

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4191021号  
(P4191021)

(45) 発行日 平成20年12月3日(2008.12.3)

(24) 登録日 平成20年9月26日(2008.9.26)

(51) Int.Cl. F I  
**G 1 0 L 15/22 (2006.01)** G 1 0 L 15/22 3 0 0 Z  
**G 1 0 L 15/06 (2006.01)** G 1 0 L 15/06 3 1 0 Z

請求項の数 7 (全 21 頁)

<p>(21) 出願番号 特願2003-401706 (P2003-401706)</p> <p>(22) 出願日 平成15年12月1日(2003.12.1)</p> <p>(65) 公開番号 特開2005-164836 (P2005-164836A)</p> <p>(43) 公開日 平成17年6月23日(2005.6.23)</p> <p>審査請求日 平成17年11月25日(2005.11.25)</p> <p>(出願人による申告)平成15年度通信・放送機構、研究テーマ「大規模コーパスベース音声対話翻訳技術の研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願</p>	<p>(73) 特許権者 393031586 株式会社国際電気通信基礎技術研究所 京都府相楽郡精華町光台二丁目2番地2</p> <p>(74) 代理人 100099933 弁理士 清水 敏</p> <p>(72) 発明者 イアン・レーン 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内</p> <p>(72) 発明者 河原 達也 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内</p> <p>(72) 発明者 松井 知子 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内</p> <p style="text-align: right;">最終頁に続く</p>
--	---

(54) 【発明の名称】ドメイン検証器のトレーニング装置、入力データのドメイン検証装置、及びコンピュータプログラム

(57) 【特許請求の範囲】

【請求項1】

複数のトピックに分類された、一つのドメインに属する機械可読なドメイン内データを記憶するための手段と、

前記ドメイン内データの前記複数のトピックの一つがそれぞれ割り当てられ、トレーニングデータが自己に割り当てられたトピックに属するか否かを判定するようにトレーニング可能な複数のトピック判定手段と、

前記ドメイン内データをトレーニングデータとして用い、前記複数のトピック判定手段をトレーニングするための第1のトレーニング手段と、

与えられる入力データが所定のドメインに属するか否かを検証するためのドメイン検証器と、

前記第1のトレーニング手段によりトレーニングされた前記複数のトピック判定手段と前記ドメイン内データとを用いて、与えられる入力データが前記ドメイン内データのドメインに属するか否かを前記ドメイン検証器が判定可能となるように、前記ドメイン検証器をトレーニングするための第2のトレーニング手段とを含み、

前記第2のトレーニング手段は、

前記複数のトピックのうちのj番目(jは自然数)の一つを選択してドメイン外に設定するためのトピック選択手段と、

前記トピック選択手段によりドメイン外に設定されたトピックに対応する前記トピック判定手段を除く、残りのトピック判定手段に対し、前記ドメインに属するトレーニングデ

10

20

ータを与え、トレーニングデータごとに前記複数のトピック判定手段による判定結果を、各トピックに属すると判定される信頼度尺度として得るための手段と、

前記判定結果を得る手段により得られた前記残りのトピック判定手段による判定結果と、前記トレーニングデータが属するトピックを示す情報とに基づいて、次式で表されるドメイン検証関数  $G_{in-domain}(X)$

【数 1】

$$G_{in-domain}(X) = \begin{cases} 1 & \text{if } \sum_{j=1}^M \lambda_j C(t_j | X) \geq \varphi \quad (\text{ドメイン内}) \\ 0 & \text{otherwise} \quad (\text{ドメイン外}) \end{cases}$$

$X$ : 入力データ

$C(t_j | X)$ : 入力データ  $X$  に対するトピック  $t_j$  の信頼度尺度

$M$ : トピックの数

によって、トレーニングデータが前記ドメイン外に設定されたトピックに属するか否かを判定できるように、前記ドメイン内検証関数  $G_{in-domain}(X)$  の係数  $\lambda_j$  ( $1 \leq j \leq M$ ,  $j \neq i$ ) の値を確率降下法により算出するための係数算出手段と、

前記複数のトピックの各々が全て一度ずつ選択されるように前記トピック選択手段、前記判定結果を得るための手段、及び前記係数算出手段が繰り返し動作するよう制御するための制御手段と、

前記制御手段による繰り返しの結果、各トピックに対して算出された前記ドメイン内検証関数  $G_{in-domain}(X)$  の係数  $\lambda_j$  ( $1 \leq j \leq M$ ,  $j \neq i$ ) の平均値  $\lambda_i$  ( $1 \leq i \leq M$ ) を算出するための平均算出手段と、

前記平均算出手段により算出された係数  $\lambda_i$  ( $1 \leq i \leq M$ ) を前記ドメイン検証器のパラメータに設定するための手段とを含む、ドメイン検証器のトレーニング装置。

【請求項 2】

前記複数のトピック判定手段の各々は、トレーニングデータ及び当該トレーニングデータが自己に割り当てられたトピックに属するか否かを示す情報が与えられると、入力データが自己に割り当てられたトピックに属するか否かを判定するようにトレーニング可能なサポートベクターマシン (SVM) を含む、請求項 1 に記載のドメイン検証器のトレーニング装置。

【請求項 3】

前記 SVM は、入力データが自己に割り当てられたトピックに属するか否かを示す情報を、所定範囲に正規化された数値として出力する、請求項 2 に記載のドメイン検証器のトレーニング装置。

【請求項 4】

前記 SVM は、入力データが自己に割り当てられたトピックに属するか否かを示す情報を、前記所定範囲内の離散値として出力する、請求項 3 に記載のドメイン検証器のトレーニング装置。

【請求項 5】

コンピュータにより実行されると、当該コンピュータを請求項 1 ~ 請求項 4 のいずれかに記載のドメイン検証器のトレーニング装置として動作させる、コンピュータプログラム。

【請求項 6】

ドメイン内データの複数のトピックの一つがそれぞれ割り当てられ、入力データが自己に割り当てられたトピックに属する信頼度尺度を出力するようにトレーニングされた複数のトピック判定手段と、

与えられる入力データが前記ドメインに属するか否かを、前記複数のトピック判定手段の出力する前記信頼度尺度の加重和をしきい値と比較することにより検証するためのドメイン検証器と、

入力データを前記複数のトピック判定手段に与え、前記複数のトピック判定手段が出力した信頼度尺度を前記ドメイン検証器に入力するためのデータ入力手段とを含み、それに

10

20

30

40

50

よって、前記データ入力手段によるデータの入力に応答した前記ドメイン検出装置の出力は、前記入力データが前記ドメインに属するか否かを示す、ドメイン検証装置。

【請求項 7】

複数のトピックに分類された、一つのドメインに属する機械可読なドメイン内データを記憶するための手段と、

入力データが自己に割り当てられたトピックに属するか否かを判定するようにトレーニングされた複数のトピック判定手段と、

与えられる入力データが所定のドメインに属するか否かを検証するためのドメイン検証器と、

前記複数のトピック判定手段、及び前記ドメインに属する入力データを用いて、与えられる入力データが前記複数のトピック判定手段のトピックからなるドメインに属するか否かを前記ドメイン検証器が判定可能となるように、前記ドメイン検証器をトレーニングするためのトレーニング手段とを含む、ドメイン検証器のトレーニング装置であって、

前記トレーニング手段は、

前記複数のトピックのうちの j 番目 ( j は自然数 ) の一つを選択してドメイン外に設定するためのトピック選択手段と、

前記トピック選択手段によりドメイン外に設定されたトピックに対応する前記トピック判定手段を除く、残りのトピック判定手段に対し、前記ドメインに属するトレーニングデータを与え、トレーニングデータごとに前記複数のトピック判定手段による判定結果を、各トピックに属すると判定される信頼度尺度として得るための手段と、

前記判定結果を得る手段により得られた前記残りのトピック判定手段による判定結果と、前記トレーニングデータが属するトピックを示す情報とに基づいて、次式で表されるドメイン検証関数  $G_{in-domain}(X)$

【数 2】

$$G_{in-domain}(X) = \begin{cases} 1 & \text{if } \sum_{j=1}^M \lambda_j C(t_j | X) \geq \varphi \quad (\text{ドメイン内}) \\ 0 & \text{otherwise} \quad (\text{ドメイン外}) \end{cases}$$

