

複数の言語情報を用いたCRFによる音声認識誤りの検出

松本 智彦[†] 佐古 淳[†] 滝口 哲也^{††} 有木 康雄^{††}

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

^{††} 神戸大学自然科学系先端融合研究環 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

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

あらまし 近年, 音声認識誤りの検出訂正には, 誤り傾向を示す n-gram を学習する「誤り傾向学習」という手法が注目されている. この手法には, 出現頻度の低い n-gram に対しては適切に学習を行うことができないという問題があると考えられる. また, 周辺の単語との意味的類似度をスコア化し, 内容語の認識誤りを検出する手法が提案されている. 本稿では, 特に後者の意味スコアについて検討し, これらの情報を統合した音声認識誤りの検出法を提案する. 日本語話し言葉コーパスによる評価実験の結果, 「誤り傾向学習」で用いる素性に意味スコアを追加することで, 特に内容語において誤り検出性能の改善が得られた. また, 単語重みと組み合わせることで意味スコアを有効に活用することができた.

キーワード 音声認識誤り検出, 信頼度, 言語情報, 意味情報, CRF

Error Detection in Speech Recognition using CRF Based on Various Linguistic Features

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

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

^{††} Organization of Advanced Science and Technology, Kobe University
Rokkodaicho 1-1, Nada-ku, Kobe, Hyogo, 657-8501 Japan

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

Abstract Recently, a learning method of n-gram showing error tendency is focused on. In this method, it is difficult to learn low frequency n-gram appropriately. On the other hand, error detection method for content words was proposed that uses semantic similarity in neighboring recognition result. In this paper, we especially examine semantic score and propose a method for speech recognition error detection that integrates these information. As a result of evaluation experiment using the Corpus of Spontaneous Japanese, error detecting performance was improved for content words in particular. Moreover, semantic score is effectively used by combining with term-weight.

Key words Speech Recognition Error Detection, Confidence Measure, Linguistic feature, Semantic feature, CRF

1. はじめに

近年, 音声メディアの増大に伴い, 音声認識技術は様々な場面で使用されるようになってきた. 例えば, 一般的な音声検索ではユーザのクエリに対して, 発話文書に

対する音声認識結果を参照することで検索結果を出力する. このとき音声認識結果に認識誤りがあると, ユーザが期待していない検索結果が出力されてしまうという問題がある. そこでユーザの負担を軽減するために, あらかじめ自動的に認識誤りを検出し訂正しておくことが好ましい.

近年、音声認識誤りの検出、訂正手法として、誤り傾向を示す言語情報、特に n-gram を Conditional Random Field(CRF) により学習する手法が提案されている [4] [5]。CRF では複数の素性を柔軟に取り込むことができるという利点があり、SVM やパーセプトロンといった識別器を用いた場合よりも良好な結果が得られたということが報告されている。しかしこれらの手法では、学習に音声認識結果と正解文書が必要となるため、大量に学習データを用意することが困難である。そのため、出現頻度の低い単語を含む n-gram については、適切に重みを学習することが困難であるという問題がある。

また、意味情報を用いた誤り検出手法として、周辺の認識結果単語との意味的類似度を用いた手法が提案されている [6]。この手法は、意味を持った内容語に対して有効であるが、機能語のような単語と共起しても不自然でない単語に対しては、単語間類似度を定義することが困難であるため用いることができない。

本稿では、特に意味情報を用いた手法について検討し、これらの情報を統合することで、誤り検出性能の改善を確認したので報告する。

以降の 2 章では音声認識スコアと言語情報を用いた音声認識誤り検出手法、3 章では意味情報を用いた誤り検出手法について述べる。4 章で誤り検出モデルについて述べ、5 章で、評価実験の条件と結果を報告し、最後に 6 章で、結論と今後の課題について述べる。

2. 誤り傾向学習

本章では、誤り傾向学習で従来用いられている音声認識スコア、誤り傾向を示す言語情報について述べる。

2.1 音声認識スコア

まず、基本的な認識誤りとして、雑音の存在や不明瞭な発話によるものが考えられる。このような誤り部分では音響スコアが低くなったり、時間的な競合候補単語が多くなったりしている可能性が高い。また、誤り部分では滅多に出現しないような n-gram が出現し、言語スコアが低くなっているということも考えられる。誤り検出では、単語グラフなどからこれらの情報を用いて各単語の信頼度を求め、信頼度の低いものを誤りとすることが多い。

