

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2022-531994  
(P2022-531994A)

(43)公表日 令和4年7月12日(2022.7.12)

(51)国際特許分類 F I テーマコード(参考)  
G 0 6 F 16/90 (2019.01) G 0 6 F 16/90 1 0 0 5 B 1 7 5

審査請求 未請求 予備審査請求 未請求 (全29頁)

(21)出願番号	特願2022-512710(P2022-512710)	(71)出願人	522087327
(86)(22)出願日	令和2年6月30日(2020.6.30)		サピエントエックス インコーポレイテッド
(85)翻訳文提出日	令和3年12月28日(2021.12.28)		アメリカ合衆国、95060、カリフォルニア州 サンタ クルーズ、スミス グレイド、4600
(86)国際出願番号	PCT/US2020/040375	(74)代理人	110000877
(87)国際公開番号	WO2020/223742		龍華国際特許業務法人
(87)国際公開日	令和2年11月5日(2020.11.5)	(72)発明者	コリーン、デイビッド
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,B	(72)発明者	マーヴィット、マクレーン
	最終頁に続く		アメリカ合衆国、95060 カリフォルニア州、サンタ クルーズ スミス グレイド ロード 4600
			最終頁に続く

(54)【発明の名称】 人工知能ベースの会話システムの生成および動作

(57)【要約】

コンピュータプロセスは、ユーザが会話プログラムを用いて、ユーザのデバイスの1または複数においてタスクを遂行するのに後で呼び出すことができる情報およびプロセスで、自動アシスタントを動的にプログラムすることを提供する。会話プログラムは、テキストまたはグラフィックソースから、自動アシスタントを生成してもよい。会話プログラムは、複数の自動アシスタントにアクセスし、ユーザの要求をアドレスするために使用するのにどれが最も適切であるかを決定する。ユーザは、視覚的文脈で表示することができる、人間の感情、顔の表情、および身振りに匹敵する、会話プログラムのためのアバターを生成することができる。結果として得られる自動アシスタントは、ソフトウェアシステム試験に使用することができる。

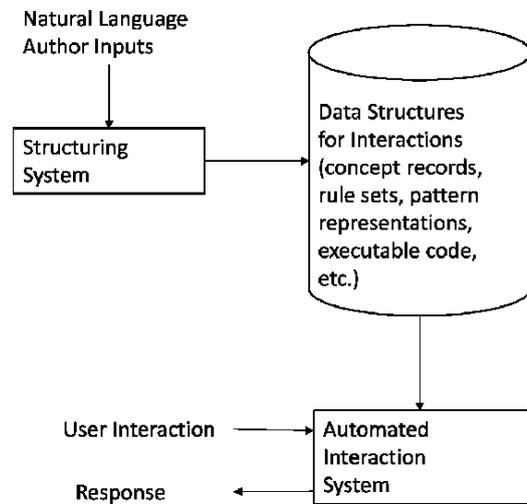


Fig. 15

**【特許請求の範囲】****【請求項 1】**

コンピュータ実装方法であって、  
実行可能命令で構成された 1 または複数のコンピュータシステムの制御下で、  
1 または複数の知識ドメインのセットから選択された知識ドメインを選択するように、  
オーサリングユーザに促す段階と、  
前記オーサリングユーザの、前記選択された知識ドメインの選択を受信する段階と、  
前記オーサリングユーザからオーサリングユーザ入力を受信する段階と、  
前記オーサリングユーザ入力を、オーサリングされた自動統合システムによって使用可能  
な、構造化された形態の複数のテキスト出力へと変換する段階と  
を備える、会話システムを生成する方法。

10

**【請求項 2】**

前記オーサリングユーザ入力を複数のテキスト出力へと変換する段階であって、第 1 のテ  
キスト出力が第 1 の認識システムの第 1 の出力であり、第 2 のテキスト出力が第 2 の認識  
システムの第 2 の出力である、段階と、  
前記選択された知識ドメインのドメイン仕様に基づいて、ドメイン固有の計画を作成する  
段階と、  
ランタイム仕様を取得する段階であって、前記ランタイム仕様が、前記選択された知識ド  
メインに対する、前記ドメイン固有の計画に基づいて構成された計画タスクフローを有す  
る、段階と、  
前記計画タスクフローを実行する段階と、  
前記オーサリングユーザ入力から入力値を生成する段階と、  
前記入力値に基づいて前記会話システムを改善する段階と、  
前記会話システムの説明をコンピュータ可読メモリに格納する段階と  
を更に備える、請求項 1 に記載の方法。

20

**【請求項 3】**

前記オーサリングユーザからの前記オーサリングユーザ入力が入力または音声入力またはテキスト入  
力を有する、請求項 1 または 2 に記載の方法。

**【請求項 4】**

第 1 の認識システムおよび第 2 の認識システムが、自動発話認識システムまたは画像認識  
システムの 1 もしくは複数である、請求項 1 から 3 のいずれか一項に記載の方法。

30

**【請求項 5】**

推論モジュールを使用し、前記会話システムと対話する対話中のユーザからの入力に基づ  
いて、計画タスクフローを動的に訂正する段階を更に備える、請求項 1 から 4 のいずれか  
一項に記載の方法。

**【請求項 6】**

前記オーサリングユーザから、1 または複数の第 1 の選択肢の第 1 のセットの中からの、  
選択された選択肢の第 1 のオーサリングユーザ選択を取得する段階と、  
前記第 1 のオーサリングユーザ選択に基づいて、計画タスクフローを調節する段階と、  
データマイニングモジュールを使用して、格納されたドメイン知識レポジトリを作成する  
段階と  
を更に備える、請求項 1 から 5 のいずれか一項に記載の方法。

40

**【請求項 7】**

前記データマイニングモジュールが、構造化テキスト、非構造化テキスト、および/または  
グラフィックのうち 1 もしくは複数を使用し、前記データマイニングモジュールの計算  
が前記会話システムの出力を変更する、請求項 6 に記載の方法。

**【請求項 8】**

クラウドソーシングによる知識ソースランキングシステムを用いて、自動ドメイン知識ソ  
ースモジュールを使用してドメイン固有の計画が生成され、前記方法が更に、  
複数の知識ソースそれぞれに対する採点値を導出する段階と、

50

前記自動ドメイン知識ソースモジュールを使用して、前記採点値に基づいて、複数のソースの中から使用する選択されたソースを動的に決定する段階と、  
前記選択されたソースを前記会話システムの出力値にマッピングする段階とを備える、請求項 1 から 7 のいずれか一項に記載の方法。

【請求項 9】

ユーザ入力に基づいて会話プログラムを動的に改善するシステムであって、  
1 または複数のプロセッサと、

実行されると前記 1 または複数のプロセッサに、

a) ユーザ入力に基づいて意図を形成させ、

b) 前記意図に基づいて計画を作成させ、前記計画が、前記意図と関連付けられた第 1 の概念オブジェクトを第 2 の概念オブジェクトへと変換する第 1 のアクションオブジェクトを有し、前記第 2 の概念オブジェクトを前記意図の目標と関連付けられた第 3 の概念オブジェクトへと変換する第 2 のアクションオブジェクトを有し、前記第 1 のアクションオブジェクトおよび前記第 2 のアクションオブジェクトが複数のアクションオブジェクトから選択され、前記第 1 のアクションオブジェクトが第 1 の第三者デベロッパによって提供され、前記第 2 のアクションオブジェクトが第 2 の第三者デベロッパによって提供され、

c) 前記計画を実行させ、

d) 前記第 3 の概念オブジェクトと関連付けられた値を出力させる、

複数の命令を格納する、非一時的コンピュータ可読媒体と

を備える、システム。

【請求項 10】

前記第 1 の概念オブジェクトが第 3 の第三者デベロッパによって提供され、前記第 2 の概念オブジェクトが第 4 の第三者デベロッパによって提供され、前記第 3 の概念オブジェクトが第 5 の第三者デベロッパによって提供される、請求項 9 に記載のシステム。

【請求項 11】

前記第 1 の概念オブジェクトが、前記第 1 の概念オブジェクトのインスタンス化を提供する第 1 のデータを含み、前記第 2 の概念オブジェクトが、前記第 2 の概念オブジェクトのインスタンス化を提供する第 2 のデータを含み、前記第 3 の概念オブジェクトが、前記第 3 の概念オブジェクトのインスタンス化を提供する第 3 のデータを含む、請求項 9 または 10 に記載のシステム。

【請求項 12】

前記第 1 のアクションオブジェクトの入力パラメータがウェブサービスパラメータにマッピングされ、ウェブサービス結果が前記第 1 のアクションオブジェクトの出力値にマッピングされる、請求項 9 から 11 のいずれか一項に記載のシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、全体として、人工知能会話システムを作成する方法、システム、およびプログラムに関する。

関連出願の相互参照

【0002】

本出願は、2019年5月2日付けの米国仮特許出願第62/842,423号の利益を主張し、全ての目的のため参照により本明細書に組み込む。

【背景技術】

【0003】

人工知能(「AI」)会話プログラムは、情報検索、技術支援(例えば、カスタマーサポート)、デバイスの制御、媒体制御、ゲームプレイ、ストーリーテリング、および人間との一般的な会話を提供するのに使用されることがある。いくつかのAI会話プログラムは、「チャットボット」、「自動アシスタント」、「デジタルアシスタント」、「対話型パーソナルアシスタント」、「インテリジェントパーソナルアシスタント」、「パーソナル

10

20

30

40

50

音声アシスタント」、「会話エージェント」などと呼ばれることがある。かかるシステムは、オーディオ、テキスト、機械的またはセンサ入力を受け入れ、同じように応答し、ならびに／あるいは記録をデータベースに格納するなど、対話の記録または説明をメモリに格納してもよい。

【0004】

異なる会話システムが、異なる目的のため、または異なる入力言語に対して作成されてもよい。ユーザが、ユーザの目的または言語に適合しない会話システムと対話している場合、結果は不満足なものであることがある。

【0005】

会話プログラムが利用する、人間のような音声またはテキスト応答を手動でオーサリングするプロセスは、時間がかかり高価である。例えば、ユーザ要求を聞き、適切に応答する自動対話システムは、プログラムのチームに、可能性がある全てのユーザ要求および適切な応答を入力させることによって、作成されることがある。これは、ユーザが日にちおよび都市を提示し、その日のその都市における予想天気を自動対話システムに応答させることができる、自動対話システムなど、レパトリが限定されたシステムの場合に実現可能なことがある。しかし、可能性がある対話のはるかに広範囲である場合、かかる自動対話システムをオーサリングするのは非常に時間がかかる場合がある。

【0006】

自動対話システムのオーサリングは時間がかかるため、オフラインで行われる場合が多く、予期されないユーザ入力に対する新しい応答は生成することができない。準備された応答の大きなコーパスが作成されてもよいが、コーパスは必然的に、予測されるユーザ対話に対する応答のみを提供することになり、その有用性が限定される。限定された予測可能な入力を取るシステムは、一般に「チャットボット」と呼ばれる。システムの可能な応答が限定され、誤った応答を送達することがあるので、システムは人間のユーザを苛立たせる場合がある。いくつかのシステムは、可能性がある回答に関してウェブページおよび他のソースからデータを抽出しようと試みることがあるが、これらのシステムは役に立たない結果を提供する場合がある。統計的な機械学習アプローチを使用する現在のシステムは、通常の会話パターンでユーザに応答できないことがある。自動対話システムを構築する際のダイアログおよび応答作成に対する、自動化されたリアルタイムアプローチが必要とされている。