$X$ : 入力データ

$C(t_j | X)$ : 入力データ  $X$  に対するトピック  $t_j$  の信頼度尺度

$M$ : トピックの数

によって、トレーニングデータが前記ドメイン外に設定されたトピックに属するか否かを判定できるように、前記ドメイン内検証関数  $G_{in-domain}(X)$  の係数  $\lambda_j$  (  $1 \leq j \leq M$  ,  $j \neq i$  ) の値を確率降下法により算出するための係数算出手段と、

前記複数のトピックの各々が全て一度ずつ選択されるように前記トピック選択手段、前記判定結果を得るための手段、及び前記係数算出手段が繰り返し動作するよう制御するための制御手段と、

前記制御手段による繰返しの結果、各トピックに対して算出された前記ドメイン内検証関数  $G_{in-domain}(X)$  の係数  $\lambda_j$  (  $1 \leq j \leq M$  ,  $j \neq i$  ) の平均値  $\lambda_i$  (  $1 \leq i \leq M$  ) を算出するための平均算出手段と、

前記平均算出手段により算出された係数  $\lambda_i$  (  $1 \leq i \leq M$  ) を前記ドメイン検証器のパラメータに設定するための手段とを含む、ドメイン検証器のトレーニング装置。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、自然言語処理に関し、特に、音声認識又は機械翻訳など、処理対象となる言語活動のドメインに依存したモデルを用いて行なう自然言語処理の改良に関する。

【背景技術】

【0002】

近年、電話と自動音声認識 ( Automatic Speech Recognit

10

20

30

40

50

ion : A S R ) とを用いた自動対話システムが広い範囲で使用される様になっている。そのための技術で現在重要と思われるのは、ユーザが複数のドメインの情報にアクセスできる様にする事で自動対話システムの適用範囲を広げ、使い勝手を向上させるための技術である(後掲の非特許文献1を参照されたい。)

【0003】

ここで「ドメイン」とは、対話が想定される種々の状況を分類する概念である。これは、いわゆる「トピック」という概念と同様であると考えると分かりやすい。たとえば旅行用の英会話のテキストブックでは、会話が行なわれる場面として「空港」、「ショッピング」、「両替」、「駅で」等というトピックに分けて、それらトピックでよく使われる例文をまとめて掲げている事が多い。こうした「トピック」と類似したものとしてドメイン

10

【0004】

簡単のために、以下の説明では、ドメインがトピックの上位の階層であって、あるドメインが複数のトピックを含むものとする。

【0005】

複数ドメインにわたって音声認識を行なう場合、ドメインごとの言語モデル等を用いる事により音声認識システムの精度と効率とがともに向上する事が知られている。こうしたアプローチは、開発のモジュール性を考える上で都合がよい。なぜなら、新しいドメイン

20

【0006】

こうしたトピック依存モデルを用いる従来の対話システムでは、主としてシステムが主導権を握る方式が採用されている。これについてはたとえば後掲の非特許文献2を参照されたい。これらシステムは、ユーザに各種の指示を出し、システムの内部状態に基づいて適切と判断された言語モデルを適用する。こうしたシステムではユーザが主導権を握る事はできず、従ってその使い勝手はよくない。ユーザがドメインを変更できる様にする事で使い勝手は向上すると考えられる。しかしその場合でもユーザは、特定のドメインに関する問合せを行なう場合、その前に明示的に必要なドメインを指定しなければならない。

30

【0007】

電話の自動受付システムでは(非特許文献3を参照されたい。)、ユーザが最初に発した音声に対する音声認識の結果からトピックを検出する処理を行なう事で、明示的ではなくドメインが指定される。対話システムでも、同様の技術を用いて必要とされるドメインを自動的に判定する事ができる。連続する発話は同一のドメインのものである場合が多いので、ドメイン依存モデルを用いた効果が期待できる。

【0008】

ドメイン依存モデルを用いて精度を向上させるシステムとして、出願人が開発したものがある(非特許文献4を参照されたい。)。このシステムでは、複数種類のドメイン依存モデルと、それらを用いて音声認識を行なう複数の音声認識装置とを予め用意し、入力される音声に対しこれら複数の音声認識装置を並列に動作させる。そして、これらから得られた複数の結果のうち、音声認識スコアが最も高いものを選択して出力する。音声認識スコアとは、音声認識装置の内部で複数の認識結果の候補から音声認識結果を選択する際に計算されるスコアの事であり、音声認識結果に付随して得る事ができるものである。

40

【0009】

【非特許文献1】S. スネフ、R. ラウ、J. ポリフローニ著、「Galaxy - II 会話システムの編成、通信、および制御」、ユーロスピーチ予稿集、1999年(S. Senef, R. Lau, J. Polizfoni, "Organization, Communication, and Control in the Galaxy - II Conversational System", Proc. Eurospeech, 1

50

999.)

【非特許文献2】F. ウェセル、A. バーダー著、「1点排除推定を用いたロバストな対話状態依存の言語モデリング」、ICASSP予稿集第2巻、741-744頁、1999年(F. Wessel, A. Baader, "Robust Dialogue-State Dependent Language Modeling using Leaving-One-Out", Proc. ICASSP Vol. 2, 1999.)

【非特許文献3】G. リッカルディ、A. ゴーリン、A. リヨリエ、M. ライリー著、「自動電話受付のための話し言葉システム」、ICASSP予稿集、第2巻、1143-1146頁、1997年(G. Riccardi, A. Gorin, A. Ljoljie, M. Riley, "A Spoken Language System for Automated Call Routing", Proc. ICASSP, Vol. 2, pp. 1143-1146, 1997)

【非特許文献4】T. タケザワ、E. スミタ、F. スガヤ、H. ヤマモト、S. ヤマモト著、「実世界における旅行会話の音声翻訳のための大規模バイリンガルコーパスに向けて」、LREC2002、147-152頁、2002年(Towards a Broad-coverage Bilingual Corpus for Speech Translation of Travel Conversation in the Real World", LREC 2002, pp. 147-152, 2002.)

【発明の開示】

【発明が解決しようとする課題】

【0010】

ある特定のドメインを自動的に選択する場合、選択結果が正しければよいが、選択結果が誤っていると音声認識が正しく行われないう問題がある。また、ドメインとして何を選択すべきかを知っているのがユーザであるとすれば、ユーザが明示的にドメインを指定することにより、ドメインに応じた適切なモデルを用いた精度の高い音声認識を実現できる。

【0011】

しかし、この場合に問題となるのは、ユーザが明示的にドメインを指定した場合でも、時に意識せずドメイン外の発話を行なうことがあるということである。ドメイン外の発話があると音声認識装置は正しく動作せず、翻訳できずに何度も発話を繰返すことを求めたり、全く誤った認識結果を出力したりする。ユーザは自分が誤ったドメインの発話を行なったことは意識していないことが多いので、音声認識装置がこのような動作をすることによりユーザの不満が大きくなる。また、ユーザが音声認識装置のドメインという概念について正しく理解をしていない場合にも同様の問題が生じ得る。

【0012】

こうした問題は、音声認識だけではなく、自然言語処理全般において、その処理対象となる文の領域ごとに異なるモデル又は辞書などを用いることによって処理の精度に大きな相違が生じるシステム全てに共通している。そこで、このような問題を解消することが必要である。

【0013】

これまで、こうした問題についてはそれほど研究されてきていない。通常は例えば処理結果に対する信頼度尺度のような尺度を定義し、この値が所定の条件を充足していないものを棄却するというアプローチを用いている。入力ドメイン内のものか、ドメイン外のものかを確認することが重要であるという視点からの研究はほとんどされていない。

【0014】

ところで、自然言語処理に関する誤りを防止する手立てとして、正誤双方を含む事例を多数集め、それらを用いて統計的な処理を行なうことにより正誤の判定を行なう判定器を用いることがよく行われている。そのためのツールとして、例えばSVM(Support Vector Machine)のようなものがよく知られている。しかし、そうした処理を行なうためには、誤りのデータを収集する必要がある。実際のところ、例えば音

10

20

30

40

50

声認識装置等におけるドメイン外の発話の取り扱い等についてはそうした誤りのデータを多数収集することは困難であり、そのために入力ドメイン外が否かを判定するための判定器を実現することは極めて難しかった。

