判別, Support Vector Machine, 音声認識

System Request Discrimination Based on AdaBoost

Atsushi SAKO[†], Tetsuya TAKIGUCHI^{††}, and Yasuo ARIKI^{††}

[†] Graduate School of Science and Technology, Kobe University Rokkodaicho 1-1, Nada-ku,
Kobe, Hyogo, 657-8501 Japan

^{††} Faculty of Engineering, Kobe University Rokkodaicho 1-1, Nada-ku, Kobe, Hyogo, 657-8501
Japan

E-mail: †sakoats@me.cs.scitec.kobe-u.ac.jp, ††{takigu,ariki}@kobe-u.ac.jp

Abstract It is difficult to discriminate a request to a system from a chat based on only simple keywords, because keywords are sometime included in utterances of explaining an usage of a system, or appear by misrecognition in chats. On another front, it has troubles to use network grammar, because it makes the usage difficult for people unaccustomed to a system due to fixed expressions of system request. In this paper, we propose a discrimination method based on AdaBoost using linguistic information obtained from speech recognition. To solve a problem of misrecognition, a confusion network is employed. The confusion network permits deep hypothesis of speech recognition. The experimental results showed that 95.7% of precision and 89.8% of recall on a task which is recognized with 42.1% of word accuracy.

Key words AdaBoost, System Request Discrimination, Support Vector Machine, Speech Recognition

1. はじめに

近年、様々な分野で音声によるインターフェイスが実用化されつつある。特に、ロボットとのコミュニケーションや、カーナビのように手を使うことが困難な機器の操作への適用が顕著である。現在の音声インターフェイス

は、以下の2つに大別できる。

- キーワードの有無によるもの
- ネットワーク文法によるもの

しかし、これらの方法では問題が生じる場合が多い。キーワード方式では、例えば機器の使い方を説明するような場合にキーワードが現れると、これをシステムへの要求

として認識してしまうことがある。また、実際にはキーワードを発話していない場合でも、音声認識の誤認識によりキーワードがわき出し、このために誤動作してしまう場合がある。ネットワーク文法方式では、キーワード方式に比べ、誤動作のわき出し（実際にはシステムへの要求がないにも関わらず誤動作してしまうこと）を減らすことができる。しかし、一方で、人手によって詳細なネットワーク文法を構築するには大きなコストがかかることや、システム要求発話の柔軟性・多様性が失われてしまい、決まった発話を暗記しないとシステムを操作できない、などの問題が生じる。

これらの問題に対し、システム要求発話かどうかを見分ける手法として、発話者の視線方向を用いるもの [1] や、発話の音響的な特徴を用いるもの [2], [3] が提案されている。視線方向を用いた手法は、ロボットと発話者の距離が離れており視線方向検出が困難な場合や、カーナビ操作のように視線をシステムへ向けることが危険な場合には利用することができない。音響的な特徴を用いた手法では、そもそも音響的な特徴だけでは判別が困難な場合が存在することや、システムが高度化・高性能化するにつれて、さらにはユーザーがシステムに習熟するにつれて、システムに対する発話が人に対する発話に近づき、判別が困難になるといった問題がある。

本稿では、音声認識によって得られた言語情報に対し、AdaBoost によってシステム要求発話を判別する手法について述べる。AdaBoost を用いることにより、ラベル付きコーパスから自動的に、システム要求発話を判別するための素性を得ることができる。また、単純なキーワード方式よりも頑健でわき出しに強く、柔軟に多様性のある発話を判別することが可能である。本研究では、音声認識結果を扱うことから、認識誤りが原因となる判別誤りが想定される。このことから、単純に 1-Best の認識結果を用いるのではなく、Confusion Network によって、より深い仮説まで用いた AdaBoost についても検討を行った。

以下、次章では本研究で用いたシステム要求判別タスクの概要について述べ、3. 章で提案手法について述べる。4. 章で評価実験と考察について述べ、5. 章でまとめる。

2. システム要求判別タスク

本章では、本研究で用いたシステム要求判別タスクの概要について述べる。本タスクでは、まず、二人以上の人間とシステムが同時に存在することを想定する。これは、ロボットを操作する際に周囲に人がいる場合や、カーナビを操作する際に助手席に同乗者がいる場合のように、自然な状況であると考えられる。二人以上の人間が互いに会話を行いながら、任意にシステムへの要求発話を行う。本研究では、“システム”として、図1のようなロボットを用いた。ロボットの機能を表1にまとめる。典型的な

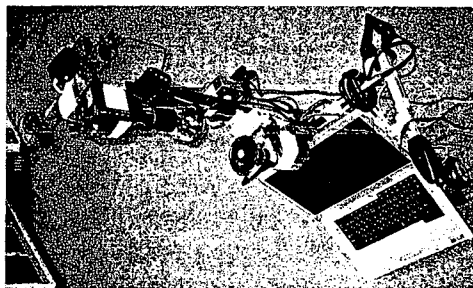


図1 本研究で用いたロボット
Fig. 1 An image of the robot.