また、n-gram モデルでは認識誤りから復帰するような情報は学習されていないため、認識誤りが起こると誤りが伝搬するということが報告されている [12]。そのため、前の単語が誤りであれば次の単語も誤りである可能性が高く、周辺単語の音声認識スコアを用いることも有効であると考えられる。

2.2 誤り傾向を示す言語情報

通常、大語彙連続音声認識では統計的言語モデルである n-gram モデルを用いる。統計的言語モデルは正しく書かれたテキストのみから学習を行っており、自然な n-gram

については学習が行われるが、不自然な n-gram については学習を行うことができない。例えば、日本語において「の-よう-を」のような n-gram が出現することはないと考えられる。しかし n-gram モデルでは、このような不自然な n-gram の確率についても、低次の n-gram からスムージングにより推定するため、認識結果に出現してしまうことがある。そこで誤り傾向学習では、音声認識結果と対応する正解文書を用いることで、不自然な n-gram についても学習を行い、誤りの検出、訂正が行われる。これにより、「の-よう-を」という n-gram は誤り部分で出現しやすかったことを学習し、不自然な n-gram の出現確率を低くすることが可能となる。

誤り傾向学習では表層単語の n-gram に限らず、他の言語情報を用いることも有効である。例えば、未然形の直後に名詞がくるといった連鎖は、誤り部分で出現する可能性が高いということなどが考えられるため、活用形品詞といった連鎖情報を用いることも有効であると考えられる。その他、音節数なども言語情報として取り入れることで、音節数の少ない単語は誤りである可能性が高いといったことも学習できると考えられる。

3. 意味情報を用いた誤り検出

本稿で用いる意味情報とは、周辺の認識結果単語を参照したときに、識別対象単語の出現が不自然でないかといった情報のことである。例えば、「裁判」「弁護士」「無罪」といった単語が周辺に出てくる中で、「大根」のように関係のなさそうな単語が出てくると、その単語の意味スコアを小さくし、誤りと判定する確率を高くするという考えをする。本章では、意味スコアの算出法と、意味スコアと単語重みの関係について述べる。

3.1 意味スコアの算出

本節では、周辺の音声認識結果を用いた意味スコアの算出法について述べる。意味スコアは、周辺の認識結果との類似度を用いて算出するが、「は」や「です」といった機能語はどのような単語と共起しても不自然でないため、意味スコアが意味をなさない。そこで本稿では、内容語として名詞、動詞、形容詞のみに意味スコアを与える。音声認識結果に出現した内容語 w の意味スコア $SS(w)$ の算出手順は次の通りである。

- (1) w の周辺に現れる内容語を文脈窓幅 N で集め、単語集合 $M(w)$ とする (w 自身も含む)。
- (2) $M(w)$ 内の各単語 w_i について、図 1 のように $M(w)$ 内の他の単語との類似度 $sim(w_i, w_j)$ を求め、その平均を $SC(w_i)$ とする。

$$SC(w_i) = \frac{1}{N} \sum_j sim(w_i, w_j). \quad (1)$$

- (3) $SC(w_i)$ の平均 $SC_{avg}(w)$ を求める。

$$SC_{avg}(w) = \frac{1}{N} \sum_i SC(w_i). \quad (2)$$

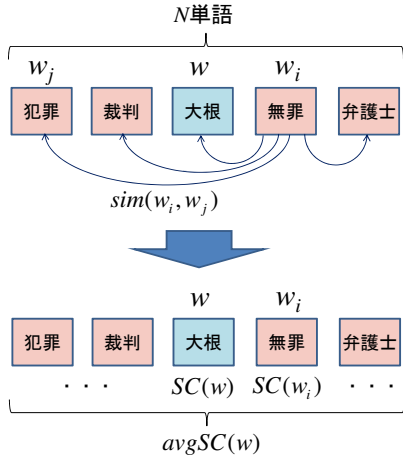


図1 意味スコアの算出

Fig.1 Calculation of semantic score

(4) $SC(w)$ と $SC_{avg}(w)$ の差を意味スコア $SS(w)$ とする.

$$SS(w) = SC(w) - SC_{avg}(w). \quad (3)$$

$SC(w)$ 自体が w と周辺単語との類似度を表しているが, $SC_{avg}(w)$ により正規化を行うことで性能が向上することを確認した. ステップ2で出てくる単語間類似度 $sim(w_i, w_j)$ の算出法については次節で述べる.