【0007】

加えて、ユーザは複数のデバイスと毎日対話し、各デバイスは、対話の嗜好およびモードの自身のセットを有する。これには、ユーザが各デバイスを別個に構成または訓練する必要がある。これらのデバイスそれぞれをどのように使用するかを学習するのにかかる時間は、ユーザにとって苛立たしいものであり得る。

【発明の概要】

【0008】

コンピュータ実装方法では、自動対話システムのオーサリングシステムは、自動対話システムの動作を駆動するのに使用される構造化データを生成し、その構造化データは、自然言語入力からの構造化システムによって形成される。自動対話システムとの対話は、自然言語入力、構造化入力などの形態であることができる。

【0009】

自動対話システムは、発話認識モジュールを使用して入力を取得することがある。自動対話システムは、提供される出力が受信された入力に対する適切および／または有用な応答であろうということを、自動対話システムがどの程度確信しているかを決定する計算を表す、出力に対する信頼値とともに、入力を取り込み、出力を与えることがある。

【0010】

いくつかの実施形態では、自動対話システムは、企業体などの特定のエンティティによる使用のため、エンティティの顧客またはユーザが自動対話システムに入力を与え、応答出力が返されるであろうとの予想によって、オーサリングされる。具体的な例では、ユーザ

10

20

30

40

50

は顧客であり、テキスト、音声などで自動対話システムに質問し、自動対話システムは、自動対話システムが質問に対する応答とみなすテキスト、音声などを出力する。

【0011】

自動対話システムは会話システムであってもよい。

【0012】

自動対話システムは、信頼値をそれぞれ有する複数の可能な応答を計算し、最も高い信頼値に対応する出力を出力してもよい。

【0013】

自動対話システムは、ユーザ音声入力など、ユーザからの入力を複数の自動発話認識装置で処理して、かかる処理に対して対応する信頼値を決定し、そのユーザとの将来の対話のため、対応する最も高い信頼値に基づいて、複数の自動発話認識装置のうち1つの自動発話認識装置を使用してもよい。

10

【図面の簡単な説明】

【0014】

本開示による様々な実施形態について図面を参照して記載する。

【0015】

【図1】例示のコンピュータシステムを示す図である。

【0016】

【図2】インターネット接続を有する、例示の車両、ロボット、または医療用デバイスのシステムを示す図である。

20

【0017】

【図3】インターネット接続を有さない、例示の電気機器、医療用デバイスまたは、飛行機を示す図である。

【0018】

【図4】インターネット接続を有さない例示の自動販売機システムを示す図である。

【0019】

【図5】インターネット接続を有する例示の自動販売機システムを示す図である。

【0020】

【図6】インターネット接続および補助ロボットを有する例示の自動販売機システムを示す図である。

30

【0021】

【図7】インターネット接続を有する例示の自動二輪車およびヘルメットのシステムを示す図である。

【0022】

【図8】インターネット接続を有する例示の代替的な自動二輪車およびヘルメットのシステムを示す図である。

【0023】

【図9】インターネット接続を有する例示の携帯電話またはスマートウォッチのシステムを示す図である。

【0024】

40

【図10】インターネット接続を有する例示の自動二輪車およびヘルメットのシステムを示す図である。

【0025】

【図11】テレフォニー用の例示の対話型音声応答または電話会議のシステムを示す図である。

【0026】

【図12】例示のスマートスピーカーまたはビデオ会議のシステムを示す図である。

【0027】

【図13】例示の会話システムを示す図である。

【0028】

50

【図14】例示のソーシャルメディアシステムを示す図である。

【0029】

【図15】自動対話システムによって使用可能なデータを構造化システムが生成する、システムを示す図である。

【0030】

以下の詳細な記載は添付図面とともに、本発明の性質および利点のより良い理解を提供する。

【発明を実施するための形態】

【0031】

本明細書に記載するシステムを使用して、会話型システムであってもよい改善された自動対話システムは、オンデマンドの情報および娯楽を提供し、電話システムでメニューの選択肢を与え、デバイスの基本制御を提供し、身体障害があるユーザに代替のコンピュータインタフェースを提供し、ならびに/あるいはユーザが話している言語を自動的に検出し、適宜反応することができる。他の入力、ユーザの感情状態を示すデータを含んでもよく、かかるデータは、かかる状態を決定する心情分析または他の技術を使用して生成されてもよく、それを次に、自動対話システムの出力を調節するのに使用することができる。

10

【0032】

図1は、本発明の一実施形態による、ユーザを別のユーザとの会話チャンネルに加入させるために実装されるプロセスのブロック図を示している。プロセスは会話管理サーバで実施される。プロセス11において、サーバは、ユーザの制御下で、ユーザデバイスが、管理された会話セッションをサポートするように構成されたプロセスを実行していることを示す、探索要求を、第1のネットワークを通じてデバイスから受信する。本明細書に記載されるように管理された会話セッションをサポートするように構成されたプロセスをユーザデバイスが実行していることを示す、ユーザデバイスから発する信号などの探索信号は、ユーザが対話しようとするリモートユーザに関する基準を特徴付けるデータを含むことができる。いくつかの例では、そのリモートユーザは、本明細書に記載するようにオーサリングされた自動対話システムである。

20

【0033】

会話セッションには、コンピュータインタフェースと対話する参加者が関与し、入力および出力の1または複数のサイクルで、参加者が何らかの入力を提供し、何らかの出力が返される。対話は、ネットワーク接続を通じてのものであってもよく、入力および/または出力は、テキスト、映像、オーディオ、および/またはデータの形態であってもよい。いくつかの会話セッションでは、2つの参加者が存在し、またいくつかの会話セッションでは、少なくとも1つの参加者が人間であり、少なくとも1つの参加者がコンピュータプロセスである。会話セッションは、会話セッションと関連付けられた論理的または物理的会話チャンネルを有してもよく、参加者は会話サーバに登録されてもよい。いくつかの実施形態では、会話サーバは、誰が会話セッションの参加者であるかに関するデータを維持する。一般的な例では、参加者に対してインタフェースを提供するハードウェア/ソフトウェアは、会話のノードに対応する「シート」とみなされてもよい。

30

【0034】

いくつかの例では、会話チャンネルは、特定の企業または組織（会話チャンネルの受益者）によってそのために操作され、会話チャンネルが利用可能であって、会話チャンネルと関連付けられた会話チャンネルの受益者に対してオープンであるか否かを表す「利用可能性ステータス」など、その状態と関連付けられたデータを有する。他の状態データは、会話チャンネルの利用可能性の格付け、会話チャンネルの受益者の応答性の格付け、および会話チャンネルの受益者と関連付けられたルールセットを含んでもよい。特定の実施形態では、利用可能性ステータスは、「離席中」、「不在」、「取込み中」、「オンライン」、および「メッセージは受信できるが直ぐに会話には応答不能」を含む、ステータスのセットから選択されてもよい。会話の関係者は、会話チャンネルをそれらの関係者に対してオープンにするために、必ずしも同時にオンラインである必要はない。

40

50

## 【0035】

ユーザによっては、会話セッション管理システムにアカウントを有する登録ユーザであってもよく、またユーザによっては、関連付けられた会話チャンネルを有する、要求される企業を有する企業ユーザの被雇用者など、特定の会話チャンネルの受益者と関連付けられてもよい。

## 【0036】

会話チャンネルは、コンピュータシステムによって、ユーザと、会話に関与するユーザによって選択された企業のシートとの間の通信路とみなされてもよい。要求される企業は、少なくとも1つ会話チャンネルを有してもよく、また要求される企業は、複数の会話チャンネルを定義してもよく、各会話チャンネルを企業との通信のカテゴリに割り当てることができる。各チャンネルは、ユーザに見えるラベルを備えることができるので、ユーザは、要求される企業の所望のチャンネルを選択し、それを通じて会話に関与することができる。

10

## 【0037】

会話セッションのシートは、個人、コンピュータプロセス、またはインストールされたアプリケーションなどのシートオペレータが配属される会話ノードに対応してもよく、シートオペレータは、会話セッションにおける所与の企業の代表であり、またはそれに対するインタフェースである。ユーザが複数のシートを有する企業と会話しようとする場合、会話管理サーバが、そのユーザを加入させて、ユーザとの会話に現在関与していない（または会話に関与するのに十分な容量を現在有する）企業のシートと会話させてもよい（当然ながら、会話中の企業を表す個人がシートを使っているものと仮定する）。

20

## 【0038】

シートは、シートと関連付けられた多数の同時会話セッションが予め指定された閾値に達した場合、ビジー状態を有してもよい。

## 【0039】

会話ホストは、会話をホストする個人アカウントが確立されており、少なくとも1つのシートが割り当てられている、個人またはグループであってもよい。

## 【0040】

コンピュータプロセスは、ソフトウェアもしくはファームウェア、またはそれらのいずれかの組み合わせの制御下で動作しているか、あるいは上記のいずれかの制御外で動作しているてもよい、コンピュータハードウェア（プロセッサ、フィールドプログラマブルゲートアレイ、もしくは他の電子組み合わせ論理、または類似のデバイス）を使用して、記載される機能をコンピュータで実行してもよい。記載される機能の全てまたは一部は、トランジスタもしくは抵抗器など、能動または受動電子構成要素によって実行されてもよい。コンピュータプロセスは、スケジュール可能なエンティティ、またはコンピュータプログラムもしくはその一部の動作を必ずしも示唆しないが、いくつかの実施形態では、コンピュータプロセスは、かかるスケジュール可能なエンティティ、またはコンピュータプログラムもしくはその一部の動作によって実装されてもよい。プロセスは、1つを超えるプロセッサまたは1つを超える（シングルもしくはマルチプロセッサ）コンピュータを使用して実行されてもよい。

30

40

## 【0041】

ユーザデバイスは、ユーザによって使用され、デスクトップユニット、ラップトップユニット、タブレット、スマートフォンとして、またはネットワークへのアクセスを有する他の任意のコンピュータとして実装される、コンピュータであってもよい。

## 【0042】

以下の記載で、様々な実施形態について記載する。説明目的で、実施形態の徹底的な理解を提供するために、具体的な構成および詳細について説明する。しかしながら、実施形態は具体的な詳細なしで実施されてもよいことも、当業者には明白となるであろう。更に、記載する実施形態が曖昧にならないようにするため、良く知られている特徴は省略または単純化されることがある。

50

## 【 0 0 4 3 】

例示のシステムの様々な構成要素は、ユーザがコンピュータ化された対話型システムと対話してもよい、コンピュータ化された対話型システムを形成するため、単独でまたは他の要素との組み合わせで使用されてもよく、コンピュータ化された対話型システムは、そのユーザからの入力を取り込み、何らかの処理および/またはデータルックアップを実行し、次にそのシステム出力をオーディオおよび/または映像の形態で、場合によっては他の出力とともに出力し、それによってユーザは、ユーザに適切に応答できる何らかの知能を有するものとして、コンピュータ化された対話型システムを知覚してもよい。これを遂行するために、コンピュータ化された対話型システムは、いくつかの訓練モードならびに動作モードを有してもよい。例えば、コンピュータ化された対話型システムは、最初に、音声フレーズを出力することができるように、また次に動作モードでそれらの音声フレーズを使用して、訓練されてもよい。