利用方法としては、少し離れた場所から「こっちに来て」とロボットを呼ぶ、「写真を撮って」と写真を撮ってもらう、などがある。

現状でロボットが受理できるコマンドは表1の通り決まった文章のみである。ただし、本研究のために収録した音声には、「こっち来て」「早く来いよ」「あっち行って」のように、ロボットは受理できないものの、同じ動作を期待する表現も含まれており、これらもシステム要求発話であるとのラベルを付与した。システム要求ではない発話としては、通常の雑談に加え、「こっちに来て、とか言うて……」「こっちに来て、向こうへ行ってだけでは……」のようにシステムへの要求発話を含むような発話も含まれている。

収録は、二人の発話者それぞれの胸元に取り付けたマイクで行った。発話数は330で、内49発話がシステム要求発話であった。書き起こしは、ひらがな、及びかな漢字混じりの二種類作成した。ひらがなのものは音響モデル適応である。

ロボットを用いた従来研究では、視線情報を用いるものも存在する。ただし、本研究では、発話者とロボットとの距離が離れており視線方向の推定が困難であったこと、また、発話者がロボットを見ずにシステム要求発話を行う場合もあること、将来的にロボットではなく、カーナビでの利用も想定していることから、システム要求発話の判別に、視線情報は用いない。

3. 提案手法

本章では、AdaBoost を用いたシステム要求判別手法について述べる。まず、AdaBoost について概説し、その後、音声認識結果から 1-Best ではない深い仮説を用いる手法として、Confusion Network について述べる。最後に 3.3 節で、AdaBoost を用いた場合と SVM を用いた場合について比較・検討を行う。

3.1 AdaBoost

AdaBoost は、いくつもの識別器を組み合わせると一つの高度な識別器を構成する *ensemble learning method*

表1 本研究で用いたロボットの機能

Table 1 Abilities of the robot.

機能	CSP による音源到来方向推定 音源方向/反対方向への移動 障害物の回避 アームによるボトルの設置 写真の撮影
コマンド例	こっちに来て 向こうへ行って 写真を撮って ついて来て ボトルを置いて

のひとつである。Schapire ら [4] が提案している学習のアルゴリズムを図2に示す。図中、 I は、 $I(true)$ ならば1, $I(false)$ ならば-1となる。 ϵ_t が0.5未満の弱学習器を見つけ続けることができれば、学習誤差0の最終学習仮説を生成できる。また、未知のサンプルに対する汎化誤差も小さくできることが実験的に報告されている [5], [6]。一方、雑音を有するサンプルの場合、過学習を起こすことが報告されている。これに対しては、AdaBoostの学習過程をマージン最大化ととらえ、SVMにおけるSoft Marginsの概念を導入した手法も提案されている [7], [8]。本研究では、認識結果を扱うため、サンプルには多くの雑音に乗っているものと考えられる。このことから、通常のAdaBoostではなく、Soft Margins付きのAdaBoostを用いることとした。

AdaBoostを用いたテキスト分類手法としては、文献 [4], [9] などが提案されている。これらの文献では、テキスト分類のための弱学習器として、Decision Stumpsが用いられている。Decision Stumpsとは、ある素性の有無に基づいて分類を行う単純な手法である。素性には、単語や単語 bi-gram, ラベル付き順序木などが用いられる。学習時には、学習サンプルを最もうまく分類するような“素性”を選択し、その際の重みを得る。識別時には、学習によって得られた全ての素性について、サンプル中にその素性があれば、クラス y に重み α の投票を行うということを繰り返し、最終的に重みの大きかったクラスと判別する。

