

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6852006号  
(P6852006)

(45) 発行日 令和3年3月31日(2021.3.31)

(24) 登録日 令和3年3月12日(2021.3.12)

(51) Int.Cl.	F I				
<b>G 1 0 L 15/10 (2006.01)</b>	G 1 0 L	15/10	5 0 0 T		
<b>G 1 0 L 13/00 (2006.01)</b>	G 1 0 L	13/00	1 0 0 M		
<b>G 1 0 L 15/22 (2006.01)</b>	G 1 0 L	15/22	3 0 0 Z		
	G 1 0 L	15/22	2 0 0 V		

請求項の数 18 外国語出願 (全 21 頁)

(21) 出願番号	特願2018-43052 (P2018-43052)	(73) 特許権者	507134068
(22) 出願日	平成30年3月9日(2018.3.9)		サウンドハウンド、インコーポレイテッド
(65) 公開番号	特開2018-151631 (P2018-151631A)		アメリカ合衆国、95054 カリフォル
(43) 公開日	平成30年9月27日(2018.9.27)		ニア州、サンタ・クララ、ベッツィー・ロ
審査請求日	平成31年3月22日(2019.3.22)		ス・ドライブ、5400
(31) 優先権主張番号	15/456, 354	(74) 代理人	110001195
(32) 優先日	平成29年3月10日(2017.3.10)		特許業務法人深見特許事務所
(33) 優先権主張国・地域又は機関	米国 (US)	(72) 発明者	ライナー・リーブ
			アメリカ合衆国、95054 カリフォル
			ニア州、サンタ・クララ、ベッツィー・ロ
			ス・ドライブ、5400
		審査官	大野 弘

最終頁に続く

(54) 【発明の名称】 ドメイン曖昧性除去を含む音声対応システム

(57) 【特許請求の範囲】

【請求項1】

音声対応システムのユーザに結果を提供する方法であって、  
 少なくとも1つのコンピュータを用いて、  
 複数のドメインに従って自然言語発話を解釈し、ドメイン毎に解釈および関連性スコアを作成するステップと、  
 各関連性スコアをそのドメインの閾値と比較し、閾値を上回る関連性スコアを有する候補ドメインのリストを決定するステップと、  
 前記候補ドメインのリストを前記ユーザに提示するステップと、  
 前記リストからドメインを選択するように前記ユーザに要求するステップと、  
 前記ユーザからのドメインの選択を受付けるステップと、  
 前記選択されたドメインでの前記発話の前記解釈に従って結果を作成するステップと、  
 前記結果を提供するステップとを含み、  
 前記リストにおける前記候補ドメインの最大数は、環境情報に基づくものである、方法。

【請求項2】

音声対応システムのユーザに結果を提供する方法であって、  
 少なくとも1つのコンピュータを用いて、  
 複数のドメインに従って自然言語発話を解釈し、ドメイン毎に解釈および関連性スコ

アを作成するステップと、

各関連性スコアをそのドメインの閾値と比較し、閾値を上回る関連性スコアを有する候補ドメインのリストを決定するステップと、

前記候補ドメインのリストを前記ユーザに提示するステップと、

前記リストからドメインを選択するように前記ユーザに要求するステップと、

前記ユーザからのドメインの選択を受付けるステップと、

前記選択されたドメインでの前記発話の前記解釈に従って結果を作成するステップと

、  
前記結果を提供するステップとを含み、

前記リストにおける前記候補ドメインの最大数は、前記ユーザが利用する前記音声対応システムの種類に基づくものである、方法。 10

【請求項 3】

前記少なくとも 1 つのコンピュータを用いて、前記選択されたドメインを表わすカウンタの値をインクリメントするステップをさらに含む、請求項 1 または請求項 2 に記載の方法。

【請求項 4】

前記候補ドメインは、前記カウンタの値に基づく順序で前記ユーザに提示される、請求項 3 に記載の方法。

【請求項 5】

前記少なくとも 1 つのコンピュータを用いて、最も最近選択された前記ドメインの指標を記憶するステップをさらに含む、請求項 1 から請求項 4 のいずれか 1 項に記載の方法。 20

【請求項 6】

前記候補ドメインは、前記最も最近選択されたドメインの前記指標に基づく順序で前記ユーザに提示される、請求項 5 に記載の方法。

【請求項 7】

前記少なくとも 1 つのコンピュータを用いてレコードをデータベース内に記憶するステップをさらに含む、

前記レコードは、

前記自然言語発話の表現と、

前記発話についての前記ドメインの選択とを含む、請求項 1 から請求項 6 のいずれか 1 項に記載の方法。 30

【請求項 8】

前記少なくとも 1 つのコンピュータを用いてレコードをデータベース内に記憶するステップをさらに含む、

前記レコードは、

前記選択されたドメインに従った前記発話の前記解釈と、

前記ドメインの選択とを含む、請求項 1 から請求項 7 のいずれか 1 項に記載の方法。

【請求項 9】

音声対応システムのユーザに結果を提供する方法であって、

少なくとも 1 つのコンピュータを用いて、 40

複数のドメインに従って自然言語発話を解釈し、ドメイン毎に解釈および関連性スコアを作成するステップと、

各関連性スコアをそのドメインの閾値と比較し、閾値を上回る関連性スコアを有する候補ドメインの数を決定するステップと、

曖昧性除去のために前記ユーザに提示するのに適度なドメインの最大数よりも前記候補ドメインの数が大きいことに応答して、前記ユーザに明確化するように要求するステップとを含み、

前記ユーザに提示するのに適度なドメインの最大数は、環境情報に基づくものである、方法。

【請求項 10】 50

音声対応システムのユーザに結果を提供する方法であって、  
少なくとも1つのコンピュータを用いて、

複数のドメインに従って自然言語発話を解釈し、ドメイン毎に解釈および関連性スコアを作成するステップと、

各関連性スコアをそのドメインの閾値と比較し、閾値を上回る関連性スコアを有する候補ドメインの数を決定するステップと、

曖昧性除去のために前記ユーザに提示するのに適度なドメインの最大数よりも前記候補ドメインの数が大きいことに応答して、前記ユーザに明確化するように要求するステップとを含み、

前記ユーザに提示するのに適度なドメインの最大数は、前記ユーザが利用する前記音声対応システムの種類に基づくものである、方法。

10

【請求項11】

音声対応システムのユーザに結果を提供する方法であって、  
少なくとも1つのコンピュータを用いて、

複数のドメインに従って自然言語発話を解釈し、ドメイン毎に解釈および関連性スコアを作成するステップと、

各関連性スコアをそのドメインの閾値と比較し、閾値を上回る関連性スコアを有する候補ドメインのリストを決定するステップと、

前記リスト上のドメイン毎に、前記ドメインの前記解釈に従った結果を作成するステップと、

20

各結果を提供するステップとを含み、

前記リストにおける前記候補ドメインの最大数は、環境情報に基づくものである、方法。

【請求項12】

音声対応システムのユーザに結果を提供する方法であって、  
少なくとも1つのコンピュータを用いて、

複数のドメインに従って自然言語発話を解釈し、ドメイン毎に解釈および関連性スコアを作成するステップと、

各関連性スコアをそのドメインの閾値と比較し、閾値を上回る関連性スコアを有する候補ドメインのリストを決定するステップと、

30

前記リスト上のドメイン毎に、前記ドメインの前記解釈に従った結果を作成するステップと、

各結果を提供するステップとを含み、

前記リストにおける前記候補ドメインの最大数は、前記ユーザが利用する前記音声対応システムの種類に基づくものである、方法。

【請求項13】

前記環境情報は、ユーザの発話時の場所、発話日時、発話時刻、および発話したユーザの身元のうち少なくとも1つを含む、請求項1、請求項9、または請求項11に記載の方法。

【請求項14】

40

少なくとも1つのコンピュータに、請求項1から請求項13のいずれか1項に記載の方法を実行させるコンピュータプログラム。

【請求項15】

音声対応システムであって、

曖昧性除去を行うために、

複数のドメインに従って自然言語発話を解釈し、

前記自然言語発話が複数のドメインで理にかなっていると判定し、

曖昧性除去のためにユーザに提示するのに適度なドメインの最大数よりも候補ドメインの数が大きいことに応答して、前記ユーザに曖昧性除去を要求する、手段を備え、前記最大数は、環境情報に基づくものであり、前記手段は、曖昧性除去を行うために、さらに、