10

## 【 0 0 4 4 】

本明細書に記載する技術は、ユーザが会話プログラムを用いて、ユーザのデバイスの1または複数においてタスクを遂行するのに後で呼び出すことができる情報およびプロセスで、自動アシスタントを動的にプログラムすることを可能にすることに関する。別の実装では、会話プログラムは、ウィキペディアなどのテキストまたはグラフィックソースから、自動アシスタントを生成するであろう。更に、会話プログラムは、複数の自動アシスタントにアクセスし、ユーザの要求をアドレスするために使用するのにどれが最も適切であったかを決定することができるであろう。別の実装は、ユーザが、会話プログラムのため、多量の視覚的文脈で表示することができる人間の感情、顔の表情、および身振りに匹敵する、アバターを生成するのを可能にするであろう。結果として得られる自動アシスタントはまた、ソフトウェアシステム試験に使用することができる。

20

## 1. 信頼値に基づいた自動発話認識エンジンの選択

## 【 0 0 4 5 】

いくつかの自動会話システムでは、複数のAI会話システムが利用可能であってもよい。ユーザを支援するのに最適なAI会話システムを選ぶことで、改善されたユーザ体験を作成することができる。

## 【 0 0 4 6 】

一実施形態では、会話システムは、例えば、異なる言語、同じ言語内の異なるアクセント、および同じ言語の方言に関して、複数の自動発話認識エンジン(「ASR」)へのアクセスを有してもよい。例えば、システムは、ポルトガル語、スペイン語、アメリカ英語、スコットランド英語、およびスペイン語圏の国出身の非ネイティブ話者が話す英語それぞれに対して、異なる自動発話認識システムを有してもよい。会話システムはまた、子ども用のASR、発話障害を持つ人向けのASRなどを有してもよい。

30

## 【 0 0 4 7 】

各ASRエンジンは、音声入力を受信し、それを出力テキストに翻訳する。ASRはまた、出力テキストに、例えば0~1の範囲であってもよいが他の範囲が想到される、信頼値をラベリングする。音声入力会話システムによって受信されると、ASRのサブセットまたは全てが、それをテキストに翻訳するのに使用されて、各ASRに対する出力および信頼値がもたらされてもよい。最上位のASRからの出力が出力として選択されてもよい。ユーザが更に音声入力を入力した場合、前に選ばれたASRにより大きい重みを与られてもよい。別の実施形態では、各音声入力は個別に扱われてもよい。ASRはまた、査定されてもよく、ASRが範囲内の信頼値を有する場合、別のASRの信頼区間が閾値の値分、その信頼値を超えるまで、前に選ばれたASRが使用されてもよい。

40

## 【 0 0 4 8 】

一実施形態では、ASRは、リモートデータセンタ(「クラウド」)に格納されてもよく、アクティブなASRのみがローカル使用のためにダウンロードされてもよいが、他の利用可能なASRの信頼区間を監視するのに、音声入力リモートデータセンタに送信されてもよく、異なるASRがアクティブなASRの信頼区間を(恐らくは閾値分)超えた場

50

合、異なる A S R がダウンロードされ、アクティブな A S R に取って代わるかまたはそれとともに稼働してもよい。ローカルシステムは、1つの A S R が、複数の音声入力にわたって、一連の高信頼区間を達成するかまたは統計的に顕著な高信頼区間を有するまでの期間、2つ以上の A S R を使用してもよい。

【0049】

別の実施形態では、会話システムが高帯域幅接続を利用可能である場合、複数の A S R がクラウド内で利用可能であってもよい。クラウドシステムは、どの A S R が最も高い信頼値を有するかを追跡し、高帯域幅接続が利用可能な場合、その A S R をローカル会話システムにダウンロードしてもよい。高帯域幅接続が利用可能でない場合、ローカル A S R が使用され、音声入力がキャッシュされてもよい。高帯域幅接続が再度利用可能になると、コールキャッシュされた対話がクラウドに送信されて評価されてもよく、異なる A S R が前の A S R よりも高い信頼値を有する場合、新しい A S R システムがローカル会話システムにダウンロードされてもよい。

10

2. 参照コーパスおよび S M E から会話システムを作成するための A I インタビュー

【0050】

一実施形態は、変換に使用可能なコーパスを生成する A I 駆動のインタビューシステムである。この実施形態では、ユーザは、テキストまたは発話のどちらかを介して、オーサリングシステムと対話する。オーサリングシステムは、例えば、新しい会話システムをどのように名付けるか、アバターをどのような見た目にするか、音声はどのような音にするべきか、対話の口語スタイルはどのようなものにするか、および新しい会話システムがどのような専門分野を有するかを含む、一連の質問を尋ねてもよい。可能なアバターは、ユーザがそこから選択してもよい可能なアバターおよびアバターの特徴のアバターデータベースに提供されてもよい。

20

【0051】

いくつかの実施形態では、自然言語オーサリング入力は、一部には、動作を自動対話システムに通知してもよい概念、ルール、パターン、および/またはアルゴリズムに対応する、構造化データを生成することによって、自動対話システムをオーサリングするのに使用される。オーサリング入力は、作者に、入力を発声によって生成させ、文、設計ルールを出力させる、特定の命令であってもよく、集計システムを使用して入力を提案してもよい。

30

【0052】

システムは、ユーザが異なる音声を選択し聞くことを可能にする、異なる音声特性によってインデックスを付けられた音声のデータベースを有してもよい。オーサリングシステムは、「くだけた」、「学術的」、および「簡潔な」など、利用可能な異なる会話スタイルを有してもよい。オーサリングシステムは、新しい会話システムがどの知識を提供するかをユーザがそこから選択する、既知の知識ベースのメニューを有してもよい。オーサリングシステムは、ユーザの応答をデータベースに格納してもよい。

【0053】

ユーザが新しい会話システムの性質を定義すると、標準的なグラフィックベースのユーザインタフェースと対話することによって、または音声によって、オーサリングシステムは、例えば情報を入力するようにユーザに求めることによって、更なる情報を知識ベースに追加するようにユーザに促してもよい。この情報は、知識ベースシステムに提供されてもよく、該システムが次に、オンラインソース、例えばウィキペディアから、更なる情報を抽出してもよい。知識ベースシステムは、他の会話システムのための知識ベースシステムと情報を交換してもよい。いくつかの実施形態では、知識ベースシステムは、書籍または技術文献に対するアクセスを有してもよく、知識ベースシステムはそれを使用して、入力情報を拡張し確認してもよい。競合するデータが見つかった場合、ユーザに競合が提示され、どの情報が優先するかを尋ねてもよい。知識ベースシステムはまた、映像ソース、ライダーソースなどから、新しい会話システムの話題に関連するテキストを抽出してもよい。

40

50

### 3. 音声オーサリングの構成要素

#### 【0054】

一実施形態では、オーサリングシステムによる会話システムの作成の一部として、オーサリングシステムは、内容領域専門家（「SME」）にインタビューし、応答を記録するように構成された、自身のAI会話インタビューシステムを有してもよい。インタビューシステムは、SME 応答に基づいて、尋ねるべき関連質問を発見する。

#### 【0055】

インタビューシステムは、対話型音声システムで有用であろう有声フレーズを含む可能性がより高い、SME からの応答を促す質問を生成するように構成される。例えば、インタビューシステムは、取得する必要がある有声フレーズの「買い物リスト」を有してもよく、SME は、質問に対する回答として話される、それらの有声フレーズに対応する尋ねるべき質問を決定する。

#### 【0056】

フレーズは、テキスト音声変換データベースに格納し、ならびに知識ベースに追加することができる。オーサリングシステムがSME にインタビューしている際、出力が知識ベースに格納されてもよく、同時に他の知識ベースと内容を比較して、例えば類似のデータがエントリされているかを検出し、その類似データを使用して、SME インタビュー中に焦点を絞った質問を生成してもよい。これらの焦点を絞った質問は、共通のエンドユーザの質問など、以前に作成された会話AIシステムとの過去のユーザ対話に基づいてもよい。

#### 【0057】

一例のシステムでは、知識ベースは、自動車ユーザに対する命令を生成するのに、発話システムによって使用されてもよい。発話システムは、いくつかの有声フレーズを既に有していることがあるが、知識システムの他の部分に次に必要ないくつかの有声フレーズは有さない。例えば、ユーザが車のシート調節とどのように対話するかを説明するデータベースが既に存在する場合、そのデータは、インタビュー質問において焦点を絞った質問を生成するのに使用されてもよい。

### 4. マルチモード会話の構成要素

#### 【0058】

マルチモード会話システムは、デバイスに依存しないユーザ音声インタフェースの個人化をコンパイルし、格納し、調査するのに使用されてもよい。

#### 【0059】

この実施形態では、ユーザの嗜好が、複数のデバイスからアクセス可能なデータストア、例えばブロックチェーン台帳に格納されてもよい。別の実施形態では、ユーザの嗜好は暗号化形式で、例えば対称鍵暗号または公開/秘密鍵暗号で格納されてもよい。セキュリティを向上するため、256ビットの暗号鍵が使用されてもよい。

#### 【0060】

一実施形態では、ユーザの誕生日、余暇活動、または好きな色などの嗜好情報が格納される。嗜好はデバイスまたはユーザ固有であってもよい。ユーザは、その日のうちに要求するための、AI会話システムの要求、例えばリマインダを行ってもよい。AI会話システムは、要求を受信し、それに時間タグをタグ付けして、情報を中央データベースに格納する。指定の時間に、ユーザが、ユーザが要求を行った元のAI会話システムの近くにいないということが起こり得る。システムは、これを認識し、応答のため、要求をユーザの近くのAI会話システムにルーティングする。

#### 【0061】

別の実施形態では、ユーザはAI会話システムに、例えば自身の好きな色が青であると伝える。色の嗜好は中央データベースに格納される。後で、ユーザが異なるAI会話システムと対話しているとき、ユーザは、ユーザの好きな色に関連する要求（例えば、「会話システム、私のUIに合った良い色のテーマを選んで」）を行ってもよい。この第2の会話システムは、中央データベースをポーリングして情報を見つけ、ユーザの好きな色である青に基づいてテーマを選んだというコメントを用いて、システムのためのUIを選んでも

10

20

30

40

50

よい。

5. ユーザの意図の決定、およびユーザの意図に基づいたAIバリエーションの選択

【0062】

会話システムは、ユーザの意図を決定し、ユーザの意図に基づいてAIバリエーションを選択する、プリプロセッサを含んでもよい。異なるAI会話システムが、他のシステムよりも、特定のタイプのユーザ対話に良く適合していることがある。プリプロセッサは、ユーザの意図を決定し、次にユーザの意図を分類してもよい。分類から、プリプロセッサは複数のAI会話システムの中から選択してもよい。ルールベースの前処理システムは、会話の一般的説明を生成してもよい。例えば、説明は、ユーザが技術的質問を尋ねていることを示し、技術文献から回答を抽出するように設計されたシステムを呼び出してもよい。技術的質問に回答するように設計されたシステムの一例は、技術文献に基づいて技術サポートの質問に回答する、スタンフォード大学で開発されたオープンソースの機械学習システムである。多くの異なるタイプのAIツールが存在するので、所与のユーザタスクに対して最良のAIツールを選択することで、向上したユーザ体験を提供してもよい。ユーザの意図（例えば、技術的質問に対する回答を見つける）の決定は、求められた情報をユーザに提供することに関して、選ばれたAI会話システムの成功を決定することができる。