3.2 潜在的意味解析

本稿では, 単語間類似度の算出に潜在的意味解析 (LSA: Latent Semantic Analysis) を用いた. LSA とは, 大量のテキストにおける単語の共起関係を統計的に解析することで, 学習データに直接の共起がない単語間の類似度についても求めることができる手法である.

学習の手順としてはまず, N 個の文書から単語文書行列 W を生成する. 本稿では W の要素 w_{ij} として tf-idf を用い, 以下の式により求める.

$$w_{ij} = tf_{ij} \cdot idf_i. \quad (4)$$

$$tf_{ij} = \frac{n_{ij}}{|c_j|}. \quad (5)$$

$$idf_i = \log \frac{N}{df_i}. \quad (6)$$

n_{ij} は文書 c_j における単語 r_i の出現回数, $|c_j|$ は文書 c_j に含まれる単語の総数, df_i は単語 r_i が出現する文書の総数である. idf_i は単語 r_i の単語重みであり, 多くの文書で出現する単語では小さく, 特定の文書でしか出現しない単語では大きくなるという特徴がある. 語彙数を M とすると, 行列 W は $M \times N$ のスパースな行列となる. そこで, この行列 W を特異値分解し, 特異値の大きなものから $R (< rank(W))$ だけ用いることで次のような近似を行う.

$$W \approx \hat{W} = USV^T. \quad (7)$$

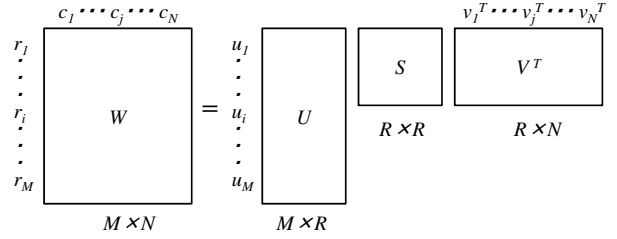


図2 特異値分解による単語文書行列の次元圧縮

Fig.2 Dimension reduction of term-document matrix by Singular Value Decomposition

特異値分解により各行列は図2のような形になっている. $U (M \times R)$ は行ベクトル $u_i (1 \leq i \leq M)$ から成る単語行列, $V (N \times R)$ は行ベクトル $v_j (1 \leq j \leq N)$ から成る文書行列, $S (R \times R)$ は特異値行列である. これは, N 次元の非常に大きな単語ベクトル r_i を, R 次元の $u_i S$ という特徴ベクトルに次元圧縮したと捉えることができる.

一般に単語 r_i と r_j の単語間類似度 $sim(r_i, r_j)$ には, 二つの単語の特徴ベクトルのコサインが用いられ,

$$sim(r_i, r_j) = \frac{(u_i S) \cdot (u_j S)}{\|u_i S\| \|u_j S\|} \quad (8)$$

により求める. $sim(r_i, r_j)$ は1に近いほど類似度が高く, -1に近いほど類似度が低いことを表す.

3.3 単語重みとの組み合わせ

本稿では名詞, 動詞, 形容詞を内容語, それ以外を機能語として扱ったが, 内容語の中にも「こと」や「する」といった頻出単語が含まれている. このような単語は, どの単語と共起しても不自然でないため, 意味スコアがあまり意味をなしていないということが考えられる. そこで, 頻出単語は単語重み (idf) が小さくなることを利用し, 単語重みと意味スコアの有効性について関係を調べた. まず, 単語重みの大きさにより内容語辞書を3つに分割した. そして, 意味スコアによって認識誤

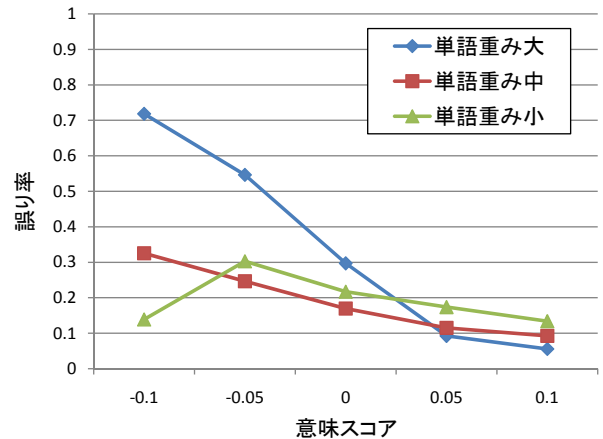


図3 単語重みと意味スコアの有効性の関係

Fig.3 Relation of term-weight and effectivity of semantic score

り単語の含まれる割合 (誤り率) がどのように変化するかを図示したものが図 3 である。この図より、単語重みの大きな単語では、意味スコアが低いほど誤り率が大きく、意味スコアが高くなるほど誤り率が小さくなるという期待していたような相関あることがわかる。しかし、単語重みが小さいほどその相関は小さくなり、意味スコアがあまり有効に働いてないことがわかる。そこで、単純に意味スコアのみで識別に用いるよりも、単語重みが高いときは意味スコアを重視するといった使い方をしたほうが有効であると考えられる。