50

前記ユーザからのドメインの選択を受付けること、および、  
前記発話を処理して所望の結果を提供すること、を行なう、音声対応システム。

【請求項 16】

音声対応システムであって、  
曖昧性除去を行うために、  
複数のドメインに従って自然言語発話を解釈し、  
前記自然言語発話が複数のドメインで理にかなっていると判定し、  
曖昧性除去のためにユーザに提示するのに適度なドメインの最大数よりも候補ドメイン  
の数が大きいことに応答して、前記ユーザに曖昧性除去を要求する、手段を備え、前記最  
大数は、前記ユーザが利用する前記音声対応システムの種類に基づくものであり、前記手  
段は、曖昧性除去を行うために、さらに、  
前記ユーザからのドメインの選択を受付けること、および、  
前記発話を処理して所望の結果を提供すること、を行なう、音声対応システム。

10

【請求項 17】

自動車プラットフォームであって、  
ユーザから音声で発せられた発話を取込むことが可能な音声キャプチャモジュールと、  
複数のドメインに従って前記発話を解釈し、ドメイン毎に関連性スコアを生成する音声  
認識モジュールと、  
音声を生成可能な音声生成モジュールとを備え、  
複数の関連性スコアの各々が対応の閾値を上回ったことに応答して、  
前記音声生成モジュールは、  
ドメインのリストを含む音声を生成し、前記リストにおけるドメインの最大数は、環  
境情報に基づくものであり、  
1つを選択することを前記ユーザに要求する、自動車プラットフォーム。

20

【請求項 18】

請求項 14 に記載のプログラムを記憶するメモリと、  
前記メモリに記憶された前記プログラムを実行するための少なくとも1つのプロセッサ  
とを含む、コンピュータ。

【発明の詳細な説明】

【技術分野】

30

【0001】

発明の分野

本発明は、自然言語発話を処理する音声対応システムの分野の発明である。具体的には  
、本開示は、音声で発せられた質問の曖昧さを除去するように音声対話を取扱う。

【背景技術】

【0002】

背景

人間によってなされた自然言語発話をコンピュータ処理システムが処理することがま  
ます一般的になっている。例えば、ウェブ検索エンジンは自然言語テキスト入力を受付け  
て処理し、視覚的結果を提供する。このようなシステムは、一般的に、1つのブラウザウ  
ィンドウで同時に多数の結果（例えば10個）を提供する。結果は画像およびテキストを  
含み得る。また、このようなシステムは、いくつかの広告を表示することもある。視覚的  
ヒューマンマシンインターフェースは、ユーザが比較的素早く消費可能な多くの情報を含  
む、豊富かつ多様な結果を提供することができる。

40

【発明の概要】

【発明が解決しようとする課題】

【0003】

音声対応システムは、ユーザが機械と対話し、自然言語の発話を音声で発することによ  
って入力を提供するものである。このような機械は一般的に、自動音声認識および自然言  
語処理技術を用いて発話を解釈する。多くの音声対応システムは生成された音声の出力も

50

行うが、そうではないものもある。音声による情報伝達速度は、視覚的ディスプレイよりも遥かに遅い。1つの自然言語表現の入力に回答して、ブラウザベースの検索エンジンが提供するすべての結果をシステムが音声で発するには、長い時間がかかるであろう。それは非常に長い時間を要するため、ユーザが音声のみによってこのような機械と対話することは非現実的であろう。

#### 【0004】

多くの視覚的システムは、意味が曖昧な表現に回答して、発話の複数の合理的な解釈の各々に該当する結果を表示する。大抵の場合、曖昧な発話に対して音声対応システムが適切な結果を提供することは不可能である。従来の音声対応システムは、曖昧な発話に直面したとき、それらの結果を作成するために最良の解釈で推測する。推測された解釈は、ユーザの意図した解釈ではないことがよくある。これは、従来の音声対応システムのユーザにとって、よくある苛立ちの原因である。

10

#### 【0005】

さらに、音声発話の意味の曖昧性を除去するために視覚的ディスプレイを使用することは、例えばディスプレイのないデバイスの場合、またはアイコンタクトを必要とせずには作動すべきデバイスの場合など、多くの状況において実際的ではない。

#### 【課題を解決するための手段】

#### 【0006】

##### 発明の概要

本発明のさまざまな実施形態は、より自然な会話型のユーザインターフェースを提供するシステムおよび方法である。それらは、可能性のある複数の主題分野（会話のドメイン）に関連して発話を解釈する。発話が複数のドメインで理にかなっている場合、システムは、ドメインのリスト、またはドメインを表わす単語を用いてユーザに回答し、次いで、どのドメインを意図したかを示すようにユーザに促す。例えば、「ポール・シモンはどこで生まれましたか？」という発話に回答して、いくつかの実施形態は、「音楽家のポール・シモンか政治家のポール・シモンのどちらですか？」と回答し得る。次の発話が音楽家（すなわち、音楽ドメイン）または政治家（すなわち、政治ドメイン）への関心を示す場合、システムはそれに従って結果を生成する。

20

#### 【0007】

いくつかの実施形態は、自然で人間のような、さまざまな方法で回答する。意味を成さない発話に回答して、いくつかの実施形態は理解できない旨を示す。不明瞭な発話に回答して、いくつかの実施形態は明確化を要求する。例えば「電話して下さい。」のように、意味を成してはいるが、より多くの情報を必要とする発話に回答して、いくつかの実施形態は、例えば「誰に電話をかけますか？」のように欠落情報の要求で回答する。理にかなった意味を1つしか含まない発話に回答して、いくつかの実施形態は、その意図された可能性の高い結果を付与する。例えば「ポール・シモンはどこで生まれましたか？」のように、可能性の高い解釈が僅かしかない発話に回答して、いくつかの実施形態は、例えば「歌手のポール・シモンはニュージャージー州のニューアークで生まれ、政治家のポール・シモンはオレゴン州のユーージンで生まれました。」のように複数の結果で回答する。例えば「サニーベールの天気は？」のように、多数の解釈を有する発話に回答して、いくつかの実施形態は、例えば「カリフォルニア州のサニーベール、テキサス州のサニーベール、ノースカロライナ州のサニーベール、ミズーリ州のサニーベール、またはノバスコシア州のサニーベールのうちのどれですか？」のように、選択肢のリストで回答し、およびユーザに曖昧さを除去するように要求する。

30

40

#### 【0008】

いくつかの実施形態は、例えば各ドメインが選択された回数をカウントすることによって、ユーザのドメイン選択に関する統計を収集する。いくつかの実施形態では多数のユーザに横断的に広くドメイン選択をカウントするのに対して、他の実施形態では、個別のユーザ毎にドメイン選択をカウントする。いくつかの実施形態は、曖昧性除去のために提示されたドメインの組み合わせ毎に、特定のカウンタを有する選択されたドメインをカウント

50

する。

【図面の簡単な説明】

【0009】

【図1A】本発明の実施形態に係る、音声対応システムでユーザ発話に対して結果を提供するための方法の概観を示すフロー図である。

【図1B】本発明の実施形態に係る、図1Aに示す全体のフローのうちの、ドメインに基づいて発話の解釈の曖昧性を除去する方法を示すフロー図である。

【図2A】本発明の実施形態に係る、選択されたドメインのカウンタのインクリメントを示す図である。

【図2B】本発明の実施形態に係る、選択されたドメインのカウンタのインクリメントを示す図である。

10

【図3】本発明の実施形態に係る、発話に関する選択のレコードの記憶を示す図である。

【図4】本発明の実施形態に係る、図1Aに示す全体のフローのうちの、さまざまな曖昧性除去技術を提供する方法を示す図である。

【図5】本発明の実施形態に係る、ユーザ、クライアント、およびサーバの間の関係を示す図である。

【図6】本発明の実施形態に係るサーバを示す図である。

【図7】本発明の実施形態に係るクライアントを示す図である。

【図8】本発明の実施形態に係る自動車の俯瞰図である。

【図9】本発明の実施形態に係る非一時的なコンピュータ読取可能媒体を示す図である。

20

【図10】本発明の実施形態に係るシステムオンチップパッケージを示す図である。

【図11】本発明の実施形態に係るシステムオンチップ内のモジュールを示す図である。

【図12】本発明の実施形態に係るコンピュータシステムのブロック図である。

【発明を実施するための形態】

【0010】

詳細な説明

用語および導入

発話は、音声で発せられた自然言語表現であり、テキストによる自然言語表現と対比される。

【0011】

30

ユーザは、自然言語の発話を音声で発することによって、音声対応システムに入力を与える。

【0012】

システムは、マイクロホンなどを介して発話を受信し、音声認識を行なう。さらにシステムは、自然言語処理を行なって発話を解釈し、解釈を生成する。解釈の処理は、音声で発せられた発話中の単語に文法ルールを適用するものである。結果として生じる解釈は、話し手の意図を表わすデータ構造の例である。