10

【0063】

一実施形態では、プリプロセッサシステムは、いくつかの特化型AI会話システムから選択して、複数のAI会話システムを備えるメタAI会話システムを作成するであろう。プリプロセッサは以下の段階を行ってもよい。ユーザが、メタAI会話システムに質問する。プリプロセッサが、ユーザの質問の意図を分析し、それをタイプによって分類する。プリプロセッサシステムは、意図タイプを、所与の意図タイプに対する応答精度によってランク付けされた、オンボードまたはオンラインAI会話システムのリストと比較する。

20

【0064】

自動対話システムは更には、入力を処理し、複数の「考え方」のうち1または複数と一致する出力を提供してもよく、それにより、出力が特定の考え方と少なくともある程度一致する。

【0065】

精度は、広範囲の値を取り、一致の信頼度に対応する値を有することができる、可変値であってもよい。この値は、会話システムがユーザの質問に最初に十分一致しているかを決定する、一致の質の推定値として使用することができる。このリストから、プリプロセッサは、質問に回答するのに最良の会話システムを選択する。別の実施形態では、プリプロセッサは複数の会話システムを選択してもよい。質問は、選ばれた特化型会話システムに送信されて処理される。特化型会話システムが回答を生成すると、回答はメタ会話システムのダイアログジェネレータによって話される。複数の会話システムが使用される場合、応答がユーザの質問に一致しているとシステムがどのくらい確信しているかで、応答を（例えば、整数または浮動小数点のスコアを使用して）採点することができる。このスコアに基づいて、アクティブな会話システムを、前に選ばれた会話システムからより高い信頼スコアを有する会話システムへと変更してもよい。

30

6. 映像に基づいた拡張現実（「AR」）または仮想現実（「VR」）のための動的キャラクターの生成

40

【0066】

ARおよびVRは両方とも3D技術であるが、ARでは、背景はライブ映像フィードであり、VRでは、背景は、単一のパノラマ画像または予め記録されたビデオ画像であることができる。どちらの場合も、コンピュータ生成されたキャラクターが映像内に存在する。それらのキャラクターは、ストーリーもしくはゲーム内のガイド、アシスタント、またはキャラクターとして行動してもよい、AI駆動の会話キャラクターであることができる。他の表示デバイスは、コンピュータ画面、投影映像、網膜ディスプレイ、および混合現実（「XR」）を含んでもよい。

【0067】

50

キャラクタ生成システムは、実世界キャラクタ（ライブキャプチャまたは予め記録されたもののどちらか）をAR/VRフィールドに追加してもよい。キャラクタ生成システムは、3Dレンダリングされた顔を、映像内の実世界キャラクタの上に重畳し、重畳された顔のアニメーションをAIシステムで駆動してもよい。

【0068】

キャラクタ生成システムは、後の映像合成で使用することができるキャラクタを生成するため、マシンビジョンを使用して以下の段階を実行してもよい。最初に、キャラクタ生成システムは、映像シーン内の既存の顔を探してもよい。次に、キャラクタ生成システムは、顔色を分析して色パレット調節層を生成してもよい。次に、顔のシェーディングおよびシャドウイングを分析して、シェーディング調節層を生成してもよい。次に、映像シーン内の各顔の3Dモデルを表す既存の3Dメッシュを修正して、映像の顔（「標的の顔」）における重要な顔の特徴と位置合わせしてもよい。

10

【0069】

これらの段階が実行された後、映像合成システムは次に、映像の顔を、それらのセントロイド、ピッチ、ロール、およびヨーに関して校正してもよい。次に、それらの映像の顔に対するオクルージョンを計算してもよい。次に、エッジぼかしに役立つアルファチャンネルを含む、オクルージョンを説明するアニメーション化した「マスク層」を生成してもよい。映像合成システムは次に、操作されたジェネリック3D顔モデルを構築してもよく、それを次に、モーフまたは他の標的化アプローチを使用して、標的顔モデルに適合させてもよい。3D顔モデルを次に、AI会話システムからの出力によって生成されるアニメーションキューに基づいて、リアルタイムでアニメーション化してもよい。

20

【0070】

映像合成システムは、映像の顔の、それらのセントロイド、ピッチ、ロール、およびヨーに関する校正から、ジェネリック顔モデルに対する位置データを受信してもよい。映像合成システムは、顔色の分析から、色情報に基づいたテクスチャマップを受信し、テクスチャマップをジェネリック顔モデルに適用してもよい。映像合成システムはまた、シェーディング情報およびシャドウイング情報の分析からシェーディング情報を受信し、シェーディング情報をテクスチャマップに適用してもよい。映像合成システムはまた、オクルージョンがあればそれを適用してもよい。映像合成システムは、自然言語理解（「NLU」）エンジンを使用して顔をアニメーション化して、発話および顔の感情を表示してもよい。映像合成システムは次に、ジェネリック顔モデルをレンダリングし、アルファブレンディングを使用してそれを背景映像の上に合成して、映像のピクセルとブレンドしてもよい。映像合成システムは、準備された（つまり「予め用意された」）応答とは対照的に、レンダリングされたキャラクタが、AI会話システムの出力と一致する、それ自体の音声でダイアログを話しているように、付随する顔のアニメーションとともに見せてもよい。映像合成システムは、AIによって駆動される完全に動的なキャラクタを提供する。

30

7. データソースから会話出力を生成するオーサリングの構成要素

【0071】

オーサリングシステムは、車のマニュアルなどの技術文書を、機械学習を使用してデータベースに変換する方法を含んでもよい。AI会話システムは、データベースから情報を検索し、それを、ユーザによって尋ねられる音声質問に回答して音声オーディオとして出力することができる。会話システムは、ユーザ質問を、データベースに格納されたものからの、既知の人間の応答アプローチのセットと比較することによって始まる、方法を使用するであろう。会話システムは次に、ユーザの名前を追加する、または人間の会話をシミュレートするようにフォーマット化するなど、個人化を追加するであろう。システムは次に、数字または他の特定のデータを挿入して、既知の人間の応答における変数を記入するであろう。特定のデータは、機械学習によって作成されたデータベースによって供給されてもよい。会話システムは次に、応答が文法的に正しいことを検証してもよい。

40

【0072】

機械学習プラットフォームは、回答データの位置を識別することによって、技術的テキスト

50

トから質問に回答することに焦点が絞られたものであってもよい。会話システムは次に、これを会話としてフォーマット化する。例えば、「V8はシリンダをいくつ有する？」という質問を尋ねられると、機械学習システムは回答が8であると識別してもよい。会話システムは次に、ユーザの名前を含めて、「V8はV字形に配列された8つのシリンダを有する」という返答を公式化してもよい。システムはまた、より徹底した回答を組み立てるため、相互参照を実行してもよい。

#### 【0073】

自動車部品組立ての分解組立図などの製図は、分類、セグメント化、およびラベリング技術を使用して分析されてもよい。一例の用途として、車のマニュアルをデータベースに読み込んで、ユーザサポートが提供されてもよい。別の用途は、自動車用の部品カタログで

10

#### 【0074】

オーサリングシステムによって実行される方法は、OCR技術を使用して既存のマニュアルをスキャンすることを含んでもよい。マニュアルは、テキスト、グラフィック、およびグラフィックを説明文に関連させるラベルを有してもよい。グラフィックと関連付けられたキャプションテキストまたは他の説明文が、識別され格納されてもよい。オーサリングシステムは、グラフィックをスキャンして、この説明文を見つけ、データベースに格納してもよい。矢印などの識別用マーカが識別されてもよく、識別マーカによって示される位置（例えば、矢印の先端）の(x, y)座標が保存されてもよい。オーサリングシステムは次に、保存されたOCRテキストをグラフィックに対応する説明文と比較して、グラフィックの説明文に対する一般的なOCRテキストへの相関があるかを決定してもよい。相関がある場合、その相関はデータベースに保存される。一実施形態では、相関はタグとして格納される。別の実施形態では、相関は、インデックスキーまたは他の相関デバイスを使用して保存されてもよい。オーサリングシステムは次に、作成されたデータベースに基づいて、質問および回答を作成してもよい。別の実施形態では、ユーザが、オーサリングシステムによって作成されたデータベースへのアクセスを有するAI会話システムと対話すると、AI会話システムは、識別されたユーザ質問に対する回答を作成してもよい。いずれの場合も、ユーザが、データベース自体またはユーザ質問への応答のどちらかに対するアクセスを有するAI会話システムと対話すると、リンク（グラフィックタグもしくはインデックス）を有する言葉または応答が識別された場合、会話システムが返答の音声ダイアログを生成すると、対応するグラフィックが表示されてもよい。グラフィックが大きい場合、識別（タグ付け）された識別用マークの(x, y)座標が拡大されるか、強調されるか、囲まれるか、または別の方法で示されて、ユーザが関連するマーカを見つけるのを助けてもよい。これはまた、画像分析および領域セグメント化で使用されてもよい。

20

30

### 8. 可変の詳細度を有するAI会話の構成要素

#### 【0075】

心情のキューは、応答の長さを変動させるなど、AI会話システムの態様を変更するのに使用されてもよい。心情はユーザの応答で変化してもよく、会話システムは、ユーザの心情を周期的に測定し、ユーザの更新された心情にしたがって応答の詳細度を変動させてもよい。会話システムは、言葉ベースの方法から、音声波形分析から、または顔の分析を実行するマシンビジョンから、心情を決定してもよい。

40

#### 【0076】

一実施形態では、会話システムは、高、中、および低の詳細度に対する設定を有してもよい。「高」の設定は、システムに「ください」応答を作成させてもよく、「中」の設定は、簡潔な直接の応答をさせてもよく、「低」の設定はビープ音またはアイコンの点滅のみで、ユーザの声が聞き取られており、アクションが実行されたことを知らせてもよい。より高度な実装は、動的な可変性またはスライドするスケールを使用して、応答の長さの変動におけるより高い粒度を作成してもよい。

#### 【0077】

50

一実施形態では、会話システムは以下の段階を行ってもよい。

【0078】

1. ユーザの発話を分析して、言葉の使用に基づいてユーザの感情状態を決定する。

【0079】

2. ユーザの感情状態を採点し、データベースに記録する。

【0080】

3. ユーザの発話を分析して、分毎に話された言葉、音量、または韻律に基づいてユーザの感情状態を決定する。他の波形または映像分析も採点に使用されてもよい。

【0081】

4. 分毎に話された言葉を、言葉の頻度と関連させた既知の感情状態のデータベースと比較する。 10

【0082】

5. ユーザの感情状態を採点し、同じデータベースに記録する。

【0083】

6. ユーザの発話を分析して、会話システムとの対話によるユーザの満足レベルを示す直接コメントを探す。例えば、会話システムに対して「静かにして」と言うユーザは、会話システムからより低い詳細度を望んでいるサインと取られる場合がある。

【0084】

7. これを採点し、満足感の変化を記録する。

【0085】

8. データベースに格納された全てのユーザデータを集約して、複合スコアを達成する。システムはこのスコアを周期的に更新するであろう。 20

【0086】

9. スコアがユーザの満足感または特定の心の状態を示す場合、会話システムの応答の詳細度レベルに変化は生じないことがあるが、他の場合は生じることがある。スコアが上下した場合、会話システムはその詳細度を適宜調節するであろう。