4. 誤り検出モデル

音声認識誤り検出とは、音声認識器によって出力された文書の各単語が正解か誤りかを推定する問題のことである。そこで認識結果に付与された図 4 のような情報から、正解部分、誤り部分で出現しやすい特徴を学習し、推定に用いることを考える。本稿では、誤り検出を各単語に対して正解か誤りかのラベルを付与していく系列ラベリング問題と考え、誤り検出モデルを Conditional Random Field(CRF) でモデル化する。

CRF では、入力記号列 x に対する出力ラベル列 y の条件付確率分布 $P(y|x)$ を次式のように定義する。

$$P(y|x) = \frac{1}{Z(x)} \exp\left(\sum_a \lambda_a f_a(y, x)\right). \quad (9)$$

ここで f_a は素性、 λ_a は素性関数に対する重みを表し、 $\sum_a \lambda_a f_a(y, x)$ は入力記号列 x に対する出力ラベル列 y のコストと考えることができる。素性とは特徴のことであり、学習データからあらかじめ設定したテンプレートに従い生成しておく。 $Z(x)$ は分配関数で、次式で与えられる。

$$Z(x) = \sum_y \exp\left(\sum_a \lambda_a f_a(y, x)\right). \quad (10)$$

パラメータ λ_a は、学習データ $(x_i, y_i) (1 \leq i \leq N)$ が与えられたとき、条件付確率分布 (9) の対数尤度、

$$\mathcal{L} = \sum_{i=1}^N \log P(y_i | x_i) \quad (11)$$

を最大にするように学習される。これは正解ラベル列のコストと他のすべてのラベル列のコストとの差が最大になるように学習することに相当する。学習は準ニュートン法である L-BFGS 法によって行われる。

識別は学習によって得られた確率分布関数 $P(y|x)$ を用いて、与えられた入力記号列 x に対する最適な出力ラベル列 \hat{y} を求める問題となる。 \hat{y} は次式をもとに Viterbi アルゴリズムにより効率的に求めることができる。

$$\hat{y} = \underset{y}{\operatorname{argmax}} P(y|x). \quad (12)$$

表層単語	所望	の	基本	周期	で	おぼ	ラッパー	...
単語事後確率	0.7	0.2	0.8	0.2	0.0	0.1	0.0	...
品詞	名詞	助詞	名詞	名詞	助詞	名詞	名詞	...
...
意味スコア	0	*	0	0.05	*	-0.1	-0.1	...
単語重み	9	*	3	6	*	6	8	...
正解ラベル	正	正	正	正	正	誤	誤	...

図 4 音声認識誤り検出

Fig. 4 Speech recognition error detection

5. 評価実験

5.1 実験条件

まず、音声認識条件について述べる。音響モデルは、日本語話し言葉コーパス (CSJ: Corpus of Spontaneous Japanese) 学会講演のうち、953 講演 (男性 787 講演 + 女性 166 講演) 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 1 に示す。音響モデルには、HMM(状態共有 triphone) を用いた。1 状態あたりの、混合分布数は 16 とした。サンプリング周波数は 16kHz、音響特徴量は 12 次元 MFCC と対数パワー、12 次元 MFCC の一次微分を加えた 25 次元である。また、言語モデルには CSJ から学習した trigram を用いた。音声認識器としては Julius [13] を用いた。

次に、意味スコアについて述べる。LSA の学習には、CSJ の書き起こし文書のうち、評価データを含まない 2672 講演のものを用いた。内容語のみを扱い、語彙数 48371 であった。文書としては内容語が 30 語程度出てくるとに区切り、文書数 76767 であった。特異値分解で

表 1 音響分析条件と HMM の仕様

Table 1 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	25ms
	フレーム周期	10ms
	窓タイプ	ハミング窓
H	タイプ	状態共有 triphone(3000 状態)
M	混合数	16 混合
M	状態	5 状態 3 ループ