【0013】

会話のドメインは主題分野を表わし、一連の文法ルールを含む。あるドメインで発話を解釈する処理は、ちょうど1つの解釈データ構造例を生成する。システムが会話の異なるドメインの文脈で発話を解釈した場合、異なる解釈が生じる。各解釈は、ドメインによる解釈に従って発話の意味を示す。ユーザは、例えば「今何時ですか?」と尋ね、または「メッセージを送信して下さい。」とシステムに指示して要求を出す。システムは、例えば時間を音声で発することによって応答する。また、システムは、例えば「誰にメッセージを送信しますか?」と尋ねることによってユーザに要求を出し、ユーザは返答として、「お母さんです。」と答えることによって応答する。1つ以上の要求および応答のシーケンスによって、メッセージの送信または時刻の通知などの結果が生成される。

40

【0014】

基本的方法

さまざまな実施形態は、コンピュータ読取可能媒体に記憶されたソフトウェアを実行す

50

るコンピュータを用いる。いくつかの実施形態は、単一のデバイスを含む。いくつかの実施形態は、ユーザに直接音声インターフェースを提供するクライアントにネットワークを介して結合されたサーバを含む。このような実施形態は、クライアントから発話を受信する。

#### 【0015】

図1Aは、本発明の実施形態に係る、音声対応システムでユーザ発話に対して結果を提供するための方法の概観を示すフロー図である。ステップ11では、自然言語理解システムによって発話を受信され、解釈される。システムは、発話の音声信号表現を長い音素シーケンスに変換する。続いて、システムは音素シーケンスをトークン化する。いくつかの実施形態では、システムは、トークン化したものを転写し、音声で発せられた発話のテキストとする。続いて、システムはトークン化された音素シーケンスを複数のドメインの意味論的文法ルールに従って解釈し、ドメイン毎に、そのドメインの文脈での発話を表わす解釈データ構造を作成する。

10

#### 【0016】

音声対応システムは、発話が意味を成す文脈の1つ以上の会話のドメインを決定し得る。各発話は、各ドメインにおいてちょうど1つの解釈を有する。いくつかの実施形態は、発話についての結果(例えば動作)を生成する前に、ユーザの意図した解釈として解釈が選択されるのを待つ。例えば、地理的事実を要求する解釈と、天気状況の確認を要する別の解釈とを有する発話について、当該実施形態は、天気状況の解釈が他方の解釈に対して選択されるまでは天気状況を検索しない。このようなシステムは、選択された解釈についての結果のみを作成することによって、処理の手間、遅延、消費電力、およびデータへのアクセスのために支払う費用を節約する。しかしながら、非常に素早く発話に応答する必要があり、かつ高度な並列システムを実行する用途では、複数の解釈の結果が作成され、次いで、選択された解釈に対応する結果のみが提供される場合もある。

20

#### 【0017】

ステップ11では、ドメインで発話がどれだけ理にかなっているかについて、各解釈が評価される。例えば「雨が降っている!」という発話の場合、システムは、天気ドメインでは高い関連性スコアを割当ててるが、料理ドメインでは低い関連性スコアを割当ててであろう。これは、システムが各ドメインにおける文法ルールに従って発話を解釈することによって行なわれる。文法ルールは、ドメインに関連する単語、およびそれらの意味を規定する。天気に関するドメインは、「raining(雨が降っている)」という単語について、他の単語とさまざまに組合わせて意味を規定する文法ルールを有する。料理に関するドメインは、「raining」という単語については意味を規定しない。システムは、文法ルールに従って発話を解釈する際に、ドメインが発話の理解にどれだけ関連しているかを示すスコアを計算する。例えば、電話に関するドメインでは、「ミシェルに電話をかけて下さい。」は高いスコアを有するであろう。なぜなら、それはコマンドを言い表わす非常に一般的な方法だからである。一方、「ミシェルに接続して下さい。」はやや低いスコアを有するであろう。なぜなら、それは電話のコマンドを言い表わすにはあまり一般的な方法ではないからである。ソーシャルネットワークのドメインでは「ミシェルに接続して下さい。」は高いスコアを有するが、「ミシェルに電話をかけて下さい。」はかなり低いスコアを有するであろう。なぜなら、それはソーシャルメディアに関するコマンドを言い表わすにはかなり珍しい方法だからである。

30

40

#### 【0018】

ステップ12は、各ドメインに割当てられた関連性スコアを、当該ドメインに対応付けられた閾値と比較する。すべてのドメインが同一の閾値を用いるように構成されてもよく、または、閾値がドメイン間で異なるように構成されもよい。ドメイン閾値を超える解釈関連性スコアを有するドメインは、候補ドメインのリストに追加される。すなわち、これらのドメインから作成された解釈のみが、発話の正しい解釈である可能性が最も高いものとして選択されるための候補である。ドメイン閾値よりも高い関連性スコアに対応する解釈を有しないドメインは、フローにおけるこの時点で検討対象から除外してもよい。

50

## 【 0 0 1 9 】

いくつかの実施形態は、選択されたドメインに過去の発話の履歴レコードを関連付ける。これは、関連性スコアリング機能を調整するのに有用である。さまざまな実施形態は、発話からキーワードを抽出し、抽出されたキーワードを、当該発話のドメイン選択とともに記憶する。キーワードに対応付けられたドメイン選択毎に、頻度数を保持してもよい。新たな発話を受信されると、システムは、当該新たな発話のキーワードを検索し、以前の発話においてキーワード毎に選択されたドメインのリストを見つけ出す。次いで、システムは、頻度情報を用いて候補ドメインのリスト上のドメインに重み付けする。いくつかのこのような実施形態は、直接、またはドメイン選択に重み付けした結果として、ユーザに尋ねることなくドメイン選択頻度情報を用いて自動的にドメインを選択する。追加でユーザに尋ねることなく、ドメイン選択頻度情報を用いて候補ドメインのリストからドメインを選択してもよい。

10

## 【 0 0 2 0 】

図 1 B および図 4 に、ステップ 1 3 をより詳細に示す。図 1 B および図 4 は、ドメインに基づいて発話の曖昧性を除去するための異なる実施形態を示す。ステップ 1 3 は、候補ドメインのリストおよびそれらに対応付けられた解釈を用いて、解釈のためのドメインを選択する。実施形態では、ステップ 1 3 は 2 つ以上のドメインを選択してもよく、そのような実施形態では、残りのステップは、選択されたドメインのうちのすべてにおいて行なわれる。

## 【 0 0 2 1 】

ステップ 1 7 は、選択されたドメインでの発話の解釈を用いて結果を作成する。いくつかの実施形態は、複数のドメインの各々での解釈に従って、複数の結果を作成する。発話が情報を要求するものとして解釈された場合、ステップ 1 7 は、要求された情報を検索し、当該検索された情報を含む結果を作成することを含み得る。発話が、行動、例えば環境の何らかの局面を変更すること（サーモスタットを下げる、音楽を 1 曲演奏する、ドライクリーニングした衣類を取込むことを思い出させる）を要求するものである場合、要求された行動を行なうためのコマンドが実行されてもよく、当該行動の完了ステータスを表わす結果が話し手への提示用に作成されてもよい。ステップ 1 8 では、ステップ 1 7 で作成された結果が話し手に提示される。複数のドメインがステップ 1 3 で選択される実施形態では、発話に回答して提供される複数の結果が存在し得る。いくつかの実施形態は、結果を音声として提供する。いくつかの実施形態は、クライアントまたはユーザインターフェース連結デバイスがテキスト読み上げ（TTS : text-to-speech）モジュールに送信し得るテキストとして、結果を提供する。いくつかの実施形態は、音声合成マークアップ言語（SSML : Speech Synthesis Markup Language）などの言語を用いてマークアップテキストを送信し、単語の強調および速度などの結果の属性を表わしてもよい。