9. 顔分析の調節を伴うアバターインタフェースの構成要素

【0087】

この構成要素を使用して、会話システムは、マシンビジョンが導出したユーザの顔の状態の分析に基づいて、アバターの顔の状態を変動させることができる。会話システムは、自然言語理解ソフトウェアを使用して、顔のジェスチャー認識に対するサポートを追加し、自分のアバターにおいて人間のような顔のジェスチャーを生成することによって、アバターの会話能力を向上させてもよい。 30

【0088】

人間は、話をする際、話し手の感情状態を伝達する顔のジェスチャーおよび微表情で話し言葉を補足する。例えば、目の瞬きは話し手が話し終わったことを合図することができる。更に、聞き手の顔の表情は、話し手の顔の表情を反映し、それと対話する場合が多い。会話システムは、ユーザの映像をキャプチャし、ユーザの顔の表情を分析し、顔の表情を認識し、それらの顔の表情に対して反応することによって、そのアバターの性能を改善してもよい。 40

【0089】

一実施形態では、会話システムは、ビデオカメラおよびマシンビジョンソフトウェアを使用して、ユーザが話す際の顔のパターンをキャプチャし分析してもよい。システムは次に、これらのパターンを既知の感情状態および表情にマッピングしてもよい。これらのパターンの分析は、発声によって表現されたユーザの意図についてのシステムのモデルを高め、それが次いで、人間のようなパターンでアバターの顔をアニメーション化するのに使用されてもよい。

【0090】

別の実施形態では、会話システムは、当該システムが話し終わったときにアバターの目の瞬きを生成して、会話システムが話し終わったことをユーザに合図してもよい。 50

## 【0091】

別の実施形態では、会話システムは、マシンビジョンを使用して、姿勢、体の位置、および手振りを含むユーザのボディランゲージを分析してもよい。会話システムは次に、これらのパターンを既知の感情状態および表情にマッピングしてもよい。このボディランゲージ情報は、発声によって表現された意図についてのシステムの理解を拡張し、システムによって使用して、アバターの顔および体を人間のようなパターンでアニメーション化することができる。ボディランゲージ情報は、人間のような応答を、アバターの発話応答ならびに顔および体のアニメーションと一致させるのに使用されてもよい。一実施形態では、アバター応答は、それらの「パーソナリティ」を構成する設計因子によって更に調整されるであろう。

10

## 【0092】

会話システムは、アバターの顔の視覚的状态に基づいて、会話システムの出力に句読点を付ける方法を使用してもよい。

## 【0093】

一実施形態では、システムは以下の段階を実行してもよい。

## 【0094】

1. ビデオカメラおよびマシンビジョンソフトウェアを使用して、ユーザの目の瞬きを検出し分析する。

## 【0095】

2. 記録された瞬きを、ユーザの発話の文字起こしを含むタイムラインと相互参照する。

20

## 【0096】

3. ソフトウェアを使用して、瞬きを発声概念およびパターンの自然言語理解分析と比較して、句読点が生じる可能性が高い場所を決定することによって、句読点に関連しない目の瞬きを拒絶する。

## 【0097】

4. ユーザの目の瞬きをキャプチャし、これらの目の瞬きを分析し分類して、どの瞬きが(ドライアイの瞬きなどとは対照的に)句読点として意図されたかを決定する。

## 【0098】

5. 句読点としての瞬きの候補を可能性が高い句読点の音声分析と比較して、偽陽性を見つける。

30

## 【0099】

6. 文の終わりに相関する目の瞬きを、会話システムが最後のユーザ発話に応答するのをトリガする、会話のハンドオーバー点として分類する。

## 【0100】

別の実施形態では、会話システムは、以下の段階で、顔認識と、会話型システムのアバターにおける顔のアニメーションの形態である、ユーザの発話におけるアバターの反応とに焦点を絞ってもよい。

## 【0101】

1. ビデオカメラおよびマシンビジョンソフトウェアを使用して、ユーザの顔のジェスチャーを検出し分析する。

40

## 【0102】

2. 記録された顔のジェスチャーを、ユーザの発話の文字起こしを含むタイムラインと相互参照する。

## 【0103】

3. 認識された顔のジェスチャーを、既知の感情状態のデータベースと比較する。

## 【0104】

4. この導出された感情状態情報を使用して、信頼、幸福、悲しみなどの変数に基づいて、採点システムに追加投入する。

## 【0105】

5. 導出されたスコアを、人間の会話で一般的に使用される既知の顔の応答のデータベ

50

スと比較する。

【0106】

これらの顔の応答を使用して、会話システムのアバターの顔を形作るプリセットされたアニメーションをトリガすることができる。

【0107】

別の実施形態は、顔のアニメーションに手振りを補足するかまたはそれと置き換えてもよい。

【0108】

別の実施形態は、顔のアニメーションに体の位置付けを補足するかまたはそれと置き換えてもよい。

10. 処理されたユーザ音声ファイルをキャッシュする構成要素

【0109】

この構成要素を使用して、会話システムは、対話型会話システムの処理されたユーザ音声ファイルを、キャッシュミスに関する低解像度音声ファイルを使用してキャッシュする。高品質の発話を生成するには、低品質の発話を生成するよりもはるかに多くの計算を行うが、場合によっては、例えばリアルタイムシステムでは、高品質の発話を生成するのに十分な時間がない。会話システムは、どの応答であるべきかを決定すると、そのオーディオ応答のバージョンが存在するかを確かめるため、ローカルレポジトリをチェックする。エントリがキャッシュに存在しない場合、システムは、低品質バージョンを生成し、それを再生し、低品質プロセスをキューアップして、ローカルレポジトリに、またはサーバ上に格納される高品質バージョンを生成する。

【0110】

会話システムは、特定のユーザに同じ応答を繰り返し用いて応答してもよいので、より高品質の音声ファイルを生成することで見かけの品質が改善される。しかし、他のユーザは一般の応答の異なるセットを有することがある。会話システムは、集中的な計算要件を有する可能性が高いので、低品質タスクは、システムの応答性に悪影響を及ぼすことなくその計算を行うのに、多くの時間を有するはずである。

【0111】

一実施形態では、システムは、話される言葉が言葉のSHAハッシュをファイル名に有するWAVファイルとして格納される、レポジトリとしてのファイルシステムを使用してもよい。システムは、応答をハッシュし、次にファイル名としてハッシュを有するファイルが存在するかを確かめるのに目を向けるであろう。

【0112】

更なる改良は、レポジトリが一杯になったときに、どのエントリを破棄するかを決める最適化された手法であろう。システムが、ファイルが使用されるたびに修正日を変更した場合、ファイルを日にちで仕分け、最も古いものを選ぶことで、「最も長い間使用されていない」エントリを識別することによってそれを破棄できる、単純な手法が提供される。あるいは、選択は使用頻度または使用予測に基づいてもよい。

11. 回帰試験の構成要素

【0113】

この構成要素を使用して、会話AIエンジンは、試験入力の可変性を扱うことができる。ソフトウェアは、回帰試験と呼ばれるプロセスで試験されてもよい。回帰試験では、既知の入力がデータベースに格納され、不具合を探すのに、試験システムの一部としてのソフトウェアプログラムによって逐次的に使用される。会話システムは、ユーザの書込みまたは音声入力を取り、この入力を処理してユーザの意図を決定し、応答を送達する。会話システムを有効に試験するため、ユーザの意図のできるだけ多くの可能な音声またはテキスト表現を試験することが望ましい。

【0114】

自然言語理解ソフトウェアエンジンを使用して、言葉および概念を変動させることによって試験入力に可変性を追加することで、会話システムをより堅牢に試験することができる

10

20

30

40

50

可変入力を作成される。一実施形態では、試験システムは、回帰試験からの入力を取り、新しい試験の形態で可変性および深度を追加する置換を生成する。これらの置換は予め規定された概念定義から生成されるであろう。一例として、「アレックスは釣りが好き？」というユーザ発話は、次の ~ m a l e \_ n a m e s の概念（アレックス、ボブ、チャーリー、デイク、アーニー、フランク）および ~ s p o r t s の概念（釣り、凧揚げ、ハイキング）によって置換することができる。結果として得られる置換は、よって、「ボブはハイキングが好き？」であってもよい。置換は回帰試験データベースに追加されて、深度および可変性が試験に追加されるであろう。いくつかの実施形態では、システムは、音声入力の韻律、アクセント、または方言を変動させることがある。

例示のハードウェア

【0115】

上述のシステムは、1または複数のコンピューティングシステム上に実装されてもよい。

【0116】

一実施形態によれば、本明細書に記載する技術は、ファームウェア、メモリ、他の記憶装置、または組み合わせにおけるプログラム命令に従った技術を実行するようにプログラムされた、1または複数の一般化されたコンピューティングシステムによって実装される。デスクトップコンピュータシステム、ポータブルコンピュータシステム、携帯用デバイス、ネットワーキングデバイス、あるいは技術を実装するハードワイヤードおよび/またはプログラム論理を組み込んだ他の任意のデバイスなど、専用コンピューティングデバイスが使用されてもよい。

【0117】

例えば、図1は、本発明の一実施形態がその上に実装されてもよい、コンピュータシステム100を示すブロック図である。コンピュータシステム100は、情報を通信するバス102または他の通信メカニズムと、情報を処理する、バス102と結合されたプロセッサ104とを含む。プロセッサ104は、例えば、汎用マイクロプロセッサであってもよい。

【0118】

コンピュータシステム100はまた、情報およびプロセッサ104によって実行される命令を格納する、バス102に結合された、ランダムアクセスメモリ(RAM)または他の動的記憶デバイスなどのメインメモリ106を含む。メインメモリ106はまた、プロセッサ104によって実行される命令の実行中、一時変数または他の中間情報を格納するのに使用されてもよい。かかる命令は、プロセッサ104にアクセス可能な非一時的記憶媒体に格納されると、コンピュータシステム100を、命令で指定された動作を実行するようにカスタマイズされた専用機械にする。

【0119】

コンピュータシステム100は更に、プロセッサ104のための静的情報および命令を格納する、バス102に結合された、読出し専用メモリ(ROM)108または他の静的記憶デバイスを含む。磁気ディスクまたは光学ディスクなどの記憶デバイス110が提供され、情報および命令を格納するため、バス102に結合される。

【0120】

コンピュータシステム100は、バス102を介して、コンピュータユーザに対して情報を表示する、コンピュータモニタなどのディスプレイ112に結合されてもよい。英数字キーおよび他のキーを含む入力デバイス114は、情報およびコマンド選択をプロセッサ104に通信するため、バス102に結合される。別のタイプのユーザ入力デバイスは、方向情報およびコマンド選択をプロセッサ104に通信するため、ならびにディスプレイ112上のカーソル移動を制御するための、マウス、トラックボール、またはカーソル方向キーなどのカーソル制御116である。この入力デバイスは、一般的に、デバイスが面内の位置を指定するのを可能にする、第1の軸(例えば、x)および第2の軸(例えば、y)の2つの軸で、2つの自由度を有する。