【0015】

それゆえに本発明の目的は、ドメイン依存の自然言語処理において、指定されたドメインではない、ドメイン外の入力があったときにも適切に判定するドメイン検証装置及びそのトレーニング装置を提供することである。

【0016】

本発明の他の目的は、ドメイン依存の自然言語処理において、指定されたドメインではない、ドメイン外の入力があったときにそれを信頼性高く検出することができるドメイン検証装置及びそのトレーニング装置であって、トレーニングにあたって誤入力に関する事例を収集する必要のないものを提供することである。

【課題を解決するための手段】

【0017】

本発明に係るドメイン検証装置のトレーニング装置は、複数のトピックに分類された、一つのドメインに属する機械可読なドメイン内データを準備するための手段と、ドメイン内データの複数のトピックの一つがそれぞれ割り当てられ、トレーニングデータが自己に割り当てられたトピックに属するか否かを判定するようにトレーニング可能な複数のトピック判定手段と、ドメイン内データをトレーニングデータとして用い、複数のトピック判定手段をトレーニングするための第1のトレーニング手段と、与えられる入力データが所定のドメインに属するか否かを検証するためのドメイン検証器と、第1のトレーニング手段によりトレーニングされた複数のトピック判定手段とドメイン内データとを用いて、与えられる入力データがドメイン内データのドメインに属するか否かをドメイン検証器が判定可能となるようにドメイン検証器をトレーニングするための第2のトレーニング手段とを含む。

【0018】

好ましくは、複数のトピック判定手段の各々は、トレーニングデータ及び当該トレーニングデータが自己に割り当てられたトピックに属するか否かを示す情報が与えられると、入力データが自己に割り当てられたトピックに属するか否かを判定するようにトレーニング可能なSVMを含む。

【0019】

さらに好ましくは、SVMは、入力データが自己に割り当てられたトピックに属するか否かを示す情報を、所定範囲に正規化された数値として出力する。

【0020】

SVMは、入力データが自己に割り当てられたトピックに属するか否かを示す情報を、所定範囲内の離散値として出力してもよい。

【0021】

好ましくは、第2のトレーニング手段は、第1のトレーニング手段によりトレーニングされた複数のトピック判定手段、及びドメイン内データを用い、かつ削除補間法を用いることにより、ドメインに属さないトレーニングデータを用いずに、ドメイン検証器をトレーニングするための手段を含む。

【0022】

第2のトレーニング手段は、複数のトピックのうちの $j$ 番目( $j$ は自然数)の一つを選択してドメイン外に設定するためのトピック選択手段と、トピック選択手段によりドメイン外に設定されたトピックに対応するトピック判定手段を除く、残りのトピック判定手段に対し、ドメインに属するトレーニングデータを与え、トレーニングデータごとに複数のトピック判定手段による判定結果を、各トピックに属すると判定される信頼度尺度として得るための手段と、判定結果を得る手段により得られた残りのトピック判定手段による判定結果と、トレーニングデータが属するトピックを示す情報とに基づいて、次式で表されるドメイン検証関数  $G_{in-domain}(X)$

10

20

30

40

50

【 0 0 2 3 】

【 数 1 】

$$G_{in-domain}(X) = \begin{cases} 1 & \text{if } \sum_{i=1, i \neq j}^M \lambda_{ij} C(t_i|X) \geq \varphi \quad (\text{ドメイン内}) \\ 0 & \text{otherwise} \quad (\text{ドメイン外}) \end{cases} \quad (1)$$

ただし  $X$ : 入力データ $C(t_i|X)$ : 入力データ  $X$  に対するトピック  $t_i$  の信頼度 $M$ : トピッククラスの数

よって、トレーニングデータがドメイン外に設定されたトピックに属するか否かを判定できるように、ドメイン内検証関数  $G_{in-domain}(X)$  の係数  $\lambda_{ij}$  ( $1 \leq j \leq M, j \neq i$ ) の値を確率降下法により算出するための係数算出手段と、複数のトピックの各々が全て一度ずつ選択されるようにトピック選択手段、判定結果を得るための手段、及び係数算出手段が繰返し動作するよう制御するための制御手段と、制御手段による繰返しの結果、各トピックに対して算出されたドメイン内検証関数  $G_{in-domain}(X)$  の係数  $\lambda_{ij}$  ( $1 \leq j \leq M, j \neq i$ ) の平均値  $\lambda_i$  ( $1 \leq i \leq M$ ) を算出するための平均算出手段と、平均算出手段により算出された係数  $\lambda_i$  ( $1 \leq i \leq M$ ) をドメイン検証器のパラメータに設定するための手段とを含んでもよい。

【 0 0 2 4 】

本発明の第 2 の局面にかかるコンピュータプログラムは、コンピュータにより実行されると、当該コンピュータを上記したいずれかのドメイン検証器のトレーニング装置として動作させる。

【 0 0 2 5 】

本発明の第 3 の局面にかかる入力データのドメイン検証装置は、ドメイン内データの複数のトピックの一つがそれぞれ割り当てられ、入力データが自己に割り当てられたトピックに属するか否かを判定するようにトレーニングされた複数のトピック判定手段と、与えられる入力データがドメインに属するか否かを検証するためのドメイン検証器と、入力データを複数のトピック判定手段に与え、複数のトピック判定手段が出力した判定結果をドメイン検証器に入力するためのデータ入力手段とを含み、それによって、データ入力手段によるデータの入力にตอบสนองしたドメイン検出装置の出力は、入力データが前記ドメインに属するか否かを示す。

【 0 0 2 6 】

本発明の第 4 の局面に係るドメイン検証装置のトレーニング装置は、入力データが自己に割り当てられたトピックに属するか否かを判定するようにトレーニングされた複数のトピック判定手段と、与えられる入力データが所定のドメインに属するか否かを検証するためのドメイン検証器と、複数のトピック判定手段、及びドメインに属する入力データを用いて、与えられる入力データが上記した複数のトピック判定手段のトピックからなるドメインに属するか否かをドメイン検証器が判定可能となるように、ドメイン検証器をトレーニングするためのトレーニング手段とを含む。

【 発明を実施するための最良の形態 】

【 0 0 2 7 】

[ 第 1 の実施の形態 ]

概論

以下に述べる第 1 の実施の形態は、入力された発話がドメイン外 (Out - Of - Domain: 以下「OOD」と呼ぶ。) の発話がドメイン内の発話かを検出することが可能な、音声認識装置とともに用いられるドメイン外検証装置に関するものである。この実施の形態にかかる装置は、対象ドメインに属する発話セットのみを用いて、入力された発話がこのドメインに属するか否かを判定する。

【 0 0 2 8 】

そのための枠組みとして、当該ドメインを複数のトピックに分け、入力された発話がこ

れらトピックの各々に属すると考えられる信頼度尺度をトピックごとに算出する。このようにして複数のトピックに対して求めた信頼度尺度をこの入力発話のドメイン判定のための特徴ベクトルとし、予めドメイン内の発話を用いて同様の処理をすることにより得られた信頼度尺度によってトレーニングされたドメイン検証器を用いてOODか否かを判定する。

【0029】

ドメイン検証器のトレーニングにはドメイン内の発話しか使用しないので、OODの発話を収集する必要はない。

【0030】

構成

以下に、本実施の形態の構成について説明する。この実施の形態は本発明を実施するための一つの形態を述べているにすぎず、各構成要素についてはこれ以外にも種々の変形が可能で、また代替が可能なものもある。それらの全てを示すのは困難であるので、可能な変形のいくつかを、実施の形態の説明の最後に掲げる。

【0031】

図1に、本実施の形態にかかるドメイン外判定装置36を用いた音声応答システムの使用環境を示す。図1を参照して、このドメイン外判定装置36を含む音声応答システム28は、ユーザの発話音声を変換するためのマイクロフォン30と、マイクロフォン30からの音声信号に対して、所定のドメイン依存モデル34を用いて音声認識を行なうための音声認識装置32と、音声認識装置32により出力された音声認識の結果が、ドメイン依存モデル34がカバーしているドメイン内のものか、OODかを判定し、OODの場合にはユーザへのフィードバックとしてドメイン外警告信号を出力するための、本実施の形態に係るドメイン外判定装置36と、ドメイン外判定装置36からのドメイン外警告信号に反応して、ドメイン外警告を視覚的に表示したり、スピーカによって音声出力したりするためのドメイン外警告部40と、ドメイン外判定装置36によって入力発話がドメイン内のものであると判定されたことに反応して、音声認識装置32による音声認識結果に基づいてユーザに対し対話のための適切な出力を行なうための応答装置38を含む。