20

30

## 【 0 0 2 2 】

いくつかの実施形態は、音声以外の方法で結果をユーザに提供する。例えば、メッセージを送信するいくつかの実施形態では、メッセージを送信せよという要求に回答して、メッセージ受信者に結果を提供する。サーモスタットの設定を変更するいくつかの実施形態では、ユーザへの音声で発せられたメッセージを用いずに、サーモスタットを調節することによって結果をユーザに提供する。

40

## 【 0 0 2 3 】

図 1 B は、本発明の実施形態に係る、図 1 A に示す全体のフローのうちの、ドメインに基づいて発話の解釈の曖昧性を除去する方法を示すフロー図である。図 1 B に示されるフローは、図 1 A のステップ 1 3 を行なうための実施形態である。ステップ 1 4 は、候補ドメインのリストをユーザに提示する。いくつかの実施形態は、例えば「音楽か政治のどちらのことですか？」のように、単にドメインに名前を付けるリストを提示する。いくつかの実施形態は、例えば「シンガーソングライターのポール・シモンか、イリノイ州の前の下院議員のポール・シモンのどちらのことですか？」のように、単なるドメイン名よりも発話に特有の情報を含むリストで応答する。

50



## 【 0 0 2 4 】

ステップ 15 は、ドメインを選択するようにユーザに要求し、次いで、選択による応答を待つ。ステップ 16 では、システムはユーザ選択を受付ける。例えば「大統領のために走った人」のように、選択応答がリスト中のドメインに概念的に関連した情報を含む場合、システムはステップ 16 および 18 に進み、関連ドメインの文脈で作成された解釈を用いて結果を作成する。

## 【 0 0 2 5 】

選択応答が、例えば「1980年代に人気があった人」のように、リストされたドメインのうち2つ以上に概念的に関連した情報を含む場合（シンガーソングライターのポール・シモンおよび政治家のポール・シモンはともに1980年代に人気があった）、システムは、リストを提示するステップおよび選択を要求するステップを繰り返す。図 1 B は、ユーザプロンプトの繰り返しを図示していない。選択応答が、例えば「天気予報は何ですか？」のように、リスト中のドメインに概念的に関連した情報を含めない場合、システムは曖昧性除去の処理を中断し、選択応答を新たな発話として扱う処理に進み、ステップ 11 から再スタートする。

## 【 0 0 2 6 】

## データの収集

データを収集し、ユーザが行なったドメイン選択に関する統計を計算することが有用な場合がある。いくつかの実施形態は、ある単語が発話中に存在するときにユーザが頻りに選択するドメインを優先して、ドメインの関連性スコアを調整する。例えば「トヨタ・タコマ」という発話の曖昧さを除去するためのドメインのリストが提示されたときに、ユーザが自動車販売店または製造業者のウェブサイトよりも消費者レビューを選択する傾向にある場合、当該システムは、「トヨタ・タコマ」という用語が発話内に存在するときに、消費者レビューに関する文法ルールスコアを高くする。結果として、当該システムは、ある発話とドメインとの間の強い相関関係を認定する。その強さが非常に大きい場合、当該システムは、ユーザに曖昧さの除去を要求する必要がない。なぜなら、ユーザの意図が明確であるからである。

## 【 0 0 2 7 】

語句の意味は進化し得るため、いくつかの実施形態は逆戻り機構を含む。いくつかのシナリオでは、特定のドメインがデフォルトとなる。例えば、発話「トヨタ・タコマ」の場合の消費者レビューである。逆戻り機構を有する実施形態は、否定的な反応で応答するユーザを考慮する。否定的な反応を収集するために、さまざまな実施形態は、音声中の否定的なキーワードを認識するか、ユーザに対して賛成ボタン/反対ボタンを提供するか、または他のユーザ挙動を監視する。発話が複数の高スコアのドメインの仮説を生じさせたにもかかわらずシステムが常に1つのドメインしか選択しなかった場合に、ユーザ反応が結果への不満を表わすとき、逆戻り機構は、発話種類が曖昧なものとして扱われるように戻す変更を行なう。

## 【 0 0 2 8 】

図 2 A は実施形態に係る統計の収集を示す。ユーザがドメイン選択を行なう度に、システムは、複数のカウンタのうちセレクトア 20 によって特定された1つのカウンタをインクリメントする。このシステムは、3つのドメイン（ドメイン A 21、ドメイン B 22、およびドメイン C 23）のうち各々にカウンタを有する。

## 【 0 0 2 9 】

図 2 B は実施形態に係る統計の収集を示す。ユーザが3つのドメインのうち2つのドメインのリストに基づいてドメイン選択を行なう度に、セレクトア 24 は、リストされたドメインのセットおよびドメイン選択を用いて、インクリメントすべき適切なカウンタを選択する。リスト上のドメインの組み合わせ毎に、各ドメイン選択に対応するカウンタが存在する。カウンタアレイ 25 はすべてのカウンタを記憶する。例えば、ドメイン A および B がリストされており、かつユーザがドメイン A を選択した場合に、カウンタドメイン A B - A カウンタがインクリメントされる。

10

20

30

40

50

## 【0030】

図3は、発話と、結果として行なわれる関連ドメインのユーザ選択とをレコード31としてデータベース32内に記憶することを示す。さまざまな実施形態は、デジタル音声サンプル、テキスト転写などのトークン化された音素のシーケンス、および解釈データ構造のうちの一つ以上として、発話を表現する。システム設計者およびドメイン固有意義論的文法の作者は、ユーザ曖昧性除去選択を用いて、それらの文法処理を改善することができる。機械学習システムは、データベース中の大量のデータ（発話の代替表現を含む）を処理することによって、ユーザ曖昧性除去選択に基づいてドメインスコアリングを操作することができる。例えば、最低限のユーザ対話で、またはユーザ対話せずに、発話について最も可能性の高いドメインを自動的に特定するためのニューラルネットワークアルゴリズムなどである。

10

## 【0031】

いくつかの実施形態は、環境情報をデータベースレコードに記憶する。環境情報のいくつかの例は、ユーザの発話時の場所、発話日時、発話時刻、および発話したユーザの身元（ユーザID）である。

## 【0032】

いくつかの実施形態は、最も最近使用されたドメインのレコードを記憶する。このような実施形態は、曖昧性除去のためのドメインのリストを提供する際に、最も最近使用されたドメインをリストの最上位に置く。いくつかのこのような実施形態は、曖昧性除去の要求とともにリストを提供する際に、ユーザ応答を受付ける。ユーザが選択によって要求に応答すると、システムはリストの提供を終了する。最も最近使用されたドメインの提供を最優先することによって、システムは、意図されたドメインを選択するためにリストを通じてユーザを待たせないことにより、ユーザの時間を節約する。

20

## 【0033】

いくつかの実施形態は、すべてのドメインが順序付けられたリストを記憶する。ドメインが使用される度に、システムは、そのドメインを順序の最上位に移動させる。このようなシステムは、曖昧性除去のためにドメインのリストをユーザに提供する際に、記憶された順序でリストを提供する。

## 【0034】

異なるシナリオにおける挙動

30

いくつかの実施形態は、曖昧な発話の種類によって異なる方法で応答する。図4は、本発明の実施形態に係る、さまざまな応答を提供する方法を示す。

## 【0035】

図4の方法は、候補ドメインの数のカウントで始まる。ステップ40では、そのカウントを用いて、システムに以下の6つの方法で挙動させる。すなわち、全般的な明確化を要求すること、複数の結果を付与すること、図1Aの方法と同様に、ドメインのリストを提供してユーザに選択を要求すること、特定の追加情報を要求すること、最も可能性の高い意図された解釈に従って結果を付与すること、および、発話不理解の旨を示すこと、である。