【0121】

10

20

30

40

50

コンピュータシステム 100 は、コンピュータシステムとの組み合わせでコンピュータシステム 100 を専用機械にさせるかまたはそのようにプログラムする、カスタマイズされたハードワイヤード論理、1もしくは複数のASICまたはFPGA、ファームウェアおよび/またはプログラム論理を使用して、本明細書に記載する技術を実装してもよい。一実施形態によれば、本明細書の技術は、プロセッサ 104 が、メインメモリ 106 に収容された1または複数の命令の1もしくは複数のシーケンスを実行するのに応答して、コンピュータシステム 100 によって実行されてもよい。かかる命令は、記憶デバイス 110 などの別の記憶媒体から、メインメモリ 106 に読み込まれてもよい。メインメモリ 106 に収容された命令のシーケンスを実行することによって、プロセッサ 104 に、本明細書に記載するプロセス段階を実行させる。代替実施形態では、ソフトウェア命令の代わりに、またはそれとの組み合わせで、ハードワイヤード回路が使用されてもよい。

10

#### 【0122】

「記憶媒体」という用語は、本明細書で使用するとき、機械を特定の形式で動作させるデータおよび/または命令を格納する、任意の非一時的媒体を指す。かかる記憶媒体は、不揮発性媒体および/または揮発性媒体を備えてもよい。不揮発性媒体は、例えば、記憶デバイス 110 など、光学または磁気ディスクを含む。揮発性媒体は、メインメモリ 106 などの動的メモリを含む。記憶媒体の一般形態は、例えば、フロッピーディスク、フレキシブルディスク、ハードディスク、ソリッドステートドライブ、磁気テープ、または他の任意の磁気データ記憶媒体、CD-ROM、他の任意の光学データ記憶媒体、穴のパターンを有する任意の物理的媒体、RAM、PROM、EPROM、FLASH（登録商標）-EPROM、NVRAM、他の任意のメモリチップもしくはカートリッジを含む。

20

#### 【0123】

記憶媒体は、伝送媒体とは異なるが、それと併せて使用されてもよい。伝送媒体は記憶媒体間での情報の転送に関与する。例えば、伝送媒体は、バス 102 を備えるワイヤを含む、同軸ケーブル、銅ワイヤ、および光ファイバーを含む。伝送媒体はまた、高周波および赤外データ通信の間に生成されるものなど、音波または光波の形態を取ることができる。

#### 【0124】

様々な形態の媒体が、1または複数の命令の1もしくは複数のシーケンスをプロセッサ 104 に搬送して実行するのに関与してもよい。例えば、命令は最初に、リモートコンピュータの磁気ディスクまたはソリッドステートドライブ上で搬送されてもよい。リモートコンピュータは、命令をその動的メモリにロードし、ネットワーク接続を通じて命令を送信することができる。コンピュータシステム 100 に対してローカルのモデムまたはネットワークインタフェースが、データを受信することができる。バス 102 は、データをメインメモリ 106 に搬送し、そこからプロセッサ 104 は、命令を検索し実行する。メインメモリ 106 によって受信された命令は、任意に、プロセッサ 104 による実行の前または後のどちらかで、記憶デバイス 110 に格納されてもよい。

30

#### 【0125】

コンピュータシステム 100 はまた、バス 102 に結合された通信インタフェース 118 を含む。通信インタフェース 118 は、ローカルネットワーク 122 に接続されたネットワークリンク 120 に結合する、二方向データ通信を提供する。例えば、通信インタフェース 118 は、ケーブルモデム、衛星モデム、または対応するタイプの電話線に対するデータ通信接続を提供するモデムであってもよい。無線リンクも実装されてもよい。任意のかかる実装では、通信インタフェース 118 は、様々なタイプの情報を表すデジタルデータストリームを搬送する、電気、電磁、または光学信号を送信および受信する。

40

#### 【0126】

ネットワークリンク 120 は、一般的に、1または複数のネットワークを通して、他のデータデバイスにデータ通信を提供する。例えば、ネットワークリンク 120 は、ローカルネットワーク 122 を通して、ホストコンピュータ 124 に対する、またはインターネットサービスプロバイダ (ISP) 126 によって操作されるデータ機器に対する、接続を提供してもよい。ISP 126 は次いで、現在は一般に「インターネット」128と呼

50

ばれる、ワールドワイドパケットデータ通信ネットワークを通して、データ通信サービスを提供する。ローカルネットワーク 1 2 2 およびインターネット 1 2 8 は両方とも、デジタルデータストリームを搬送する、電気、電磁、または光学信号を使用する。様々なネットワークを通る信号、およびネットワークリンク 1 2 0 上の、通信インタフェース 1 1 8 を通る信号は、デジタルデータをコンピュータシステム 1 0 0 に、またそこから搬送するものであり、伝送媒体の例示の形態である。

【 0 1 2 7 】

コンピュータシステム 1 0 0 は、ネットワーク、ネットワークリンク 1 2 0、および通信インタフェース 1 1 8 を通して、メッセージを送信し、プログラムコードを含むデータを受信することができる。インターネットの例では、サーバ 1 3 0 は、インターネット 1 2 8、ISP 1 2 6、ローカルネットワーク 1 2 2、および通信インタフェース 1 1 8 を通して、アプリケーションプログラムに対して要求されたコードを伝送してもよい。受信したコードは、受信されたままプロセッサ 1 0 4 によって実行され、および/または記憶デバイス 1 1 0 に、もしくは後の実行のために他の不揮発性記憶装置に格納されてもよい。

10

【 0 1 2 8 】

本明細書に記載するプロセスの動作は、本明細書における別段の指示、または文脈による明確な矛盾がない限り、任意の好適な順序で実行することができる。本明細書に記載するプロセス（あるいはその変形および/または組み合わせ）は、実行可能命令を用いて構成された、1 または複数のコンピュータシステムの制御下で実行されてもよく、ハードウェアもしくはその組み合わせによって、1 もしくは複数のプロセッサ上で集合的に実行する、コード（例えば、実行可能命令、1 もしくは複数のコンピュータプログラム、または1 もしくは複数のアプリケーション）として実装されてもよい。コードは、例えば、1 または複数のプロセッサによって実行可能な複数の命令を備えるコンピュータプログラムの形態で、コンピュータ可読記憶媒体に格納されてもよい。コンピュータ可読記憶媒体は非一時的であってもよい。

20

【 0 1 2 9 】

図 2 ~ 図 1 4 は、実施形態の他の態様について記載している。

【 0 1 3 0 】

図 1 5 は、自動対話システムによって使用可能なデータを構造化システムが生成する、システムを示している。図示されるように、作者は、自然言語の作者入力を、概念の記録、ルールセット、パターンの説明、およびユーザ入力（音声、テキスト、データなどであってもよい）を取り込み、応答出力を提供する、自動会話システムなどの自動対話システムの動作を形成する実行可能コードなどの、データ構造を構築することができる、オーサリングシステムなどの構造化システムに提供することができる。構造化システムは、作者の側で詳細なプログラミングを要することなく構築することができる、自動対話システムの実装を可能にする。特定の例では、自然言語処理を、特定のドメインまたは使用のための質問・回答システムを構築している、オーサリングシステムで使用することができる。オーサリングシステムは、質問を尋ね、作者の応答を得ること、それらの応答を自然言語の作者入力として処理することなどによって、作者と対話し、それらを構造化形式データとして格納し、その構造化形式データから、自動対話型システムを形成するであろう概念、パターン、ルール、実行可能コード、またはルーチンなどを計算してもよい。ユーザは次に、その自動対話型システムを使用して対話することができる。

30

40

【 0 1 3 1 】

本開示の実施形態は、以下の節を考慮して記載することができる。

1 . 実行可能命令で構成された 1 または複数のコンピュータシステムの制御下で、1 または複数の知識ドメインのセットから選択された知識ドメインを選択するように、オーサリングユーザに促す段階と、  
オーサリングユーザの、選択された知識ドメインの選択を受信する段階と、  
オーサリングユーザからオーサリングユーザ入力を受信する段階と、

50

オーサリングユーザ入力を、オーサリングされた自動統合システムによって使用可能な、構造化された形態の複数のテキスト出力へと変換する段階と  
を備える、会話システムを生成するコンピュータ実装方法。

2. オーサリングユーザ入力を複数のテキスト出力へと変換する段階であって、第1のテキスト出力が第1の認識システムの第1の出力であり、第2のテキスト出力が第2の認識システムの第2の出力である、段階と、

選択された知識ドメインのドメイン仕様に基づいて、ドメイン固有の計画を作成する段階と、

ランタイム仕様を取得する段階であって、ランタイム仕様が、選択された知識ドメインに対する、ドメイン固有の計画に基づいて構成された計画タスクフローを有する、段階と、  
計画タスクフローを実行する段階と、

ユーザ入力から入力値を生成する段階と、

入力値に基づいて会話システムを改善する段階と、

会話システムの説明をコンピュータ可読メモリに格納する段階と

を更に備える、節1の方法。

3. オーサリングユーザからの入力が音声入力またはテキスト入力を有する、節1または2の方法。

4. 第1の認識システムおよび第2の認識システムが、自動発話認識システムまたは画像認識システムの1もしくは複数である、節1～3のいずれかの方法。

5. 推論モジュールを使用し、会話システムと対話する対話中のユーザからの入力に基づいて、計画タスクフローを動的に訂正する段階を更に備える、節1～4のいずれかの方法。

6. オーサリングユーザから、1または複数の第1の選択肢の第1のセットの中からの、選択された選択肢の第1のオーサリングユーザ選択を取得する段階と、

第1のオーサリングユーザ選択に基づいて、計画タスクフローを調節する段階と、

データマイニングモジュールを使用して、格納されたドメイン知識レポジトリを作成する段階と

を更に備える、節1～5のいずれかの方法。

7. データマイニングモジュールが、構造化テキスト、非構造化テキスト、および/またはグラフィックのうち1もしくは複数を使用し、データマイニングモジュールの計算が会話システムの出力を変更する、節6の方法。

8. クラウドソーシングによる知識ソースランキングシステムを用いて、自動ドメイン知識ソースモジュールを使用してドメイン固有の計画が生成され、方法が更に、

複数の知識ソースそれぞれに対する採点値を導出する段階と、

自動ドメイン知識ソースモジュールを使用して、採点値に基づいて、複数のソースの中から使用する選択されたソースを動的に決定する段階と、

選択されたソースを会話システムの出力値にマッピングする段階と

を備える、節1～7のいずれかの方法。

9. ユーザ入力に基づいて会話プログラムを動的に改善するシステムであって、

1または複数のプロセッサと、

実行されると1または複数のプロセッサに、

a) ユーザ入力に基づいて意図を形成させ、

b) 意図に基づいて計画を作成させ、計画が、意図と関連付けられた第1の概念オブジェクトを第2の概念オブジェクトへと変換する第1のアクションオブジェクトを有し、第2の概念オブジェクトを意図の目標と関連付けられた第3の概念オブジェクトへと変換する第2のアクションオブジェクトを有し、第1のアクションオブジェクトおよび第2のアクションオブジェクトが複数のアクションオブジェクトから選択され、第1のアクションオブジェクトが第1の第三者デベロッパによって提供され、第2のアクションオブジェクトが第2の第三者デベロッパによって提供され、

c) 計画を実行させ、

10

20

30

40

50

d) 第3の概念オブジェクトと関連付けられた値を出力させる、複数の命令を格納する、非一時的コンピュータ可読媒体とを備える、システム。

10. 第1の概念オブジェクトが第3の第三者デベロッパによって提供され、第2の概念オブジェクトが第4の第三者デベロッパによって提供され、第3の概念オブジェクトが第5の第三者デベロッパによって提供される、節9のシステム。