3.2 Confusion Network

本研究では、音声認識結果からのシステム要求判別を行う。そのため、音声認識における認識誤りが問題となることが考えられる。そこで、1-Bestの認識結果を用いるのではなく、より深い仮説を用いる。深い仮説を用いる手法としてN-Best, ワードグラフなどが存在するが、本研究では、情報をコンパクトに表現できる Confusion Networkを用いる。Confusion Networkはワードグラフを圧縮したものであると考えられる。Confusion Networkの具体例を図3に示す。

Input: n examples $Z = \{(x_1, y_1), \dots, (x_n, y_n)\}$

Initialize: $w_1(z_i) = 1/n$ for all $i = 1 \dots n$

Do for $t = 1, \dots, T$,

(1) Train a base learner with respect to weighted example distribution w_t and obtain hypothesis $h_t : x \mapsto \{-1, 1\}$

(2) Calculate the training error ϵ_t of h_t :

$$\epsilon_t = \sum_{i=1}^n w_t(z_i) \frac{I(h_t(x_i) \neq y_i) + 1}{2}$$

(3) Set

$$\alpha_t = \log \frac{\epsilon_t}{1 - \epsilon_t}$$

(4) Update example distribution w_t :

$$w_{t+1}(z_i) = \frac{w_t(z_i) \exp\{\alpha_t I(h_t(x_i) \neq y_i)\}}{\sum_{j=1}^n w_t(z_j) \exp\{\alpha_t I(h_t(x_j) \neq y_j)\}}$$

Output: final hypothesis:

$$f(x) = \frac{1}{\|\alpha\|_1} \sum_{t=1}^T \alpha_t h_t(x)$$

図2 AdaBoostのアルゴリズム

Fig. 2 AdaBoost algorithm.

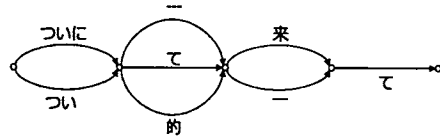


図3 Confusion Networkの具体例

Fig. 3 An example of Confusion Network.

3.3 Support Vector Machines との比較

SVMは、図4のようにクラス間のマージンを最大化するように識別境界を推定する手法である。また、Kernel法と組み合わせることにより、非線形の識別境界も取り扱うことが可能であり、高い汎化能力と識別性能を持つことが知られている。本研究においても、SVMを用いてシステム要求発話を判別することが可能である。

SVMでは、サポートベクターと呼ばれる一部の学習データ(事例)によって識別境界が決まる。言い換えれば、SVMは、クラス間のマージンを最大化するようなサポートベクターを推定する手法と言える。テキスト分類問題を考えた場合、サポートベクターは“文章”として求まる。すなわち、SVMでは、“文章”を基にシステム要求発話の判別を行う。

一方、AdaBoostは、“素性”を基に判別を行う手法である。AdaBoostによる学習では、“文章”ではなく、“素性”となる単語や単語 bi-gramなどが求まる。

このことは、2つの意味を持っていると考えられる。ひとつは、AdaBoostでは素性を基に判別を行うため、得られた素性から、学習データ中に存在しない事例につい

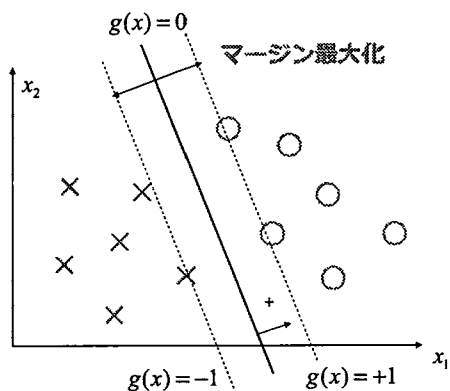


図4 SVMのマージン最大化
Fig.4 SVM: Margin maximization.

でも判別を行うことができるということである。例えば、「こっちに来てとか」という学習データから、「とか → 雑談」という素性が得られたとする。このとき、「向こうへ行ってとか」という発話学習データ中に存在しない場合でも、これを正しく雑談と判別できる可能性が高い。

もうひとつは、AdaBoostの方がキーワード・スポットティングとの相性がよいことである。本研究のように、雑談を取り扱う場合、未知語は判別性能を低下させる大きな問題となる。そこで、キーワード・スポットティングを用いて、未知語に対処することが考えられる。通常であれば、どの単語をキーワードとし、どの単語をガベージとすればよいかは人手で判断する必要がある。しかし、AdaBoostによって素性を学習した場合、判別に必要なキーワードは素性語であり、それ以外は判別に不必要なガベージと考えることができる。これに対し、SVMでは、素性となる単語ではなく、文章のレベルで学習がなされるため、サポートベクターとなる文章の中から関係のありそうな単語を選択するか、文章中に含まれる全ての単語をキーワードとする必要がある。