## 【0036】

40

候補ドメインがゼロである場合、つまり、対応付けられたドメインの閾値よりも大きな関連性スコアが割当てられたドメインが無い場合、当該方法は、システムが発話を理解しなかったことを示す（41）。例えば「フェットチーネ 多数 攪乳器 ジュース ロマン アトリウム ボルト」という発話は、文法的正しくないか、または如何なるドメインにおいても意味のあるものではない。したがって、それはすべてのドメインで低い関連性スコアを有するであろう。不理解の旨を示すために、システムは、例えばユーザに「申し訳ありません。まだできません。」、または「申し訳ありません。理解できません。」と伝えてもよい。視覚的ディスプレイを用いるいくつかの実施形態は、発話についてウェブ検索エンジンで検索を行なうことによって得られた結果を（例えば、発話のテキスト転写によって）表示してもよい。この時点で、当該発話の処理が終了する。

50

## 【 0 0 3 7 】

ちょうど1つの候補ドメインが存在する場合、次いでステップ42において、システムは、当該システムによる結果の作成のために十分な情報を発話が含んでいるか否かを検討してもよい。例えば「メッセージを送信せよ」のような発話は、システムが有用な結果を作成するための情報を十分には含んでいない。さらなる情報が必要な場合には、ステップ43においてユーザは必要な追加情報を促される。例えば、システムは「誰にメッセージを送信しますか？メッセージで何を伝えますか？」のような特定の追加情報を要求し得る。

## 【 0 0 3 8 】

ちょうど1つのドメインが存在し、かつ、システムによる結果の作成のために発話が必要とする追加情報が無い場合には、当該1つの候補ドメインが、ステップ44で使用するドメインとして選択される。例えば「猫には尻尾がありますか？」のような発話については、システムが回答を決定し、「はい。」と応答する結果を作成するために何ら追加情報を必要としない。例えば「たった今、バージニア州のアレクサンドリアで雨が降っていますか？」のような極めて具体的な発話ならば、疑義は生じない。天気ドメインなどの1つのドメインのみが、閾値を上回るスコアを有するであろう。

## 【 0 0 3 9 】

候補ドメインの数が少なくとも2つであり、かつ、ユーザに提示するのに適度なドメインの最大数以下である場合、システムはステップ45に進む。例えば、ユーザに提示するのに適度なドメインの最大数は、2、3、4、5、またはそれ以上であり得る。いくつかの実施形態では、最大数は可変である。ユーザに提示するのに適度なドメインの最大数の値はシステムの種類に依存してもよく、いくつかのシステムでは、場所、時刻、ユーザID、および他の環境情報に依存する。例えば、車におけるいくつかのシステムの場合、多くの曖昧性除去ステップの遅延を回避する必要があるため、3つ以上のドメインが閾値を上回るスコアを有する場合に全般的な明確化を要求する。一方、リラックスした夕方に使用されるいくつかのホームアシスタントシステムの場合、6つ以上のドメインが閾値を上回って初めて全般的な明確化を要求する。

## 【 0 0 4 0 】

いくつかの実施形態では、すべてのドメインが関連しているとステップ45が判定した場合、システムは、ドメインのリストを提供せずに、かつ、ユーザに曖昧性除去の要求をせずに、ステップ46に進む。システムは複数の結果を提供する。リスト上のドメインが関連しないとステップ45が判定した場合には、システムはステップ47に進み、図1Bに示すようにユーザにリストを提供して曖昧性除去の要求を行なう。ドメインのうちすべてを使用することが決定された場合、候補ドメイン同士の類似度合い、および発話において表現された概念同士の類似度合いが考慮される。例えば「ジョージアの首都は？」という発話の場合、国に関する情報のドメイン、および米国の州に関する情報のドメインがともに閾値を上回るスコアを付与するであろう。なぜなら、「ジョージア」は米国の州の名称であるとともに、国の名称でもあるからである。このようなシナリオにおいて、これら2つの候補ドメインはともに地理に関係するため、ステップ45は、これら2つの候補ドメインを類似概念を有するものとして識別する。システムは、例えば「国のジョージアの首都はトビリシで、米国の州のジョージアの州都はアトランタです。」のように複数の結果で応答してもよい。例えば「コロラド州のデンバーの高さは？」のような発話の場合、天気ドメインおよび地理ドメインの両方が、閾値を上回るスコアを有する。これらのドメインは密接には関連していないが、「コロラド州のデンバー」という概念は両方のドメインに共通するため、システムは、続いて「デンバーは標高1マイルであり、その最高気温は氷点下40度です。」のように複数の結果で応答してもよい。

## 【 0 0 4 1 】

適度な数の非類似のドメインに関連し、ひいては曖昧性を除去する要求が必要となり得る発話の例としては、以下のものがある。

## 【 0 0 4 2 】

「本日の最高のパフォーマーは誰ですか？」（これは、「演劇か音楽のどちらですか？」という曖昧性除去の要求を受けるであろう。）

「ブライアン・ウィルソンは、2006にどこでプレイしましたか？」（これは、「ビーチ・ボーイズのブライアン・ウィルソンか、ジャイアンツおよびドジャーズのピッチャーのブライアン・ウィルソンのどちらですか？」という曖昧性除去の要求を受けるであろう。）

「トウガラシをどのくらいホット（hot）にしますか？」（「温度的な熱さか、ピリ辛さのどちらですか？」という曖昧性除去の要求を受けるであろう。）

候補ドメインの数が、ユーザが提示されたドメインのうちの1つを選択可能なようにユーザに対して適度に提示され得る数を超える場合、システムはステップ48に進む。例えば、システムは、「トヨタ・タコマ」のような発話を受信した後に、「より具体的にお願  
10  
いします。」のように全般的な明確化を要求することによって、ユーザに応答し得る。いくつかのシステムは、例えば「『トヨタ・タコマ』に関して何を知りたいですか？」のような発話からの情報を利用する。他の例としては、「Turkey」という発話への応答があるだろう。「Turkey」という発話は、動物、レシピ、地理、歴史、ニュース、天気、およびスポーツのドメインのうちのすべてにおいて意味を持つため、これらのドメインで高スコアが割当てられ得る。

#### 【0043】

いくつかのシステムは、全般的な明確化を要求した後に、次の発話を新たな発話として  
20  
取扱う。いくつかのシステムは、全般的な明確化を要求した後に、応答発話からの情報を、正しい解釈の構築のために明確化を必要とした発話に追加する。そうすることにより、システムは、「それはどのくらい信頼性が高いですか？」という次の発話を「トヨタ・タコマ」ピックアップトラックに言及するものとして解釈し、自動車の消費者レビューのデータソースから得た結果を付与するであろう。

#### 【0044】

共通の頭字語を使用する発話は、たとえ1つのドメイン内であっても、明確化を必要とする多くの可能な解釈を有する傾向にある。例えば「CSUはどこですか？」という発話  
30  
では、頭字語「CSU」は、特にコロラド州立大学（Colorado State University）またはカリフォルニア州立大学（California State University）を指し得る。カリフォルニア州立大学の場合、23の異なるキャンパス所在地がある。

#### 【0045】

いくつかの実施形態は、異なるシナリオにおける挙動を決定するために、ローカライゼーションおよびユーザプロファイリングを適用する。ローカライゼーションは、全地球的な位置のうちの特定の範囲、および環境の種類（例えば、家、オフィス、車両など）を決定することを含み得る。プロファイリングは、個人情報（例えば、年齢、性別、人種、自宅住所など）、ならびに、最近使用したドメイン、特定のキーワード、および着信メッセージなどの情報を含み得る。いくつかの実施形態は、ローカライゼーション情報値およびプロファイリング情報値を重みとしてドメイン仮説に適用する。例えば、いくつかの実施形態は、男性ユーザの場合には、スポーツドメインにプラスに重み付けするとともに、ファッションドメインにマイナスに重み付けする。いくつかの実施形態は、山岳地域のユーザの場合にはスキードメインにプラスに重み付けし、海岸地域のユーザの場合には水泳ドメインにプラスに重み付けする。いくつかの実施形態は、ローカライゼーション情報値およびプロファイリング情報値を、ユーザ選択のために提示するのに適度なドメイン数を増加/減少する重みとして適用する。例えば、いくつかの実施形態は、ユーザが車の中にいる場合、運転者の気を散らす可能性のある長い選択肢リストを提示するのではなく、曖昧性除去のための候補ドメインの数を（たとえその選択肢の数が電車の乗客にとっては許容範囲であったとしても）2つに減らして、全般的な明確化の要求を生じさせる。いくつかの実施形態は、ローカライゼーション情報値およびプロファイリング情報値をドメイン類似性の計算に適用する。  
40