11. 第1の概念オブジェクトが、第1の概念オブジェクトのインスタンス化を提供する第1のデータを含み、第2の概念オブジェクトが、第2の概念オブジェクトのインスタンス化を提供する第2のデータを含み、第3の概念オブジェクトが、第3の概念オブジェクトのインスタンス化を提供する第3のデータを含む、節9または10のシステム。

10

12. 第1のアクションオブジェクトの入力パラメータがウェブサービスパラメータにマッピングされ、ウェブサービス結果が第1のアクションオブジェクトの出力値にマッピングされる、節9～11のいずれかのシステム。

#### 【0132】

「A、B、およびCの少なくとも1つ」、または「A、BおよびCの少なくとも1つ」の形態の語句など、接続的な文言は、別段の具体的な提示、または文脈による別段の明確な矛盾がない限り、その他の点では、項目、用語などが、AもしくはBもしくはCのいずれか、またはAおよびBおよびCのセットの任意の空でないサブセットのいずれであってもよいことを提示するのに一般に使用されるような文脈で理解される。例えば、3つの部材を有するセットの具体例では、接続的な語句「A、B、およびCの少なくとも1つ」なら

20

#### 【0133】

あらゆる例の使用、または本明細書で提供される例示的な文言（例えば、「など」）は単に、本発明の実施形態をより良好に明らかにしようとするものであり、別段の特許請求がされない限り、本発明の範囲に対する限定を課すものではない。本明細書における文言は、いずれかの特許請求されない要素を本発明の実践に必須であるとして示すものとして解

30

#### 【0134】

上述の明細書では、実装ごとに変動してもよい多数の特定の詳細を参照して、本発明の実施形態について記載してきた。したがって、明細書および図面は、限定的ではなく例示の意味で解釈されるべきである。本発明の範囲の唯一かつ排他的な指標、および本発明の範囲であることが本出願人らによって意図されるものは、本出願に由来する特許請求の範囲のセットの、あらゆる後の補正を含むかかる請求項が由来する特定の形態の、逐語的な等価の範囲である。

#### 【0135】

更なる実施形態を、本開示を読んだ後に当業者が想起することができる。他の実施形態では、上述の発明の組み合わせまたは下位組み合わせを有利に行うことができる。構成要素の例示の配置は例示の目的で示されており、組み合わせ、追加、再配置などが本発明の代替実施形態で想到されることが理解されるべきである。したがって、例示的实施形態に関して本発明について記載してきたが、当業者であれば、多数の修正が可能であることを認識するであろう。

40

#### 【0136】

例えば、本明細書に記載するプロセスは、ハードウェア構成要素、ソフトウェア構成要素、および/またはそれらの任意の組み合わせを使用して実装されてもよい。したがって、明細書および図面は、限定的ではなく例示の意味で解釈されるべきである。しかしながら、それらに対して、特許請求の範囲で説明するような本発明のより広範な趣旨および範囲

50

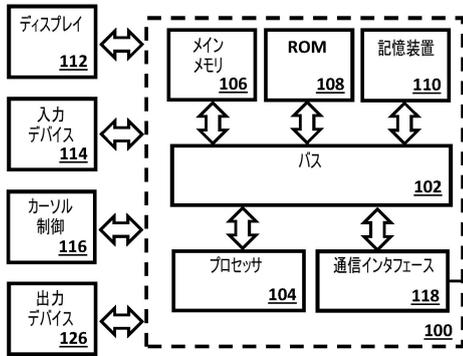
から逸脱することなく、様々な修正および変更が行われてもよいこと、また本発明は、以下の特許請求の範囲の範囲内の全ての修正および等価物を網羅しようとするものであることが、明白となるであろう。

【 0 1 3 7 】

本明細書に列挙する、出版物、特許出願、および特許を含む全ての参照文献は、各参照文献が参照により組み込まれるものと個別かつ具体的に示され、その全体が本明細書で説明された場合と同じ範囲で、参照によりここに組み込まれる。

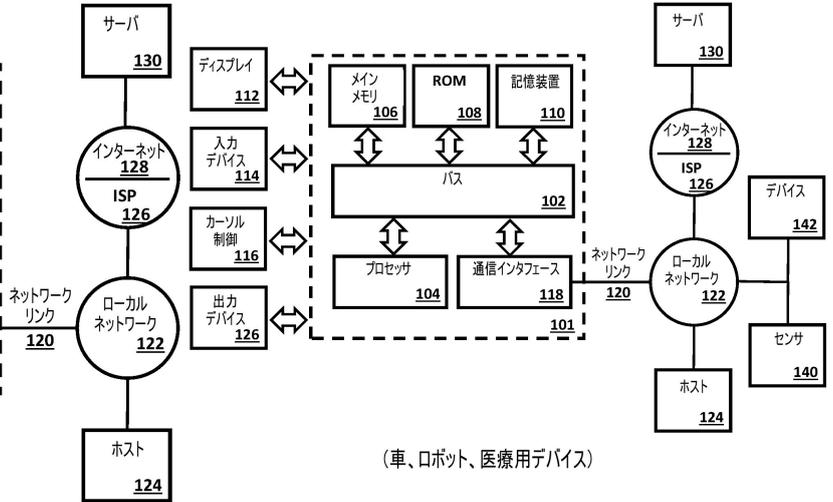
【 図 面 】

【 図 1 】



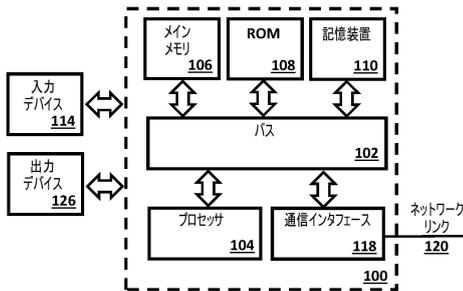
(コンピュータ)

【 図 2 】



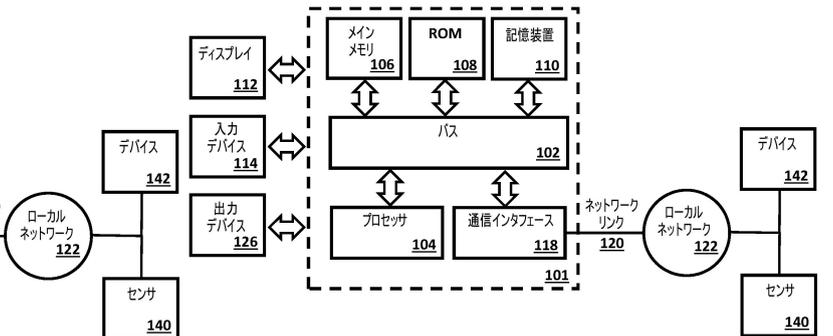
(車、ロボット、医療用デバイス)

【 図 3 】



(電気器具、医療用デバイス、飛行機)

【 図 4 】



(自動販売機)

10

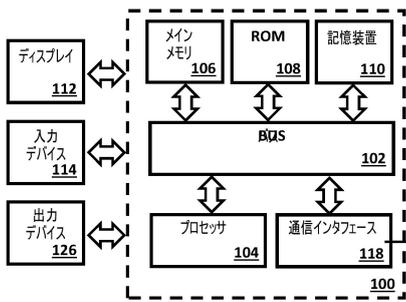
20

30

40

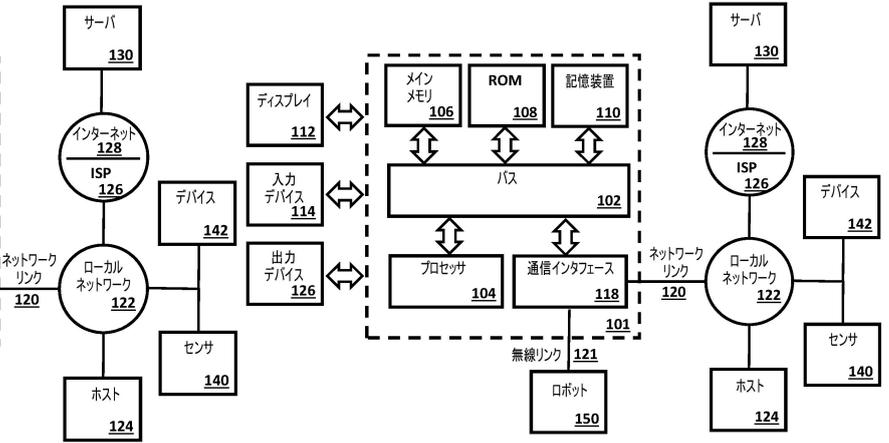
50

【 図 5 】



(インターネット自動販売機)

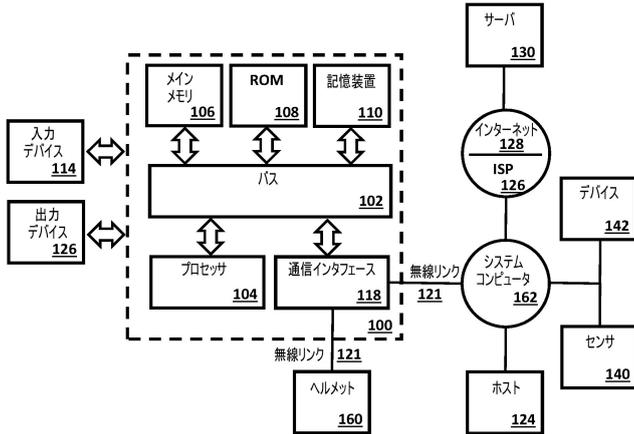
【 図 6 】



(自動販売機+ロボット)

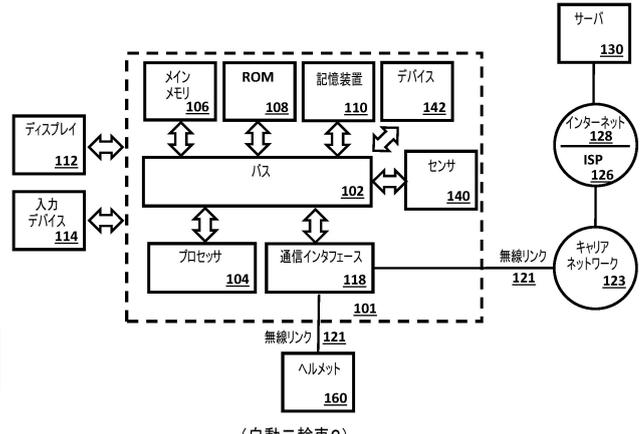
10

【 図 7 】



(自動二輪車1)

【 図 8 】



(自動二輪車2)