【0032】

この音声応答システム28で使用されるドメイン外判定装置36は、既に述べたように予めこのドメインに対応してトレーニングされている。図1には、そのためのドメイン外判定トレーニングユニット22も示してある。後述するように、ドメイン外判定装置36は処理対象のドメインに含まれる複数のトピックの各々に対して、入力発話とそのトピック内のものであると考えられる信頼度尺度を出力するクラス分類器26と、その結果に基づいて入力発話が対象ドメイン内のものかOODかを判定するドメイン検証器24とを使用する。ドメイン外判定トレーニングユニット22は、ドメイン依存モデル34のドメインに対応する発話を含む発話セット20を用いて、クラス分類器26とドメイン検証器24とをトレーニングする機能を持つ。

【0033】

図2に、発話セット20の構成と、ドメイン外判定トレーニングユニット22の詳細なブロック図を示す。図2を参照して、発話セット20は、それぞれ別々のトピック(第1のトピックから第Mのトピックとする。)の発話からなるM個の発話セット70-1~70-Mに分割されている。

【0034】

ドメイン外判定トレーニングユニット22は、与えられた発話の各々に対して、上記したM個のトピックに属するか否かをトピックごとに信頼度尺度として出力するためのクラス分類器52と、上記したようにクラス分類器52が信頼度尺度を出力するように、発話セット20を用いてクラス分類器52をトレーニングするためのクラス分類器トレーニング部50とを含む。

【0035】

10

20

30

40

50

本実施の形態の装置では、クラス分類器 5 2 に対して与えられる発話データは、発話を構成する 1 - 、 2 - 、及び 3 - グラムを用いて得られる特徴ベクトルの形である。この特徴ベクトル内の各語にはさらに、品詞情報 ( Part Of Speech : POS ) が付加されている。

【 0 0 3 6 】

ドメイン外判定トレーニングユニット 2 2 はさらに、ある発話に対してトレーニング終了後のクラス分類器 5 2 が出力する M 個の信頼度尺度を要素とする信頼度ベクトルが与えられると、当該発話が発話セット 2 0 のドメインに属するか否かを判定するためのドメイン検証器 5 6 と、トレーニング後のクラス分類器 5 2 及び発話セット 2 0 を用いて、後述する削除補間 ( Deleted Interpolation ) 法に基づいたトレーニング方法により、発話セット 2 0 のドメインに属さない発話をトレーニングデータとして用いずにドメイン検証器 5 6 をトレーニングするためのドメイン検証器トレーニング部 5 4 とを含む。本実施の形態で使用する削除補間法については後述する。

10

【 0 0 3 7 】

クラス分類器 5 2 は、発話セット 2 0 の各トピック 7 0 - 1 ~ 7 0 - M にそれぞれ対応して設けられ、発話の特徴ベクトルが与えられると、当該発話がその担当するトピックに属するか否かに関する信頼度尺度を出力するための、第 1 ~ 第 M の SVM 7 4 - 1 ~ 7 4 - M を含む。本実施の形態では、各 SVM に対し、発話の特徴ベクトルの要素と、その発話がこの SVM に割当てられたトピックに属するか否かを示す情報とからなるベクトルを与えることによりトレーニングが行なわれる。この際、トピック内の発話を正解データ、それ以外のトピックに属する発話を誤りデータとする。

20

【 0 0 3 8 】

トレーニング済みの SVM による分類では、発話の特徴ベクトルの張る空間内において与えられた特徴ベクトルにより表される点と、SVM による判別のためのハイパープレーンとの間の距離を用いて分類の尺度とする。入力がトピック内であればこの距離の値は正、それ以外の場合には負とする。また、SVM の出力である距離は  $[- , ]$  の範囲の値であるが、これにシグモイド変換を行なって  $[ 0 , 1 ]$  の範囲に正規化することにより、信頼度尺度が得られる。

【 0 0 3 9 】

クラス分類器トレーニング部 5 0 は、クラス分類器 5 2 の第 1 ~ 第 M の SVM 7 4 - 1 ~ 7 4 - M をそれぞれトレーニングするための第 1 ~ 第 M の SVM トレーニング部 7 2 - 1 ~ 7 2 - M を含む。

30

【 0 0 4 0 】

ドメイン検証器 5 6 は、処理対象の発話 X を受けたトレーニング後のクラス分類器 5 2 から出力される M 個の信頼度尺度を要素とするベクトルに基づき、次の式  $G_{in-domain}$  を適用することで当該発話 X がドメイン外か否かを判定する。

【 0 0 4 1 】

【数 2】

$$G_{in-domain}(X) = \begin{cases} 1 & \text{if } \sum_{i=1}^M \lambda_i C(t_i|X) \geq \varphi \quad (\text{ドメイン内}) \\ 0 & \text{otherwise} \quad (OOD) \end{cases} \quad (1)$$

40

ただし X : 入力データ

$$\lambda_i = \frac{\sum_{j=1, j \neq i}^M \lambda_{ij}}{M-1} \quad (1 \leq j \leq M, j \neq i)$$

$C(t_i|X)$ : 入力データ X に対するトピック  $t_i$  の信頼度

M : トピッククラスの数

ドメイン検証器トレーニング部 5 4 は、ドメイン検証器 5 6 を削除補間法によりトレーニングする。式 ( 1 ) 中の係数  $\lambda_{ij}$  はこのトレーニングにより得られる。ドメイン検証器トレーニング部 5 4 の詳細を説明する前に、削除補間法の概要について説明する。

50

【 0 0 4 2 】

ドメイン検証器 5 6 のトレーニングには、削除補間法を用いるので、発話セット 2 0 中のデータのみ、すなわち、対象ドメイン内のデータのみを用いる。OOD のデータは用いない。ドメイン検証器トレーニング部 5 4 によるトレーニングではさらに、この削除補間法に、GPD ( gradient probabilistic descent : 確率降下法 ) を組み合わせて用いる。

【 0 0 4 3 】

この方法のアルゴリズムを擬似コードで書けば以下ようになる。

[ 1 , M ] 中のトピック j ごとに

10

{

トピック j を一時的な OOD に設定する

他のトピックをドメイン内に設定する

トピック j を OOD、他のトピックをドメイン内と判別するように、式 ( 2 ) に  
よってドメイン検証モデル  $G_{in-domain}(X)$  をトレーニングする

すなわち、GPD を使い、式 ( 2 ) の識別エラーが最小化するように重み (  $\lambda_{1j}$   
, ...,  $\lambda_{ij}$ , ...,  $\lambda_{mj}$  ) (  $i = j$  ) を計算する

【 数 3 】

$$G_{in-domain}(X) = \begin{cases} 1 & \text{if } \sum_{i=1, i \neq j}^M \lambda_{ij} C(t_i | X) \geq \varphi \quad (\text{ドメイン内}) \\ 0 & \text{otherwise} \quad (OOD) \end{cases} \quad (2)$$

20

ただし  $1 \leq j \leq M$

}

$\lambda_{ij}$  (  $1 \leq i \leq M$  ) の各々を、上記した繰り返しにより得られた値 (  $\lambda_{i1}, \dots, \lambda_{ij}, \dots,$   
 $\lambda_{im}$  ) (  $1 \leq j \leq M, j \neq i$  ) の平均値として求める。

つまり、あるトピックを一時的に OOD とし、このトピックに対応する SVM をモデルから削除して、残りの SVM を用い発話セット 2 0 の全ての発話に対する信頼度尺度を求める。発話が、OOD に設定されたトピック以外の残りのトピックに属するような判定が  
上記したドメイン内検証モデル  $G_{in-domain}(X)$  で行なえるように、重み  $\lambda_{ij}$  (  $1 \leq i$   
 $, j \leq M, j \neq i$  ) を GPD により算出する。最後に、 $\lambda_{ij}$  (  $1 \leq i \leq M$  ) の各々を、こ  
うして得た重み  $\lambda_{ij}$  (  $1 \leq j \leq M, j \neq i$  ) を平均することにより求める。