#### 【0046】

10

20

30

40

50

### クライアントサーバモデル

図5は、クライアント-サーバインターフェースを用いる実施形態を示す。ユーザ51はクライアント52に対して発話し、クライアント52はマイクロホンで発話を受信する。また、ユーザは、クライアントから音声で発せられた応答をスピーカを通して聞く。クライアント52は、インターネットなどのネットワーク53を介してサーバ54に情報伝達する。サーバは、さまざまなドメインの文法ルールに従って発話を処理する。

【0047】

図6はサーバの実施形態を示す。サーバは、ラック61と5つのサーバブレード62とを備える。

【0048】

図7はクライアントの3つの例を示す。図7aは携帯電話71を示す。携帯電話は電池式であるため、電池が切れないようにするためには、複雑な計算を最小限にすることが重要である。したがって、インターネットを介して携帯電話71がサーバに接続されてもよい。いくつかのユースケースでは、携帯電話は情報を提供可能な視覚的ディスプレイを有する。しかしながら、携帯電話はスピーカも有し、いくつかのユースケースでは、携帯電話は音声のみによって発話に応答してもよい。

【0049】

図7bは、ホームアシスタントデバイス72を示す。それは据え付けの電源に差込み得るため、携帯電話よりも高度なローカル処理を行なうためのパワーを有する。電話71と同様に、ホームアシスタントデバイス72は、特殊なドメイン、特に、有用な結果を作成するためにダイナミックデータを必要とするドメインに従って発話を解釈するために、クラウドサーバを利用してもよい。ホームアシスタントデバイス72はディスプレイを有しないため、それは音声のみのデバイスである。

【0050】

図7cは自動車73を示す。自動車は、ワイヤレスネットワークを介してインターネットに接続可能であってもよい。しかしながら、信頼性が高いワイヤレスネットワークにつながるエリアから出た場合、自動車73は、ローカル処理のみを用いて確実に発話を処理し、応答し、適切な結果を付与しなければならない。結果として、自動車は、自然言語発話処理のために局所的にソフトウェアを実行する。多くの自動車は視覚的ディスプレイを有しているが、危険な方法で運転者の気を散らすことを回避するために、自動車73は、音声のみの要求および応答によって結果を提供する。

【0051】

### 自動車システム

自動車におけるヒューマンマシンインターフェースの必要条件是、素人によって動作される場合であっても、機械のシステムが安全に動作する必要があるため、特に要求が厳しい。また、自動車は特に困難な環境も生じさせる。なぜなら自動車は、窓を開けた状態で反響するトンネルを通過して他の騒々しい車両の近くを走行することもあるからである。しかしながら、他のいくつかの音声対応システムは如何なる方向からの音声も受信しなければならない1つのマイクロホンまたは密集した複数のマイクロホンを備える小型デバイスであるのに対して、自動車は発話を行なう人々を取囲んでいる。したがって、自動車は発話を受信するための取囲みマイクロホンを有しており、それらは離間距離によって騒音を打消し、複数の人々のうちのだれが話しているかを検出する。携帯電話は個人用であるため、多くの話し手のうちの1人を特定することは、困難ではあるが通常は問題にならない。

【0052】

いくつかの供給業者は、電子サブシステムを製造して自動車産業界に販売している。車両におけるエレクトロニクスの量は急速に増えている。多くのこのようなサブシステムは、プログラム作製者が有用な機能を開発し、維持し、向上させることを可能にするハードウェアプラットフォームおよびソフトウェアプラットフォームからなる。いくつかの自動車プラットフォームは、自動車が音声対応デバイスとして動作することを可能にするハー

10

20

30

40

50

ドウェアサブモジュールまたはソフトウェアサブモジュールを含む。自動車プラットフォームのいくつかの実施形態は、音声キャプチャモジュールを含む。音声キャプチャモジュールは、マイクロホンから音声を受信し、ノイズリダクションアルゴリズムを実行し、エコーキャンセルを行ない、スピーカ音声分離を行なう。

【0053】

実施形態は、音声認識モジュールをさらに含む。それは、音声キャプチャモジュールからの処理された音声を用いて、1つ以上の仮説音素シーケンスと、各仮説の強さを表わすスコアとを生成する。続いて、音声キャプチャモジュールは、統計的言語モデルに従って仮説音素シーケンスをトークン化したものを計算し、転写または音声単語シーケンスを生成する。いくつかの実施形態ではネットワーク接続を介してアクセス可能なサーバを補助のために利用するのに対して、他の実施形態では、すべての音声認識機能を局所的に行なう。音声認識モジュールは、音声認識モジュールからの1つ以上のスコア化された仮説に基づく単語シーケンスを用いて、複数のドメインの文法ルールに従って発話を解釈する。音声認識モジュールは、ドメインにおける各文法ルールに従って単語シーケンスの関連性を計算し、音声認識モジュールからの仮説見込みスコアによってそれを重み付けする。音声認識モジュールは、最も高いスコアの文法ルールに従った解釈を用いて、ドメインの解釈およびスコアを提供する。続いて、音声認識モジュールは、上述の実施形態に従って解釈を選択するか、または自動車内のユーザに曖昧性除去選択を要求して曖昧性除去選択を自動車内のユーザから受け取る。

【0054】

自動車プラットフォームは、選択された解釈を用いて、発話の結果としてのさまざまな機能呼び出す。いくつかのこのような機能としては、ナビゲーション動作を行なうこと、暖房装置または窓などの他の自動車サブシステムを動作させること、および、天気状況または交通状況に関する質問などの質問に回答することがある。

【0055】

いくつかの実施形態によれば、自動車プラットフォームは、音声生成モジュールをさらに含む。音声認識モジュールは音声生成モジュールを呼び出して、上述の実施形態に従って曖昧性除去を要求する。さらに、音声認識モジュールは音声生成モジュールを呼び出して、天気状況などの情報の結果、および、コマンド発話に従って暖房装置の設定が変更されたことなどの確認の結果を提供する。

【0056】

さまざまな実施形態では、音声キャプチャモジュール、音声認識モジュール、および音声生成モジュールによって行なわれる機能は、このようなモジュールの異なる例において行なわれる。

【0057】

図8は、運転者が自動車の左側に座った状態で道路の右側を運転するように設計された自動車80の俯瞰図を示す。自動車80は2つの前部座席81を有しており、それぞれが1人の人を収容可能である。また、自動車80は数人の人を収容可能な後部座席82を有する。自動車80は、速度およびエネルギーレベルなどの基本情報を表示する運転者情報コンソール83を有する。また、自動車80は、例えばナビゲーションマップ上の閲覧箇所および入力箇所など、音声では素早く行なうことができない複雑な人間の対話のためのタッチボードコンソール84を有する。

【0058】

自動車80は、サイドパーマイクロホン85および天井取付けコンソールマイクロホン86を有する。それらすべては、自動車に組込まれたデジタルシグナルプロセッサが運転者または前部座席の乗客からの音声同士を区別するためのアルゴリズムを実行できるように、音声を受信する。また、自動車80は、後部座席の乗客からの音声を受信する後部天井取付けコンソールマイクロホン87を有する。

【0059】

また、自動車80は、スピーカを備える車載用音声システムを有する。スピーカは音楽

10

20

30

40

50

を再生可能であり、ユーザコマンドに対して音声で発せられた応答および結果の音声を生成することも可能である。また、自動車 80 は、組み込みマイクロプロセッサを有する。それは、他の機能の中でも特に、図 1 のアルゴリズムを実行するようにプロセッサに命令する、非一時的なコンピュータ読取可能媒体に記憶されたソフトウェアを実行する。

#### 【0060】

##### コンピュータ実装

いくつかの実施形態では、ネットワークを介して接続されたクライアントおよびサーバが用いられ、ユーザはクライアントと対話するが、サーバが発話を処理してそれらの意味を解釈する。いくつかのサーバは、データセンタにおけるプロセッサのラックであり、多数のクライアントおよび多数のユーザからの発話を同時に処理可能である。いくつかの実施形態は、ユーザから直接発話を取込む同一のデバイス上で、すべての処理を局所的に行なう。