4. 実験

本章では、提案手法を用いたシステム要求発話の判別実験について述べる。実験は大きく分けて2つ行った。

(1) クリーンな書き起こしテキストを用いたシステム要求発話の判別 (判別性能の上限)。

(2) 音声認識結果を用いたシステム要求発話の判別。以下、それぞれの実験結果について述べる。

4.1 クリーンな書き起こしテキストによるシステム要求判別

まず、提案手法の判別性能の上限値として、クリーンな (認識誤りの含まれていない) 書き起こしテキストを用いたシステム要求発話の判別実験を行った。かな漢字

表2 クリーンな書き起こしテキストによるシステム要求判別結果

Table 2 Results of system request discrimination using clean transcription.

	適合率	再現率	F 値
AdaBoost (uni-gram)	94.0%	95.9%	0.95
SVM	97.9%	95.9%	0.97

混じりの書き起こしテキストを Mecab [10] によって形態素に分割し、AdaBoost, SVM でそれぞれ学習を行った。なお、SVM においては単語の出現頻度をベクトル化し、特徴とした。また、カーネル関数には Gauss Kernel を用いた。判別実験は、コーパスの分量の不足もあり、one-leave-out 法により行った。ここで用いた素性は単語 uni-gram である。実験結果のうち、F 値が最も高かったケースを表2に示す。表の通り、SVMの方がAdaBoostよりも高い識別性能を示している。特に、SVMでは、システム要求のわき出しを抑えられる傾向がみられた。

4.2 音声認識結果からのシステム要求判別

次に、音声認識結果に対し、システム要求発話判別実験を行った。まず、音声認識の条件と結果について述べる。

4.2.1 音声認識条件と結果

ベースラインの音響モデルは、日本語話し言葉コーパス (CSJ: Corpus of Spontaneous Japanese) モニター版 [11] のうち、男性話者 200 名の講演音声を用いて作成した。音響分析条件と HMM の仕様を表3に示す。これらの条件で音響モデルを作成し、さらに、MLLR+MAP [12] により音響モデル適応を行った。音響モデル適応は、テストセットを含めた適応をクロズド、含めない適応をオープンとした。ただし、どちらの場合も適応データの話者は、テストセットと同一のものを用いた。適応データの分量は、クロズドの場合で約 10 分、オープンの場合で約 5 分であった。

言語モデルは、実験で用いた発話を書き起こしたテキストから作成した。ただし、テストセットに対しオープンとなるように、話者 B の発話を用いて話者 A の認識用言語モデルを作成した。

音声認識による単語正解精度を表4に示す。音響モデルをオープンにした場合は単語正解精度が 30% を下回ってしまう。将来的に、適応に用いるデータを増やした際、オープンでもクロズド程度の性能に近づくことを期待して、音響モデルはクロズドの方を用いることとした。

4.2.2 音声認識結果を用いたシステム要求判別

前節によって得られた認識結果に対し、提案手法によるシステム要求判別を行った。学習とテストはクリーンな書き起こしテキストの場合と同様に one-leave-out 法により行った。なお、本節の実験では Confusion Network は用いず、1-Best の結果から判別を行った。Confusion Network を用いた場合の結果は次節で示す。素性は、Ad-

表 3 音響分析条件と HMM の仕様

Table 3 Condition of acoustic analysis and HMM specification.

音響分析	サンプリング周波数	16kHz
	特徴パラメータ	MFCC(25 次元)
	フレーム長	20ms
	フレーム周期	10ms
	窓タイプ	ハミング窓
H	タイプ	244 音節
M	混合数	32 混合
M	母音 (V)	5 状態 3 ループ
M	子音+母音 (CV)	7 状態 5 ループ

表 4 音声認識結果の単語正解精度.

Table 4 Word accuracy of the speech recognition results.

	音響モデル	
	オープン	クローズド
単語正解精度	29.8%	42.1%

表 5 音声認識結果を用いたシステム要求判別結果

Table 5 Results of system request discrimination using recognized transcription.