30

【 0 0 4 4 】

ここで、一時的に OOD に設定されたトピックの発話は、トレーニングにおいては誤り ( OOD ) の例として用いられ、残りのトピックの発話は、いずれも正しい例 ( ドメイン内 ) として用いられる。

【 0 0 4 5 】

このようにして重み  $\lambda_{ij}$  (  $1 \leq i, j \leq M, i \neq j$  ) を算出した後、これらを全て平均する。各重み  $\lambda_{ij}$  については  $M - 1$  個の値が算出されるので、重み  $\lambda_i = \lambda_{ij}$  (  $1 \leq j$   
 $\leq M, \text{ただし } j \neq i$  ) / (  $M - 1$  ) として計算できる。これら重み  $\lambda_i$  が求められること  
により、ドメイン検証器 5 6 のトレーニングが完了する。

40

【 0 0 4 6 】

図 2 を参照して、ドメイン検証器トレーニング部 5 4 は、上記した処理を行なうため、発話セット 2 0 のトピックをひとつずつ OOD に指定して繰り返し処理を行ない、重み  $\lambda_i$  を算出する制御を行なうためのトレーニング制御部 6 0 と、発話セット 2 0 の発話の各々に対するクラス分類器 5 2 からの出力のうち、トレーニング制御部 6 0 により指定されたトピック以外のトピックの信頼度尺度を要素とする信頼度ベクトルを作成するための信頼度ベクトル作成部 6 2 とを含む。

【 0 0 4 7 】

50

ドメイン検証器トレーニング部54はさらに、信頼度ベクトル作成部62により発話セット20の発話ごとに得られた信頼度ベクトルを用いて、上記した様にGPDを用いて重み $w_{ij}$  ( $1 \leq i, j \leq M$ ,  $i \neq j$ )を算出するための係数算出部64と、係数算出部64により算出された重み $w_{ij}$ の合計を上記したように $M - 1$ で除することにより平均するための係数平均部66とを含む。

【0048】

クラス分類器トレーニング部50の処理を実現するアルゴリズムを図3にフローチャート形式で示す。図3を参照して、まずステップ100で初期化を行なう。コンピュータで当該処理を実行する場合には、ここでワークメモリの確保、変数への初期値の設定などを行なう。

10

【0049】

ステップ102で、繰返し制御変数 $i$ に0を代入する。続くステップ104～ステップ108の処理は、 $i = 1 \sim M$ までの繰返し処理である。

【0050】

まずステップ104で変数 $i$ に1を加算し、変数 $i$ の新たな値とする。続くステップ106で、変数 $i$ の値がトピック数 $M$ を超えたか否かを判定する。変数 $i$ の値が $M$ を超えていれば処理を終了する。さもなければステップ108に進む。

【0051】

ステップ108では、 $i$ 番目のトピックの発話セットを正解、他のトピックの発話セットを誤りとして、 $i$ 番目のSVMの訓練を行なう。訓練が終了すればステップ104に戻り、次のトピックのSVMの訓練を行なう。

20

【0052】

以上のアルゴリズムにより、クラス分類器52に含まれる第1～第 $M$ のSVMの訓練を行なうことができる。

【0053】

ドメイン検証器トレーニング部54によるドメイン検証器56の訓練を実現するアルゴリズムを図4に示す。図4を参照して、ドメイン検証器トレーニング部54による処理では、最初にステップ120で初期化が実行される。続いてステップ122で繰返し変数 $j$ に0が代入される。以下のステップ124～130は、削除補間法に従った繰返し処理である。

30

【0054】

まずステップ124で、変数 $j$ に1を加算する。続いてステップ126で変数 $j$ がトピック数 $M$ を超えたか否かを判定する。変数 $j$ が $M$ を超えていればステップ132に進み、さもなければステップ128に進む。

【0055】

ステップ128では、 $j$ 番目のSVM72- $j$ をモデルから取除く。続くステップ130で、このように $j$ 番目のSVM72- $j$ を除いたほかのSVM72- $k$  ( $1 \leq k \leq M$ ,  $k \neq j$ )を用いて、発話セット20の全発話を用いてGPDによって重み $w_{ij}$  ( $1 \leq i \leq M$ ,  $j \neq i$ )を算出する。この後、ステップ124に戻る。

【0056】

ステップ126で変数 $j$ がトピック数 $M$ を超えていると判定された場合、ステップ132において $w_{ij}$  ( $1 \leq j \leq M$ ,  $j \neq i$ )の和を $M - 1$ で除算することにより平均し、 $w_i$  ( $1 \leq i \leq M$ )を算出し処理を終了する。以上によりドメイン検証器56の訓練が終了する。

40

【0057】

このドメイン検証器56によるドメイン外検証の基本的な考え方について説明する。ある発話が、あるトピックに属し、他のトピックには属さない場合、その発話から得られる特徴ベクトルをクラス分類器52の全てのSVMに与えると、理想的には、対応のSVMのみの出力が「正解」、他のSVMの出力は全て「誤り」となるであろう。従って、正解を「1」、誤りを「0」とすれば、それらを要素とするベクトルは、正解となった要素の

50

トピックに対応する軸の、座標「1」の近傍を指すであろう。

【0058】

同様に、各トピックのいずれかに属する発話に対してクラス分類器52から得られた信頼度尺度を要素とする信頼度ベクトルは、各トピックに対応する軸の座標「1」の付近に集まるであろう。

【0059】

これに対し、いずれのトピックにも属さない発話、すなわちこのドメインに属さない発話の場合には、得られる信頼度ベクトルの要素はいずれも0に近い値となるであろう。従って、信頼度尺度の張る空間内において、ある発話の信頼度ベクトルが原点付近にあればその発話はドメイン外、原点付近になく各軸の近くに沿った付近にある信頼度ベクトルの場合には、その発話はドメイン内、と判定することができる。

10

【0060】

実際には、そのように理想的な分類はできないので、上記したように削除補間法を用いて各重み  $w_i$  を算出する。

【0061】

なお、上のように考えた場合、各軸ごとに、ドメイン外かドメイン内かを判定するためのしきい値は異なってしかるべきである。ドメイン検証器トレーニング部54による  $w_i$  の算出は、このように軸ごとに異なるしきい値を、各要素ごとの重みに置き換えて算出することである。

【0062】

図5に、簡略のために  $M = 3$  の場合の信頼度ベクトル空間を模式的に示す。この空間の軸が  $x_1 \sim x_3$  であるとする。  $x_1$  軸に対応するトピックに属する発話の場合、信頼度ベクトルの終点（始点を原点とする。）はほぼ部分空間150のあたりに位置するであろう。同様に、  $x_2$  軸に対応するトピックに属する発話の場合、信頼度ベクトルの終点は部分空間150のあたりに位置するであろう。  $x_3$  軸に対応するトピックに属する発話の場合、その信頼度ベクトルの終点は部分空間154のあたりに位置するであろう。

20

【0063】

いずれのトピックにも属さない発話の場合、その信頼度ベクトルの終点は、原点周辺の領域156のあたりに位置するであろう。

【0064】

従って、図5に示すように領域150、152、及び154と、領域156とを分離する平面160（次元数が4以上になれば超平面）を定めることができれば、発話の信頼度ベクトルの終点がこの平面により分割される領域の原点側にあるか、それ以外かによって、領域外か否かを判定できる。

30

【0065】

平面160と各軸との切片は、  $w_1 \sim w_3$  で示すように互いに異なった値となる。このように、各軸との切片を決めるのが、式(1)により表される加重和における重み  $w_i$  としきい値  $th_i$  である。

【0066】

図1示すクラス分類器26及びドメイン検証器24は、図2に示されるドメイン外判定トレーニングユニット22においてトレーニングの済んだクラス分類器52及びドメイン検証器56にそれぞれ相当する。

40

【0067】

ドメイン外判定装置36

一方、図1に示すドメイン外判定装置36は、このクラス分類器26及びドメイン検証器24を用いて、入力発話がOODか否かを判定するためのものであり、以下の構成を有する。図6を参照して、ドメイン外判定装置36は、マイクロフォン30から入力される発話音声に対し、ドメイン依存モデル34を用いた音声認識を行なうことにより音声認識装置32により出力される音声認識結果の発話テキストデータを入力とする。