10

#### 【0061】

スマートフォン、自動ホームアシスタントデバイス、および自動車などの多くの種類のユーザ - 対話型デバイスが利用可能である。自動車およびスマートフォンなどのいくつかのモバイルデバイスは、サーバに接続するモバイルネットワークの範囲内にある場合はクライアントのみとして用いられ得るが、モバイルネットワークの範囲外にある場合はローカルのみする方法を用いることができる。自動車は、従来のシステムにとっては特に困難な条件を生じさせ、特に本発明の利益を享受する。なぜなら、安全運転のためには、ディスプレイを必要としないこと、および、ユーザが注意散漫になったり、混乱したり、苛立ったりしないことが求められるからである。

20

#### 【0062】

図 9 は、さまざまな実施形態に好適な非一時的なコンピュータ読取可能媒体の例を示す。いくつかの実施形態は、磁気ディスクプラッタ 91 を用いてコンピュータコードを記憶および提供する。いくつかの実施形態は、フラッシュ RAM ストレージデバイス 92 を用いてコンピュータコードを記憶および提供する。

#### 【0063】

図 10 は、パッケージ化されたシステムオンチップデバイス 101 を 2 つの向きで示す。これらのデバイスは、マイクロホン、スピーカ、ネットワーク接続、非一時的なコンピュータ読取可能媒体、および他の入力 / 出力装置への結合を提供するプリント回路基板に取付けられるボール・グリッド・アレイ (ball grid array) を有する。図 10 において、一方の向きはボール・グリッド・アレイを示し、他方の向きはプリント回路基板とは反対の方向を向くデバイス面を示す。

30

#### 【0064】

図 11 は、システムオンチップデバイス 101 内のハードウェア機能モジュールのブロック図を示す。デバイス 101 は、2 つのコンピュータプロセッサ 111 および 112 を含む。いくつかの実施形態では、プロセッサは ARM 命令セットアーキテクチャに従う。いくつかの実施形態では、プロセッサは x86 命令セットアーキテクチャに従う。インターコネクト 113 は、コンピュータプロセッサ 111 および 112 を、他のハードウェアモジュールに接続する。1 つのモジュールは、ランダムアクセスメモリ (RAM) 114 に対するインターフェースコントローラである。別のモジュールは、フラッシュ RAM および磁気ハードディスクドライブなどの不揮発性メモリ 115 に対するインターフェースである。1 つのモジュールは、システムオンチップをマイクロホン、スピーカ、ディスプレイ、および手動入力装置に結合する I/O コントローラインターフェース 116 である。1 つのモジュール 117 は、ワイヤレスネットワークモデムおよびイーサネット (登録商標) ポートなどの、インターネットを介してシステムオンチップ 101 を遠隔サーバに結合可能なネットワーキングデバイスに対するインターフェースである。

40

#### 【0065】

いくつかの実施形態は、図 12 に示すような、複数のチップおよび他のコンポーネントを備えた汎用コンピュータシステム上で行なわれてもよい。図 12 は、そのようなコンピ

50

ユーザシステムの例のブロック図である。コンピュータシステム 1210 は、典型的には、バスサブシステム 1212 を介して複数の周辺機器と通信する少なくとも一つのプロセッサ 1214 を含む。これらの周辺機器は、例えば、ストレージサブシステム 1224 (例えば、メモリデバイスおよびファイルストレージサブシステムを含む)、ユーザインターフェース入力装置 1222、ユーザインターフェース出力装置 1220、およびネットワークインターフェースサブシステム 1216 を含む得る。入力装置および出力装置は、コンピュータシステム 1210 とのユーザ対話を可能にする。ネットワークインターフェースサブシステム 1216 は、外部ネットワークに対するインターフェース (通信ネットワーク 125 に対するインターフェースを含む) を提供し、通信ネットワーク 125 を介して他のコンピュータシステムにおける対応のインターフェースデバイスに結合される。

10

**【0066】**

ユーザインターフェース入力装置 1222 は、キーボード、ポインティングデバイス (例えばマウス)、トラックボール、タッチパッドまたはグラフィックタブレット、スキャナ、ディスプレイと一体化されたタッチスクリーン、音声入力装置 (例えば音声認識システム、マイクロホン、および他の種類の入力装置を含む得る。一般的に、「入力装置」という用語の使用は、コンピュータシステム 1210 内または通信ネットワーク 125 上に情報を入力するすべての可能な種類のデバイスおよび方法を含むことを意図している。

**【0067】**

ユーザインターフェース出力装置 1220 は、ディスプレイサブシステム、プリンタ、ファックス装置、または非視覚表示 (例えば音声出力装置) を含む得る。ディスプレイサブシステムは、陰極線管 (CRT)、フラットパネルデバイス (例えば、液晶ディスプレイ (LCD)、投影機、または、可視画像を作成するための何らかの他のメカニズムを含む得る。また、ディスプレイサブシステムは、例えば音声出力装置による非視覚表示を提供し得る。一般的に、「出力装置」という用語の使用は、コンピュータシステム 1210 から、ユーザまたは別の装置またはコンピュータシステムに情報を出力するすべての可能な種類のデバイスおよび方法を含むことを意図している。

20

**【0068】**

ストレージサブシステム 1224 は、本明細書に記載の処理に係る質問提案のための推測質問を作成するロジックを含む、本明細書に記載のモジュールのうちのいくつかまたはすべての機能を提供するプログラミング構造およびデータ構造を記憶する。これらのソフトウェアモジュールは、一般的に、プロセッサ 1214 単独によって、または他のプロセッサとの組み合わせによって実行される。

30

**【0069】**

ストレージサブシステムで使用されるメモリ 1226 は、プログラム実行中に命令およびデータを記憶するためのメインランダムアクセスメモリ (RAM) 1230 と、固定命令が記憶されている読み出し専用メモリ (ROM) 1232 とを含む複数のメモリを含む得る。ファイルストレージサブシステム 1228 は、プログラムおよびデータファイルのための永続ストレージを提供することができ、ハードディスクドライブ、フロッピー (登録商標) ディスクドライブ (関連する取外し可能媒体と併用)、CD-ROM ドライブ、光学ドライブ、または取外し可能媒体カートリッジを含む得る。ある実施形態の機能を実現するモジュールは、ストレージサブシステム 1224 におけるファイルストレージサブシステム 1228、または、プロセッサがアクセス可能な他の装置に記憶され得る。

40

**【0070】**

バスサブシステム 1212 は、コンピュータシステム 1210 のさまざまなコンポーネントおよびサブシステムを意図通りに互いに通信させるためのメカニズムを提供する。バスサブシステム 1212 は単一のバスとして概略的に示されているが、バスサブシステムの代替的な実施形態は、複数のバスを用いてもよい。

**【0071】**

コンピュータシステム 1210 は、ワークステーション、サーバ、コンピューティングクラスタ、ブレードサーバ、サーバファーム、または他の任意のデータ処理システムもし

50



くはコンピューティングデバイスを含むさまざまな種類のものであり得る。コンピュータおよびネットワークの性質は刻々と変化するので、図12に示すコンピュータシステム1210の図は、好ましい実施形態を説明するためのほんの一例として意図したものである。図12に示したコンピュータシステムよりも多い、または少ないコンポーネントを有する、他の多数のコンピュータシステム1210の構成が可能である。

【0072】

本出願に記載された実施形態および例は、実現可能な唯一のものではない。当業者ならば、本出願に記載された実施形態および例に従って、他のさまざまな適用例を供する他の実施形態および変形例を認識するであろう。変更例および変形例は、開示された特徴の任意の関連する組み合わせを含む。