	適合率	再現率	F 値
AdaBoost (uni-gram)	91.8%	91.8%	0.92
AdaBoost (bi-gram)	93.8%	91.8%	0.93
SVM	97.8%	91.8%	0.95

aBoost では uni-gram, 及び bi-gram を用いた. SVM ではクリーンな書き起こしテキストの場合と同様に, 単語の出現頻度をベクトル化したものを特徴として用いた. 結果を図 5 に示す. また, F 値が最も高かったケースを表 5 に示す. 全体としての識別性能は SVM が優れている. 特に, わき出しに強い特性を持っていることがわかる. これは, 文書全体から識別を行うため, システム要求発話に見られるような単語が混じり込んでも, それに大きな影響を受けないためと考えられる. ただし, この点については, 4.3 節で述べるような欠点も存在する.

一方, AdaBoost の利点として, 素性語が明確に得られることがあげられる. このときに選択された素性語を表 6 に列挙する. システム要求に投票を行う素性には bi-gram が多く選択されていることがわかる. これは, システム要求発話がある程度決まったフレーズによって行われているためと考えられる. また, <s> や </s> との組み合わせが多数存在することから, システム要求発話は, 一旦会話を区切った上で為されているものと考えられる. AdaBoost において, uni-gram を用いるよりも bi-gram を用いる方が高い性能を示した. 特に, bi-gram を用いることでわき出しを減らすことができていた.

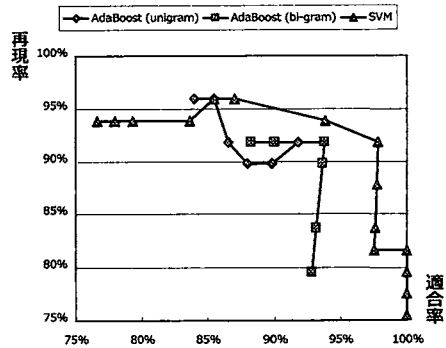


図 5 音声認識結果を用いたシステム要求判別結果

Fig. 5 Result of system request discrimination using recognized transcription.

表 6 AdaBoost によって選択された素性語

Table 6 Selected features by AdaBoost.

システム要求	雑談
<s>+ここ <s>+向こう <s>+写真 ください+</s> ここ+は+</s> て+ください 場所 来+て 来い+</s> etc.	あーこれららつて とかないなんかまあ やつ 旨っ etc.

4.2.3 Confusion Network の利用による判別性能の改善

本節では, Confusion Network を用いた場合のシステム要求判別について述べる. 簡単のため, AdaBoost に対してのみ考察を行う. まず, 前節において, 音声認識誤りが原因となって判別に失敗していた例として, 「ついでに 来て → 雑談」というものがあった. これは, 本来, 「ついで」と発話しているところを「ついでに」を誤って認識したために判別に失敗した例である. この発話について, Confusion Network を構築した例が 3.2 節で取り上げた図 3 である. 2 番目のパスに正解単語である「ついで」が含まれていることがわかる. このようなケースにおいて, Confusion Network を用いることで精度が改善すると期待される. 表 7 に, Confusion Network を用いた場合の結果を示す. 表中の 1-Best は, 前節で F 値が最も高かった場合である. Confusion Network を用いて深いパスまでを対象とすることにより, 判別性能が改善できることがわかる.

4.2.4 システムに要求された動作の種別認識

本稿では, これまで, システム要求発話であるか否かの判別について述べてきた. しかし, 通常は, システム要求発話であると判別した上で, さらにどの動作を望んでいるかまで認識する必要がある. 本節では, AdaBoost を利用して, 動作の種別まで認識した場合の実験結果につ

表 7 Confusion Network による判別性能の改善
Table 7 Improvements of the accuracy using Confusion Network.

	適合率	再現率	F 値
1-Best	93.8%	91.8%	0.93
Confusion Network	93.8%	93.8%	0.94

表 8 未知の「システム要求発話+雑談素性」の判別結果
Table 8 Results of unseen chat discrimination.