【0068】

50

ドメイン外判定装置 36 は、入力の発話テキストデータの特徴ベクトルを作成するための特徴ベクトル作成部 180 を含む。特徴ベクトル作成部 180 による特徴ベクトルの作成は、図 2 においてクラス分類器 52 に与えられるものと同様である。すなわち、1 - , 2 - , 及び 3 - グラムと品詞 (word + POS) とを要素として並べたものである。

【0069】

ドメイン外判定装置 36 はさらに、特徴ベクトル作成部 180 から与えられる特徴ベクトルに対し、クラス分類器 26 により作成されるトピックごとの M 個の信頼度尺度を要素として持つ信頼度ベクトルを作成するための信頼度ベクトル作成部 182 と、信頼度ベクトル作成部 182 の作成した信頼度ベクトルを入力として、ドメイン検証器 24 (図 2 のドメイン検証器 56 と同様) を用いて式 (1) に示す演算を行ない、入力発話が OOD いか否かを判定するためのドメイン検証部 184 と、ドメイン検証部 184 による判定結果に応じて、ドメイン外警告部 40 又は応答装置 38 を適切に動作させる処理を行なうための検証結果処理部 186 とを含む。

【0070】

図 7 に、ドメイン外判定装置 36 を実現するためのアルゴリズムをフローチャート形式で示す。コンピュータプログラムによりドメイン外判定装置 36 を実現する場合には、このフローチャートに従った処理を実行するようにすればよい。

【0071】

まずステップ 200 において、音声認識結果の入力 X を受ける。ステップ 202 で、この入力 X に対する特徴ベクトルを作成する。前述したとおり、この特徴ベクトルは入力発話テキストデータの 1 - , 2 - , 及び 3 - グラムと品詞情報とを要素とする。

【0072】

ステップ 204 で、第 1 ~ 第 M のトピックに対応する SVM を用い、この入力 X に対するトピックごとの信頼度尺度  $C(t_i | X)$  ( $1 \leq i \leq M$ ) を算出する。

【0073】

ステップ 206 で、ドメイン内検証モデル (ドメイン内検証関数)  $G_{in-domain}(X)$  をこの信頼度尺度を要素とする信頼度ベクトルに対して適用する。その結果は、式 (1) に示すように入力 X がドメイン内であれば 1、それ以外の場合には 0 である。

【0074】

ステップ 208 で、ステップ 206 において算出されたドメイン内検証関数  $G_{in-domain}(X)$  の値が 1 か否かを判定する。1 であればステップ 210 で音声認識結果に対する適切な応答処理を実行し、処理を終了する。0 であればステップ 212 で入力された発話が、この音声認識装置 32 の対象とするドメイン外であったことを示すドメイン外警告表示を行ない、処理を終了する。

【0075】

[動作]

以上説明システムの動作は 3 つのフェーズに分かれる。第 1 のフェーズは、ドメイン外判定トレーニングユニット 22 によるクラス分類器 52 (クラス分類器 26) のトレーニングフェーズである。第 2 のフェーズは、ドメイン外判定トレーニングユニット 22 において、クラス分類器 52 のトレーニングの終了後、ドメイン検証器 56 (ドメイン検証器 24) のトレーニングを行なうフェーズである。第 3 のフェーズは、このようにトレーニングの終了したクラス分類器 52 (クラス分類器 26) 及びドメイン検証器 56 (ドメイン検証器 24) を備えたドメイン外判定装置 36 を用い、音声応答装置において入力された音声が発話ドメイン外か否かを判定するフェーズである。以下、これらについて順番に説明する。

【0076】

ドメイン外判定トレーニングユニット 22 の動作

ドメイン外判定トレーニングユニット 22 は以下のように動作する。この処理に先立ち、発話セット 20 が準備されており、かつその中の各発話が第 1 ~ 第 M のトピックからなる発話セット 70 - 1 ~ 70 - M に分類されているものとする。この分類は、手作業で行

10

20

30

40

50

なってもよいし、何らかの手段で機械的に行なってもよい。

【0077】

発話が第1～第Mの発話セット70-1～70-Mに対応するトピックに属するか否かを判定できるように、第1～第MのSVM74-1～74-Mに対するトレーニングが実行される。第1～第MのSVMトレーニング部72-1～72-Mが、それぞれ担当するトピックに属する文を正解、それ以外のトピックに属する文を誤り、として各SVMに与えることによりSVM74-1～74-Mのトレーニングが行なわれる。SVMのトレーニングには、一般に入手可能なSVMツールを用いることができる。

【0078】

トレーニングが済むと、クラス分類器52の各SVM74-1～74-Mは、自己が担当するトピックに属する文が入力されると1、それ以外の場合には0となる信頼度尺度を出力するようになる。

【0079】

ドメイン検証器56のトレーニング

クラス分類器52の各SVMのトレーニングが終了すると、ドメイン検証器56のトレーニングが可能になる。

【0080】

トレーニング制御部60は、まず1番目のトピックを選択し、これを一時的なOODとし、残りの第2～第Mのトピックをドメイン内とする。そして、発話セット20に含まれる発話の各々を、第1～第MのSVM74-1～74-Mに対して与える。第1～第MのSVM74-1～74-Mは、これに対してそれぞれ信頼度尺度を出力する。

【0081】

信頼度ベクトル算出部62は、第1～第MのSVM74-1～74-Mから出力される信頼度尺度からなる信頼度ベクトルを発話ごとに生成し、その発話が第1のトピックに属するか否かを示す情報とともに係数算出部64に与える。

【0082】

係数算出部64は、信頼度ベクトル算出部62から与えられる全ての発話に対する信頼度ベクトルと、それぞれの発話が第1のトピックに属するか否かを示す情報とに基づき、GPDにより第1のトピックに関するM-1個の重み $w_{12} \sim w_{1M}$ を算出し係数平均部66に与える。

【0083】

次にトレーニング制御部60は、2番目のトピックを選択し、これを一時的なOODとし、残りのトピックの発話データをドメイン内データとして第1～第MのSVM74-1～74-Mに与える。第1のトピックの場合と同様に、クラス分類器52、信頼度ベクトル算出部62、及び係数算出部64を用い、第2のトピックに関するM-1個の重み $w_{21}, w_{23}, \dots, w_{2M}$ を算出し係数平均部66に与える。

【0084】

トレーニング制御部60は、以下同様にして第3～第Mのトピックの各々に関して、当該トピックに対応する重みを除くM-1個の重みを算出し係数平均部66に与える。

【0085】

係数平均部66は、ドメイン検証器トレーニング部54による重みの計算が終了すると、 $1 \leq i \leq M$ の全てのiに関して、重み $w_{ij}$  ( $1 \leq j \leq M, i \neq j$ )の合計を計算し、さらにその和をM-1で除算することにより、平均の重み $\bar{w}_i$  ( $1 \leq i \leq M$ )を算出する。この重み $\bar{w}_i$ を全て算出することでドメイン検証器56のトレーニングが終了したことになる。

【0086】

ドメイン外判定装置36の動作

マイクロフォン30により入力された音声に対し、音声認識装置32がドメイン依存モデル34を用いて音声認識を行ない、音声認識結果の発話データをテキスト形式で特徴ベクトル作成部180に与える。ドメイン依存モデル34のドメインは、信頼度ベクトル作

10

20

30

40

50

成部 182 及びドメイン検証部 184 のクラス分類器 26 及びドメイン検証器 24 がトレーニングを行なった際のドメインと一致しているものとする。

【0087】

特徴ベクトル作成部 180 は、入力されるテキストデータの 1 - 、 2 - 、及び 3 - グラムと品詞情報とからこの入力の特徴ベクトルを作成し信頼度ベクトル作成部 182 に与える。

【0088】

信頼度ベクトル作成部 182 は、与えられた特徴ベクトルを各 SVM に与えることにより、トピックごとの信頼度尺度を要素とする信頼度ベクトルを作成しドメイン検証部 184 に与える。

10

【0089】

ドメイン検証部 184 は、与えられた信頼度ベクトルをドメイン検証部 184 に含まれるドメイン検証器 24 に与える。従ってドメイン検証部 184 は、前述したとおり、入力ドメイン依存モデル 34 のドメイン外であれば 0、ドメイン内であれば 1 をそれぞれ出力し検証結果処理部 186 に与える。