表 2 誤り検出モデルの実験に用いたデータ

Table 2 Experimental data of error detection model

	学習	評価
講演数	150	10
文書数	52692	2667
単語数	484405	22522
内容語数	187154	8782
機能語数	297251	13740
誤り率	23.6%	25.8%

表 3 正解検出性能

Table 3 Performance of correct detection

素性				全単語			内容語			機能語		
認識スコア	言語	意味	単語重み	適合率	再現率	F 値	適合率	再現率	F 値	適合率	再現率	F 値
	×	×	×	0.841	0.910	0.874	0.853	0.927	0.888	0.833	0.898	0.864
×		×	×	0.861	0.925	0.892	0.873	0.925	0.898	0.853	0.925	0.888
×	×			0.754	0.985	0.854	0.807	0.964	0.878	-	-	-
		×	×	0.891	0.915	0.903	0.904	0.915	0.910	0.883	0.915	0.898
			×	0.893	0.920	0.906	0.908	0.925	0.916	0.883	0.816	0.899
				0.894	0.920	0.907	0.911	0.927	0.919	0.883	0.914	0.898

表 4 誤り検出性能

Table 4 Performance of error detection

素性				全単語			内容語			機能語		
認識スコア	言語	意味	単語重み	適合率	再現率	F 値	適合率	再現率	F 値	適合率	再現率	F 値
	×	×	×	0.661	0.507	0.574	0.650	0.459	0.538	0.667	0.531	0.592
×		×	×	0.726	0.571	0.640	0.682	0.546	0.607	0.750	0.585	0.657
×	×			0.642	0.076	0.136	0.642	0.220	0.328	-	-	-
		×	×	0.736	0.679	0.706	0.701	0.671	0.686	0.754	0.683	0.717
			×	0.748	0.684	0.714	0.729	0.681	0.704	0.758	0.685	0.719
				0.749	0.688	0.717	0.738	0.694	0.716	0.754	0.684	0.718

は 100 次元に圧縮した。また、意味スコアを求める際の文脈窓幅は、前の 10 個、後ろの 10 個に対象単語を加えた 21 とした。

次に、誤り検出モデルの実験条件を述べる。学習と評価に用いたデータは表 2 のようになっている。素性には、音声認識スコアとして、識別対象単語の単語事後確率に加え、前後 2 単語の単語事後確率も用いた。言語情報としては、表層単語の 1-gram, 2-gram, 3-gram, 品詞の 1-gram, 2-gram, 活用形-表層単語, 活用形-品詞といった連鎖, 読み, 音素数の 1-gram, 2-gram を用いた。意味情報としては、意味スコアをそのまま用いたときと単語重みと組み合わせたときそれぞれで評価を行った。CRF のツールには CRF++ [14] を用いた。CRF++ では連続値の素性に対応していないため、単語事後確率, 意味スコア, 単語重みについては 10 段階程度に離散化して用いた。

評価実験としては、正解, 誤りそれぞれに対する検出性能の比較を行った。評価値には適合率, 再現率, F 値の三つを用いた。また、意味スコアは内容語のみを対象にしているため、全単語, 内容語のみ, 機能語のみでの評価をそれぞれ行った。

5.2 実験結果

正解検出性能の実験結果を表 3 に、誤り検出性能の実験結果を表 4 に示す。認識スコアと言語情報を用いたものが従来手法、それに意味情報を加えたものが提案手法となっている。実験結果より、意味スコアを追加することで正解検出, 誤り検出ともに性能が向上していることがわかる。また、単語重みと組み合わせて用いることにより、意味スコアの効果が大きくなっている。特に内容

語のみでの評価では改善率が大きく、誤り検出では F 値で $(0.716 - 0.686) = 0.030$ の改善が見られた。機能語には意味スコアが求められないという素性を一意に与えたが、そのことによる機能語の検出性能への悪影響は見られなかった。

全単語での評価では、意味情報の追加による検出性能の改善はわずかであった。この理由としては、内容語の中でも特に意味をもった単語の割合が少なかったことが挙げられる。しかし、このような単語のほうが頻出単語に比べてより多くの情報を持っており、音声認識結果を用いた処理を行う場合に影響力が大きいと考えられる。

次に、意味スコアを加えたことにより検出性能が改善した例をあげる。周辺に「接尾」「活用」「語彙」「助詞」といった単語が出てくる言語に関する内容の講演では、「丹後(単語)」や「イチゴ(一語)」のような認識誤りに対して誤りのラベルを与えることができた。また「音楽」「歌っ」「弾い」といった単語が周辺に出てきたとき、「ギター」に対して正解のラベルの与えることができた。正解のラベルが与えられるようになった単語には、周辺に同じ単語が出現しているものが多かった。