	AdaBoost (uni-gram)	SVM (uni-gram)
識別率	100.0%	21.4%

いて簡単に述べる。学習は、one-vs-rest 法を用いて行った。これは、注目する動作とそれ以外の識別器を、動作の数だけ構成する手法である。動作の種類は、呼ぶ・遠ざける・写真撮影・連れる・ボトル操作・場所記憶・停止・初期化の 8 種類である。

実験の結果、システム要求判別性能は維持したまま、高精度な動作の種類認識ができることがわかった。動作の数が少なく、似通った文章による依頼発話が少ないこともあり、動作の種類認識性能は 100%であった。

4.3 素性に基づく判別の効果

最後に、学習データ中に存在しない雑談発話の認識について行った実験について述べる。まず、クリーンな書き起こしテキストを用いて識別器を構成する。次に、「システム要求発話+雑談素性」となるような文章が正しく雑談と識別されるかについて調べた。具体的には、「向こうへ行つてとか」「ついて来てって言う」となどの文章である。同一のフレーズが学習コーパス中には存在しない文章のみを選択した。AdaBoost と SVM による結果を表 8 に示す。

この結果、AdaBoost の方が SVM よりも高い性能を示した。これは、AdaBoost が「素性」に基づいて判別を行うため、雑談素性に大きな影響を受けるためと考えられる。SVM は、「文章」全体に基づいて判別を行うことから、雑談素性が付け加わっただけでは雑談と判別されにくいと考えられる。

5. ま と め

本稿では、AdaBoost を用いたシステム要求判別手法について述べた。AdaBoost を用いることで、システムへの要求発話と雑談とを 9 割以上の割合で、高精度に判別することが可能であった。判別性能の点では Support Vector Machines に及ばないものの、ほぼ同等の性能を持つものと考えられる。また、AdaBoost 独自の特徴として、素性として用いた単語が明確に得られることからキーワード・スポットティングへの応用が容易であることがあげられる。さらには、「とか」「って言う」とのように、端的に雑談であることを示す素性を学習することに

より、学習コーパス内に「システム要求発話+とか」のような文章が現れない場合でも、正しく雑談発話と判別することが可能であった。

今後の課題として、大規模なコーパスを構築し実験すること、タスクの難易度を上げることがあげられる。また、今回は言語情報のみから判別を行ったが、音響情報などと組み合わせて判別を行う手法についても検討を行う必要がある。

文 献

- [1] 堀内靖雄, 庵原彩子, 西田昌史, 市川薫, “自然対話における聞き手の反応と話し手のうなずき・言語情報・韻律情報との関係に関する予備的検討,” 情処学研報, SLP-52, pp.93-98, July 2004.
- [2] 伊藤敏彦, 山田 真也, 荒木健治, 音声対話認識率や状況の違いによる音声対話の言語的・音響的特徴の比較, 情報処理学会研究会報告 2005-SLP-56, pp.101-106(2005 年 5 月).
- [3] 杉本夏樹, 北岡教英, 中川聖一, 音響特徴を用いた対システム発話と対人間発話の識別, 電子情報通信学会, 総合大会, D-14-9, pp.133 (2006.3)
- [4] R.Schapire, Y.Freund, P.Bartlett, and W.Lee, “Boosting the margin: A new explanation for the effectiveness of voting methods,” Annals of Statistics, vol.26, no.5, pp.1651-1686, Oct. 1998.
- [5] Y.Freund and R.Schapire, “Experiments with a new Boosting algorithm,” Proc. 13th International Conference on Machine Learning Bari, Italy Morgan Kaufmann, pp.148-146, July 1996.
- [6] H.Schwenk and Y.Bengio, “Adaboosting neural networks,” Proc. ICANN'97, vol.1327 of LNCS Berlin Springer, pp.967-972, Oct. 1997.
- [7] G.Ratsch, T.Onoda, and K.-R. Muller, “Soft Margin for AdaBoost,” Machine Learning, vol.42, no.3, pp.287-320, March 2001.
- [8] 小野田崇, “Boosting の過学習とその回避,” 電子情報通信学会論文誌, Vol.J85-D2, No.5, pp. 776-784, 2002 年 5 月.
- [9] 工藤拓/松本裕治, “半構造化テキストの分類のためのブースティングアルゴリズム,” 情報処理学会論文誌, Vol.45, NO.9, 2004 年 9 月.
- [10] “MeCab: Yet Another Part-of-Speech and Morphological Analyzer,” <http://chasen.org/taku/software/mecab/>
- [11] 古井貞照, 前川喜久雄, 伊佐原均, “『話し言葉工学』プロジェクトのこれまでの成果と展望,” 第 2 回話し言葉の科学と工学ワークショップ, pp.1-6, 2002.
- [12] 緒方淳, 有木康雄, “音素事後確率に基づく信頼度を用いた音響モデルの教師なし適応,” 信学技報, SP2001-105, 2001.