【0090】

検証結果処理部 186 は、ドメイン検証部 184 の出力が 1 であれば応答装置 38 を制御して、音声認識の結果に対応した適切な応答を出力させる。一方、ドメイン検証部 184 の出力が 0 であればドメイン外警告部 40 に対し、警告を出力する。ユーザは、ドメイン外警告を受けた場合にはその問に対する答えを得ることを直ちにあきらめるであろう。ユーザが、何度も発音しなおしたりすることがなくなり、ユーザの不満が大きくなることを防止できる。

20

【0091】

また、ドメイン検証器 56 のトレーニングには、ドメイン内の発話データしか必要としない。ドメイン外のデータという、収集が困難なデータを用いなくても、ドメイン外か否かの判定を行なうことができ、システムを実現する上での困難が解消される。

【0092】

[変形例]

上記した実施の形態の装置では、SVM への入力など、発話の特徴ベクトルとして 1 - 、 2 - 、及び 3 - グラムと品詞情報とからなるものを用いた。しかし本発明はそのようなものに限定されるわけではない。例えば、1 - グラムのみ、2 - グラムのみ、4 - グラム以上、及びそれらの任意の組み合わせ、並びにそれらと品詞情報との組み合わせ等も用いることができる。

30

【0093】

さらに、単語としては、単語の原形のみ、単語の表層構造等を用いることもできる。さらに、単語の変種を含ませることも可能である。

【0094】

また、上記実施の形態では、トピッククラスの分類に SVM を用いている。しかし本発明はそのような実施の形態に限定されるわけではない。例えば、トピック依存の単語 N - グラム、又は LSA (Latent Semantic Analysis: 潜在意味分析) 等の手法を用いることもできる。

40

【0095】

さらに、上記した実施の形態の手法ではトピックにより明確な判定が困難な場合が生ずることもあり得る。そうした場合には、そのトピックに特化した検証装置を作成し使用することもできる。例えば、上記した実施の形態でのドメイン内検証モデル(ドメイン内検証関数)  $G_{in-domain}(X)$  ではうまくモデル化できないようなトピックが存在した場合には、そのトピック専用の関数を付加してもよい。その場合には、その関数の重みを、当該トピックをよく検出できるようにトレーニングする必要がある。

【0096】

ドメイン検証では、トピック専用の関数が存在する場合には最初に最も入力文が属する

50

信頼度が高いトピックを求め、そのトピックに対し専用の検証用の関数を適用する。それ以外の場合には、式(1)を適用する。

【0097】

〔実験〕

上記したような種々の条件を組合わせて、ドメイン検証の実験を行なった。実験には、出願人において準備したコーパスを用いた。このコーパスの概略構成についてテーブル1に示す。

【0098】

【表1】

テーブル1

ドメイン: 基本的な旅行会話
領域内: 11トピック(移動、宿泊、…)
OOD: 1トピック(ショッピング)
トレーニングセット: 11トピック、149540文(ドメイン内データのみ)
語彙サイズ: 17000語
テストセット: 領域内: 1852発話
OOD: 138発話

ここで、WERは単語誤り率、SERは文誤り率、OOVは語彙外を示す。

【0099】

この実験では、音声翻訳システムにおいて「ショッピング」をOODとした。11のドメイン内トピックを用いて、音声認識のための言語モデル及びトピッククラス分類器をトレーニングした。

【0100】

ドメイン内テストセット及びドメイン外テストセットに対する認識結果をテーブル2に示す。

【0101】

【表2】

テーブル2

	発話数	WER(%)	SER(%)	OOV(%)
ドメイン内	1852	7.26	22.4	0.71
ドメイン外	138	12.49	45.3	2.56

表2から分るように、ドメイン内データセットと比較するとOODデータセットは誤り率及び語彙外率はいずれも高くなっているが、発話の半数以上が正しく認識されている。これは、ドメイン全体に対応した言語モデルを音声認識において使用しているためである。これから、OODセットがドメイン内タスクに関連しており、その結果それらの間の区別は困難であろうと思われる。

【0102】

システム性能は以下の基準により測定した。

【0103】

FRR (False Rejection Rate) : OODに分類されたドメイン外の発話の率

FAR (False Acceptance Rate) : ドメイン内に分類されたOODの発話の率

EER (Equal Error Rate) : FRRとFARとが等しくなった点における誤り率

また、テーブル3に、トピッククラス分類に関して上記した種々の方式、及び発話の特徴ベクトルに関して前述した種々のフォーマットに関するトピック分類の性能について示す。

10

20

30

40

50

【 0 1 0 4 】

【表 3】

テーブル3: 特徴ベクトルと分類モデルによる比較

分類方式	トークンセット	特徴セット	特徴数	EER(%)
SVM	基本形	1-グラム	8771	29.7
SVM	全語	1-グラム	9899	23.9
SVM	語+POS	1-グラム	10006	23.3
SVM	語+POS	1, 2-グラム	40754	21.7
SVM	語+POS	1, 2, 3-グラム	73065	19.6
LSA	語+POS	1-グラム	10006	24.8
LSA	語+POS	1, 2-グラム	40754	25.2
LSA	語+POS	1, 2, 3-グラム	73065	24.2
N-グラム	語+POS	1-グラム	10006	23.3
N-グラム	語+POS	1, 2-グラム	40754	24.1
N-グラム	語+POS	1, 2, 3-グラム	73065	23

テーブル3から分るように、トピッククラス分類器としてSVMを用い、特徴ベクトルとして語+POSの1, 2, 3-グラムを用いた場合が最も低いEERを示した。次に説明するOOD検証においては、SVMを用い、語+POS及び1, 2, 3-グラムからなる特徴ベクトルを用いた実験を行なった。

【 0 1 0 5 】

実験では、全てのSVMの判定結果が否のときに発話がOODと判定するベースラインシステムと、上記した実施の形態に示したようにドメイン内のデータのみを用いてドメイン検証のトレーニングした装置(提案方式)と、ドメイン内データだけでなく、OODデータも用いてドメイン検証のトレーニングを行なったもの(参照方式)とを用い、結果の比較を行なった。その結果、EERにおいてベースラインシステムでは27.7%、参照システムが21.2%であるのに対し、本発明の実施の形態に係るシステムではEER=19.6%となった。すなわち、ベースラインシステムに対しEERにして6.5ポイントの改善を示す一方、参照システムとほぼ同等のEERを示した。本実施の形態に係るシステムでは、領域内発話データしか用いずにドメイン検証のトレーニングを行なったのであるから、この結果は十分に高いものといえることができる。

【 0 1 0 6 】

なお、上記した実施の形態に係るドメイン判定学習ユニット22(図1参照)は、通常のハードウェア構成及びオペレーティングシステムを備えたコンピュータ上で実行されるプログラムにより実現可能である。当該プログラムには、上記したドメイン判定学習ユニット22の機能を実現するために必要な全ての命令を含んでいる必要はなく、例えばオペレーティングシステムに備えられている命令を呼び出すことにより、所望の機能を実現するものでもよい。すなわち、コンピュータのハードウェア及びソフトウェア資源を利用して上記したドメイン判定学習ユニット22の各機能を実現するものであればよい。

【 0 1 0 7 】

また、図1に示す音声応答システム28及びそこで使用されるドメイン外判定装置36もマイクロフォン30及び音声処理専用のボードを除き、一般的な構成のコンピュータ及びソフトウェアにより実現可能である。

【 0 1 0 8 】

そして、そうしたプログラムによりプログラムされたコンピュータは、本発明に係るドメイン検証器のトレーニング装置又はドメイン検証装置となる。

【 0 1 0 9 】

今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味および範囲内

10

20

30

40

50

でのすべての変更を含む。

【図面の簡単な説明】

【0110】

【図1】本発明の第1の実施の形態に係るドメイン外判定装置を用いた音声応答システムの概略構成を示すブロック図である。

【図2】本発明の第1の実施の形態に係るドメイン判定学習ユニット22のブロック図である。

【図3】本発明の第1の実施の形態に係るドメイン判定学習ユニット22のクラス分類器トレーニング部50によるトレーニング処理のアルゴリズムを示すフローチャートである。

【図4】本発明の第1の実施の形態に係るドメイン判定学習ユニット22のドメイン検証器トレーニング部54によるトレーニング処理のアルゴリズムを示すフローチャートである。