逆に悪くなったり検出できなかった誤りとしては、周辺に「こと」や「いう」のような頻出単語が多く、意味スコアがあまり意味をなしていないものが見られた。また、周辺に類似度の高い認識誤りがあることで、誤りにもかかわらず意味スコアが大きくなってしまっているのではないと思われるものも見られた。そのため、意味スコアを求める際、参照する側の単語重みや認識スコアなども考慮するといったことや、文脈窓幅についても検討する必要がある。

その他の検出できなかった誤りとしては、認識誤りの伝播により自然な n-gram を形成しているようなものが見られた。本稿では、音声認識スコアとして音響スコアと言語スコアを統合した単語事後確率のみを用いたが、このような誤りに対しては、前後の単語の認識スコアが低ければ言語情報の重みを小さくし、音響スコアを重視するような素性が必要なのではないと思われる。

6. おわりに

本稿では、特に内容語の音声認識誤り検出に対して有効であると考えられる、意味スコアについて検討を行った。CRF により誤り傾向学習と意味スコアを組み合わせることで、音声認識誤りを検出する手法を提案した。日本語話し言葉コーパスによる評価実験で、提案手法による検出性能の改善が確認された。また単語重みと組み合わせることで、意味スコアを有効に活用することができた。

今後は、他にも誤り検出に有効な情報がないか検討していく予定である。また、本稿では 1-best の出力結果のみに対して誤り検出を行ったが、複数の音声認識仮説に対して誤り検出を行い、誤り訂正を行う手法についても考えていきたい。

文 献

- [1] H. Jiang, “Confidence Measures for Speech Recognition: A Survey”, *Speech Communication*, vol.45, pp.455-470, 2005.
- [2] Lukas Burget, Petr Schwarz, Pavel Matejka, Mirko Hanneman, Ariya Rastrow, Christopher White, Sanjeev Khudanpur, Hynek Hermansky, and Jan Cernocky, “Combination of strongly and weakly constrained recognizers for reliable detection of OOVs”, *ICASSP*, 4081-4084, 2008.
- [3] 山口 辰彦, 酒向 慎司, 山本 博史, 菊井 玄一郎, “信頼度尺度に基づく音声認識誤りの検出および訂正”, *電子情報通信学会技術研究報告*, SP-65, pp.7-12, 2003.
- [4] 山本 仁, “条件付確率場を用いた信頼度基準による認識誤り検出”, *日本音響学会講演論文集*, pp.63-64, 2006-9.
- [5] Akio Kobayashi, Takahiro Oku, Shinichi Homma, Shoei Sato, Toru Imai, and Tohru Takagi, “Discriminative Rescoring Based on Minimization of Word Errors for Transcribing Broadcast News”, *ISCA*, pp.1574-1577, 2008.
- [6] Diana Inkpen, and Alain Desilets, “Semantic Similarity for Detecting Recognition Errors in Automatic Speech Transcripts”, *HLT/EMNLP*, pp.49-56, 2005.
- [7] Jerome R. Bellegarda, “Latent Semantic Mapping”, *IEEE Signal Processing*, 5(22), pp.70-80, 2005.
- [8] 松浦 優彦, 上原子 正利, 小柳 滋, “ベクトル空間モデルにおける単語重み決定の一般化”, *DEWS*, 2007.
- [9] John Lafferty, Andrew McCallum, and Fernando Pereira, “Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data”, *ICML*, pp.282-289, 2001.
- [10] 太田 健吾, 土屋 雅稔, 中川 聖一, “フィルターの書き起こしのないコーパスからのフィルター付き言語モデルの構築”, *情報処理学会研究報告*, SLP-67-1, pp.1-6, 2007.
- [11] 竹内 孔一, 岡田 和也, 川添 愛, コリアー・ナイジェル, “統計的手法を利用した伝染病検索システムの構築に向けて”, *言語処理学会第 13 回年次大会*, pp.578-581, 2007.
- [12] 浅見 太一, 野田 善昭, 高橋 敏, “ピットフォールエラーに着目した音声認識誤りの分析”, *日本音響学会講演論文集*, pp.53-54, 2008-3.
- [13] 李 晃伸, “julius” <http://julius.sourceforge.jp/>
- [14] 工藤 拓, “CRF++” <http://crfpp.sourceforge.net/>