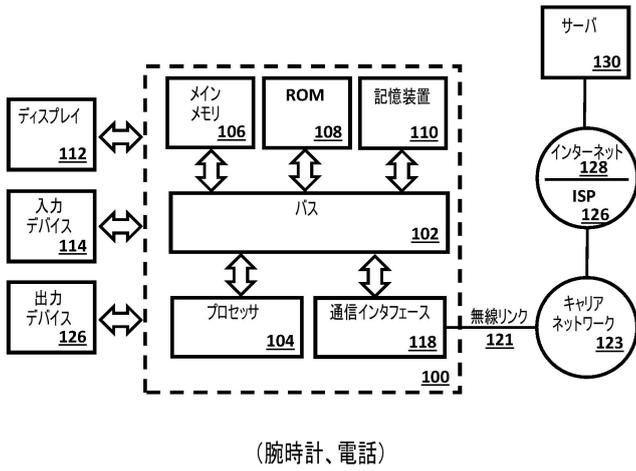
20

30

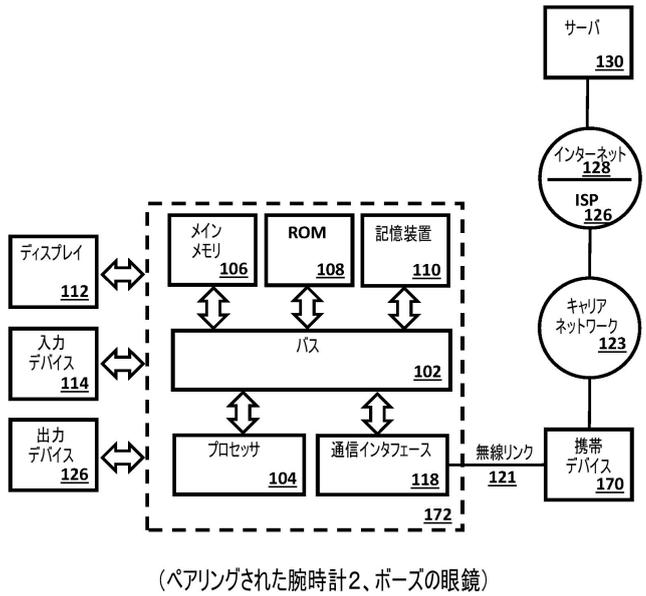
40

50

【 図 9 】

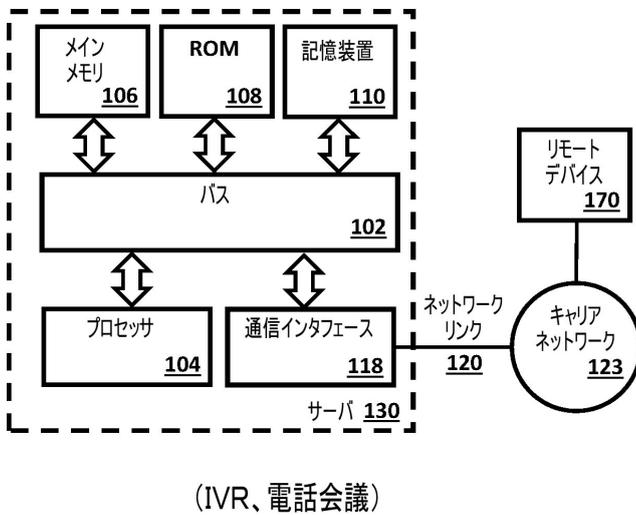


【 図 10 】

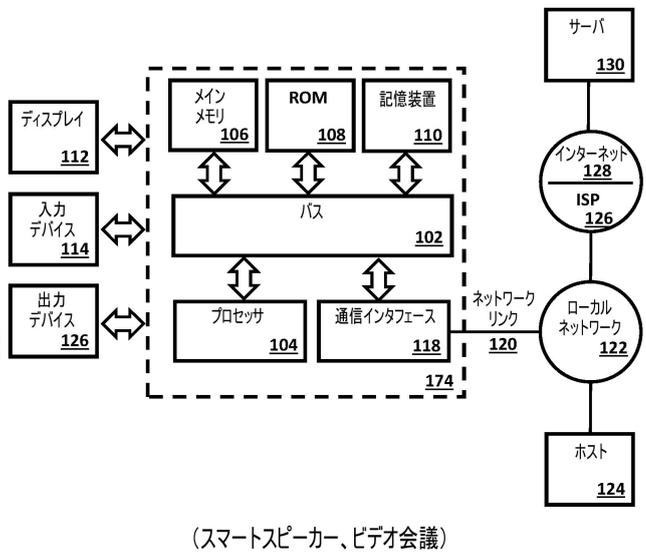


10

【 図 11 】



【 図 12 】



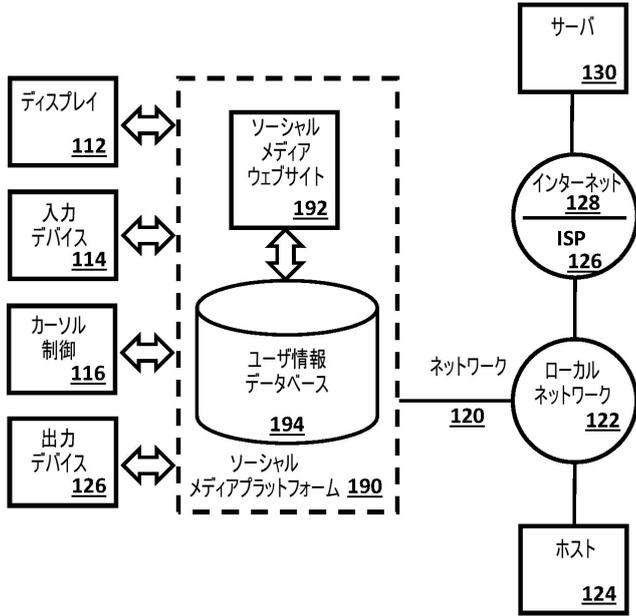
20

30

40

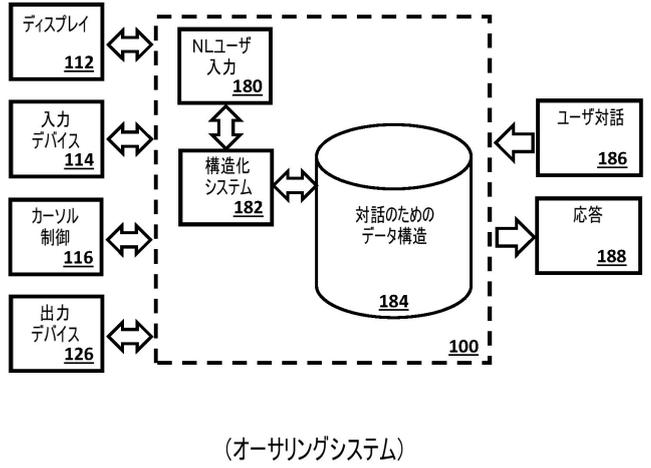
50

【 図 1 3 】



(ソーシャルメディアシステム)

【 図 1 4 】

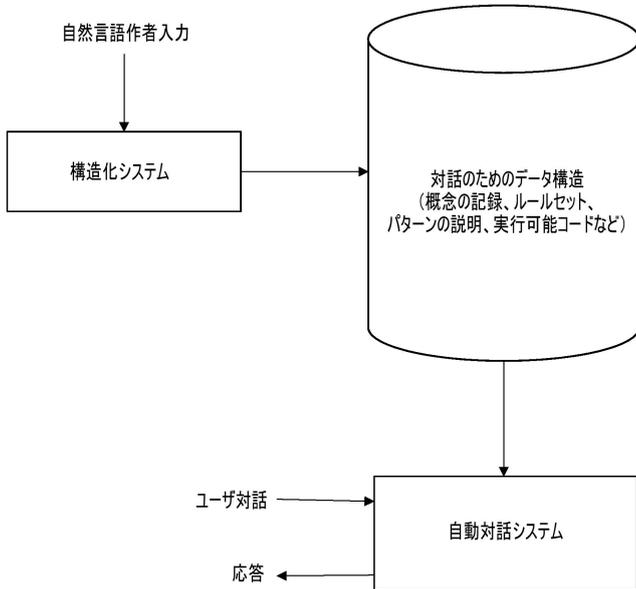


(オーサリングシステム)

10

20

【 図 1 5 】



30

40

50

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 20/40375

**A. CLASSIFICATION OF SUBJECT MATTER**  
 IPC - G10L 13/04; G10L 25/27; G10L 25/48; G11B 20/10 (2020.01)  
 CPC - G06N 5/02; G06N 5/022; G06N 5/04; G06Q 30/02; G06Q 30/0203; G06Q 30/0251; G06Q 30/0256; G06Q 40/04; G10L 13/00; G10L 15/08; G10L 25/27; G10L 25/48; G11B 20/10527

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**  
 Minimum documentation searched (classification system followed by classification symbols)  
 See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
 See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
 See Search History document

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2012/0041903 A1 (BEILBY ET AL.) 16 February 2012 (16.02.2012) Entire document, especially para [0056], [0057], [0058], [0059], [0073], [0074], [0115], [0119], [0120], [0125], [0128], [0149], [0156], [0159], [0179], [0236].	1-8

Further documents are listed in the continuation of Box C.       See patent family annex.

\* Special categories of cited documents:  
 "A" document defining the general state of the art which is not considered to be of particular relevance  
 "D" document cited by the applicant in the international application  
 "E" earlier application or patent but published on or after the international filing date  
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  
 "O" document referring to an oral disclosure, use, exhibition or other means  
 "P" document published prior to the international filing date but later than the priority date claimed  
 "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  
 "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  
 "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art  
 "&" document member of the same patent family

Date of the actual completion of the international search  
28 October 2020 (28.10.2020)

Date of mailing of the international search report  
**20 NOV 2020**

Name and mailing address of the ISA/US  
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents  
P.O. Box 1450, Alexandria, Virginia 22313-1450  
Facsimile No. 571-273-8300

Authorized officer  
Lee Young  
Telephone No. PCT Helpdesk: 571-272-4300

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 20/40375

**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1.

Group I: Claims 1-8, directed to a computer-implemented method for generating a conversation system.

Group II: Claims 9-12, directed to a system for a dynamically improving a conversation program based on user input.

(see extra sheet)

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-8

**Remark on Protest**

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 20/40375

In continuation of Box No. III Observations where unity of invention is lacking (Continuation of Item 3 of first sheet):

The groups of inventions listed above do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical feature of the Group I invention is prompting an authoring user to select a selected knowledge domain from a set of one or more knowledge domains; receiving the authoring user's selection of the selected knowledge domain; receiving authoring user input from the authoring user; and converting the authoring user input into a plurality of text outputs in a structured form, usable by an authored automated integration system, not required by Group II.

10

The special technical feature of the Group II invention is a) form an intent based on a user input; b) create a plan based on the intent, wherein the plan comprises a first action object that transforms a first concept object associated with the intent into a second concept object and comprises a second action object that transforms the second concept object into a third concept object associated with a goal of the intent, wherein the first action object and the second action object are selected from a plurality of action objects, and wherein the first action object is provided by a first third-party developer and the second action object is provided by a second third-party developer; c) execute the plan, and d) output a value associated with the third concept object, not required by Group I.

Groups I-II share the technical features of, in various combination, conversation, and user input.

However, these shared technical features fail to represent a contribution over the prior art of US 2018/0314689 A1 to SRI International (hereinafter "SRI"), which discloses conversation, and user input (para [0089]-[0090], [0113] - "the system is able to let the user specify the subject matter of the conversation, and to have the user direct the flow of the conversation"; "the person is speaking to the system, thus the person's input is user audio input"; "aided by ontologies and domain-specific knowledge").

Thus, the inventions listed as Groups I-II lack unity of invention because they do not share a same or corresponding special technical feature providing a contribution over the prior art.

20

30

40

50

---

フロントページの続き

G,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE  
,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA  
,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,R  
W,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

レード ロード 4 6 0 0

Fターム(参考) 5B175 EA01