【図5】本発明の第1の実施の形態に係るドメイン判定学習ユニット22のドメイン検証器56によるドメイン判定の原理を示す図である。

【図6】本発明の第1の実施の形態に係るドメイン外判定装置36のブロック図である。

【図7】ドメイン外判定装置36を含む音声応答システム28により実行される処理のアルゴリズムを示すフローチャートである。

【符号の説明】

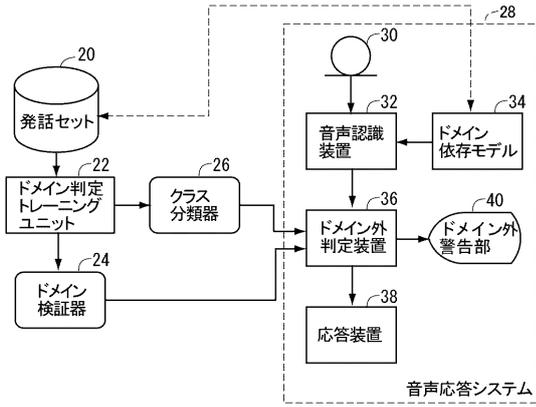
【0111】

20 発話セット、22 ドメイン判定学習ユニット、24 ドメイン検証器、26 クラス分類器、28 音声応答システム、30 マイクロフォン、32 音声認識装置、34 ドメイン依存モデル、36 ドメイン外判定装置、38 応答装置、40 ドメイン外警告部、50 クラス分類器トレーニング部、52 クラス分類器、54 ドメイン検証器トレーニング部、56 ドメイン検証器、60 トレーニング制御部、62 信頼度ベクトル算出部、64 係数算出部、66 係数平均部、70-1~M 発話セット、72-1~M SVMトレーニング部、180 特徴ベクトル作成部、182 信頼度ベクトル作成部、184 ドメイン検証部、186 検証結果処理部

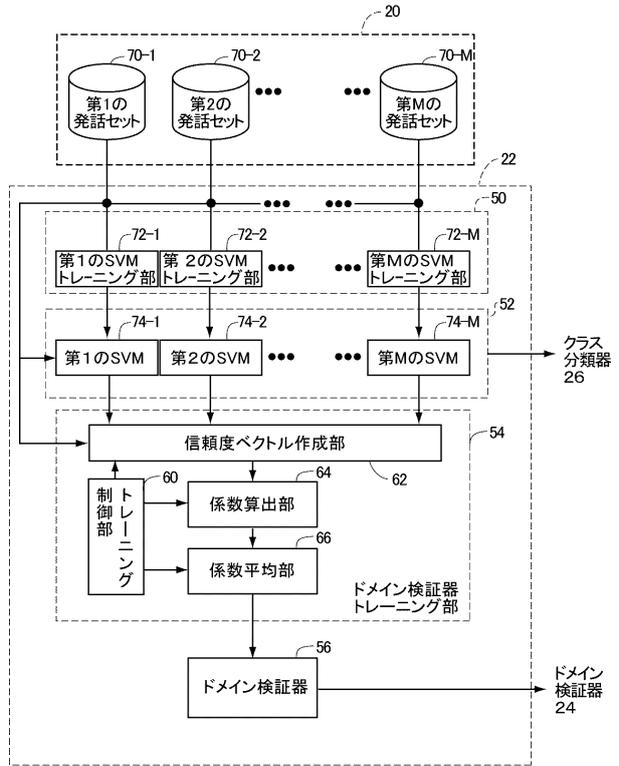
10

20

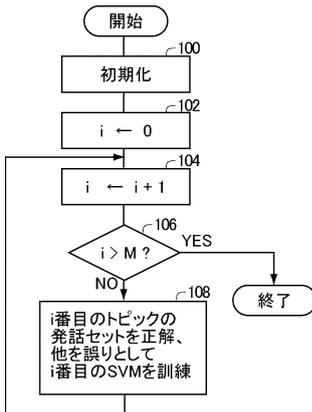
【図1】



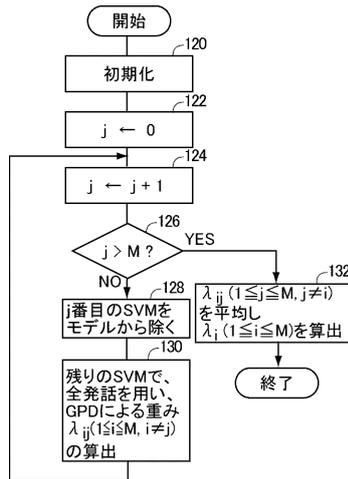
【図2】



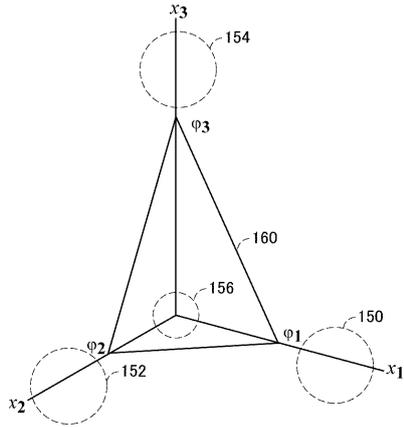
【図3】



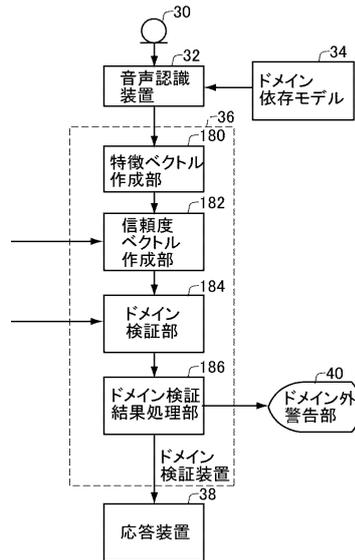
【図4】



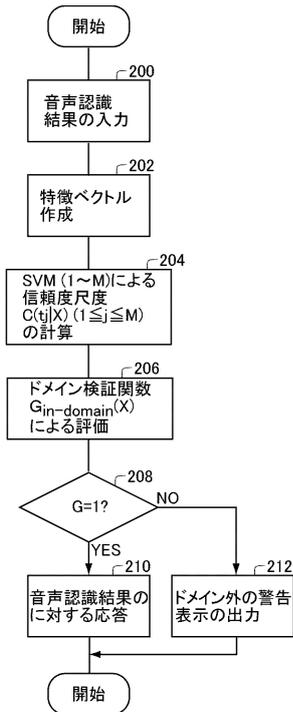
【図5】



【図6】



【図7】



## フロントページの続き

(72)発明者 中村 哲

京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

審査官 間宮 嘉誉

(56)参考文献 特開2004-198597(JP,A)

特開平10-116094(JP,A)

講演音声認識のための言語モデルの検討,加藤一臣,河原達也,日本音響学会2001年春季研究発表会講演論文集 I,日本,社団法人 日本音響学会,2001年 3月14日,p.27-28

Ian R.LANE, Tatsuya KAWAHARA, Tomoko MATSUI, LANGUAGE MODEL SWITCHING BASED ON TOPIC DETECTION FOR DIALOG SPEECH RECOGNITION, Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003年 4月 6日, Vol.1, p. I-616 - I619

平博順, 春野雅彦, Support Vector Machineによるテキスト分類における属性選択, 情報処理学会論文誌, 日本, 社団法人 情報処理学会, 2000年 4月15日, Vol.41, No.4, p.1113-1123

水谷博之, リジェクトを考慮した場合の最小識別誤り学習法の提案, 電子情報通信学会技術研究報告, 日本, 社団法人 電子情報通信学会, 1998年 2月20日, Vol.97, No.559, p.37-42

磯谷亮輔, 嵯峨山茂樹, 自立語と付属語の連鎖統計モデルを用いた音声認識のための候補選択, 電子情報通信学会技術研究報告, 日本, 社団法人 電子情報通信学会, 1993年 6月18日, Vol.93, No.88, p.73-78

Nam Soo KIM, Chong Kwan UN, Statistically Reliable Deleted Interpolation, IEEE Transactions on Speech and Audio Processing, 1997年 5月, Vol.5, Issue.3, p.292-295

(58)調査した分野(Int.Cl., DB名)

G06F 17/28

G10L 15/00-17/00

IEEE

JSTPlus(JDreamII)

JST7580(JDreamII)