【図1A】

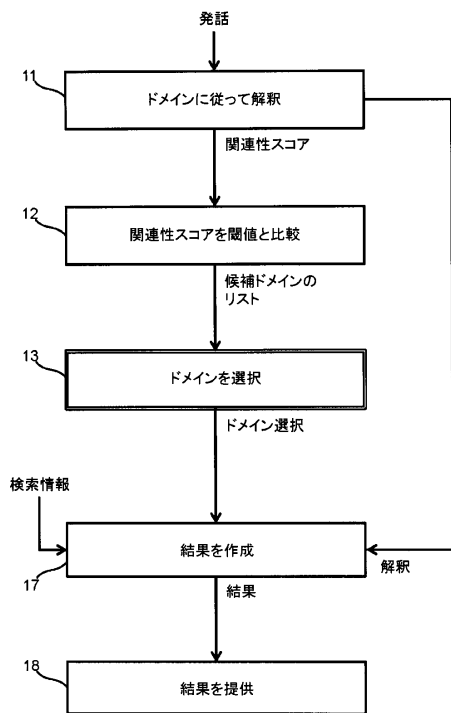


FIG. 1A

【図1B】

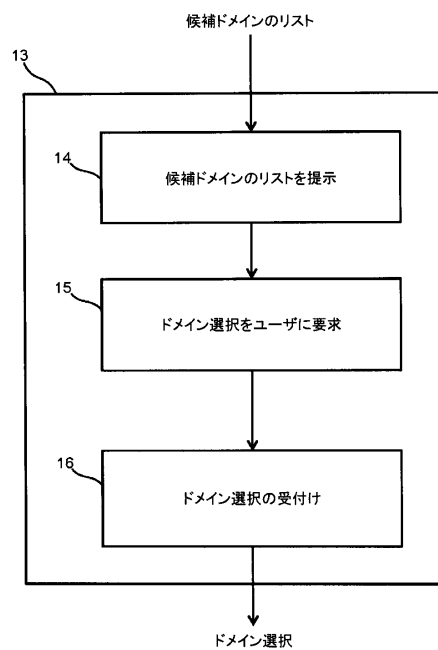


FIG. 1B

【図 2 A】

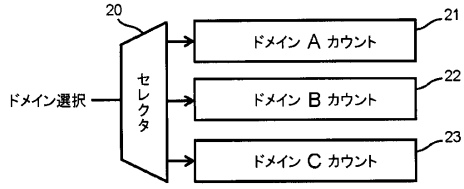


FIG. 2A

【図 3】

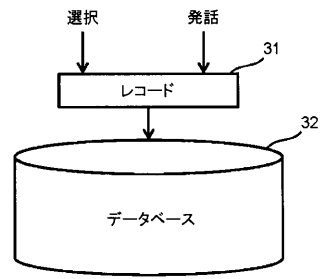


FIG. 3

【図 2 B】

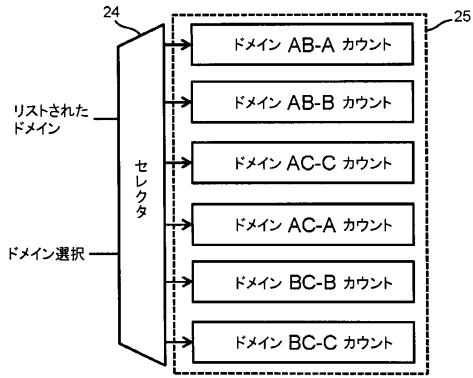


FIG. 2B

【図 4】

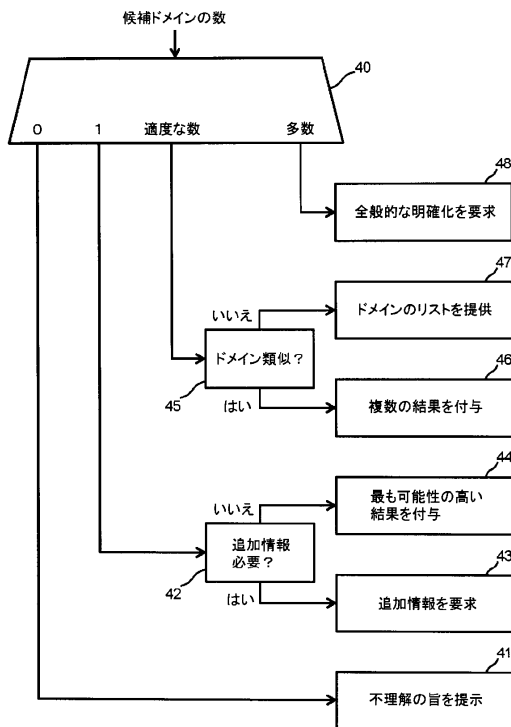


FIG. 4

【図 5】

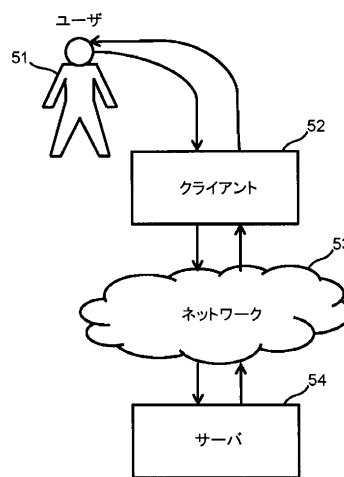


FIG. 5

【図6】

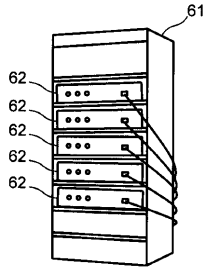


FIG. 6

【図7】

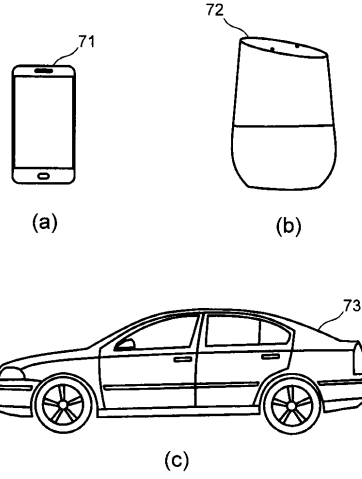


FIG. 7

【図8】

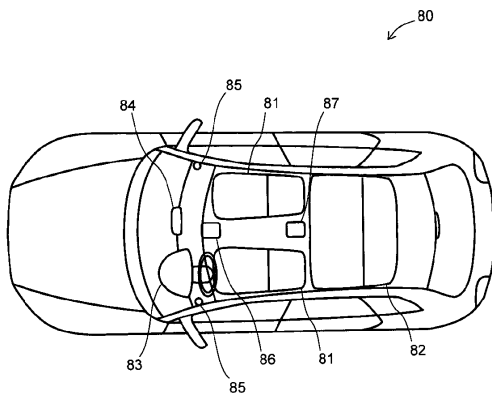


FIG. 8

【図10】

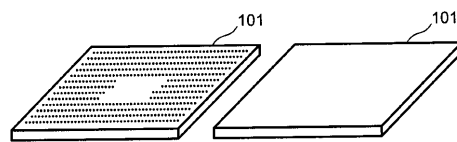


FIG. 10

【図9】

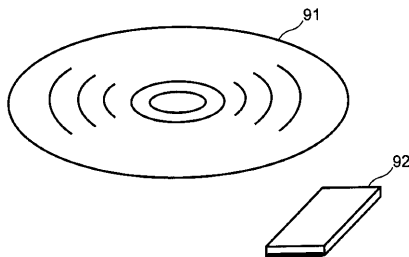


FIG. 9

【図11】

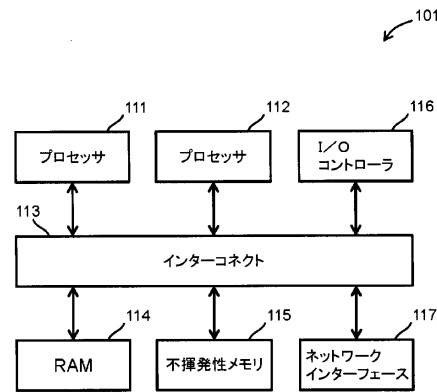


FIG. 11

【 図 1 2 】

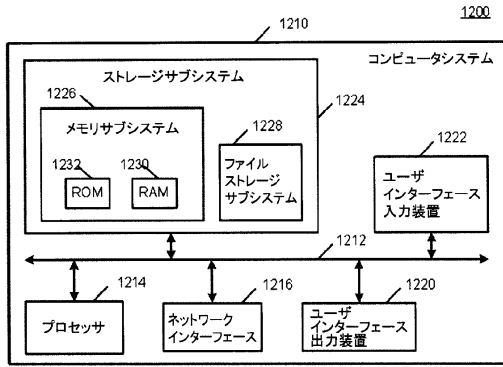


FIG. 12 - コンピュータシステム

---

フロントページの続き

- (56)参考文献 特開2009-009170(JP,A)  
国際公開第2007/088877(WO,A1)  
特開2010-078934(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10L 15/10  
G10L 13/00  
G10L 15/22