

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7263376号
(P7263376)

(45)発行日 令和5年4月24日(2023.4.24)

(24)登録日 令和5年4月14日(2023.4.14)

(51)国際特許分類	F I			
G 0 6 F 40/56 (2020.01)	G 0 6 F	40/56		
G 0 6 F 40/35 (2020.01)	G 0 6 F	40/35		
G 0 6 F 16/90 (2019.01)	G 0 6 F	16/90	1 0 0	
G 0 6 F 3/16 (2006.01)	G 0 6 F	3/16	6 5 0	
	G 0 6 F	3/16	6 2 0	
請求項の数 17 (全33頁) 最終頁に続く				

(21)出願番号	特願2020-546386(P2020-546386)	(73)特許権者	502208397
(86)(22)出願日	平成30年3月5日(2018.3.5)		グーグル エルエルシー
(65)公表番号	特表2021-515938(P2021-515938 A)		Google LLC
(43)公表日	令和3年6月24日(2021.6.24)		アメリカ合衆国 カリフォルニア州 9 4 0 4 3 マウンテン ビュー アンフィシ
(86)国際出願番号	PCT/US2018/020943		アター パークウェイ 1 6 0 0
(87)国際公開番号	WO2019/172878		1 6 0 0 Amphitheatre P
(87)国際公開日	令和1年9月12日(2019.9.12)		arkway 9 4 0 4 3 Mounta
審査請求日	令和2年11月2日(2020.11.2)		in View, CA U.S.A.
		(74)代理人	100108453
			弁理士 村山 靖彦
		(74)代理人	100110364
			弁理士 実広 信哉
		(74)代理人	100133400
			弁理士 阿部 達彦
最終頁に続く			

(54)【発明の名称】 自動化されたアシスタントによる以前の対話コンテキスト間の遷移

(57)【特許請求の範囲】

【請求項1】

1つまたは複数のプロセッサを使用して実行される方法であって、
 ユーザによって操作されるコンピューティングデバイスの1つまたは複数の入力構成要素において、前記ユーザと、前記プロセッサのうちの1つまたは複数によって少なくとも部分的に実装された自動化されたアシスタントとの間の第1の人間対コンピュータの対話中に前記ユーザから自由形式の自然言語入力の1つまたは複数のインスタンスを受信するステップと、
第1の対話コンテキストと関連する、前記ユーザと前記自動化されたアシスタントとの間の前記第1の人間対コンピュータの対話の1つまたは複数の第1のパラメータを記憶するステップと、
 前記コンピューティングデバイスの前記1つまたは複数の入力構成要素において、前記ユーザと前記自動化されたアシスタントとの間の第2の人間対コンピュータの対話中に、前記ユーザから自由形式の自然言語入力の1つまたは複数のさらなるインスタンスを受信するステップと、
第2の対話コンテキストと関連する、前記ユーザと前記自動化されたアシスタントとの間の前記第2の人間対コンピュータの対話の1つまたは複数の第2のパラメータを記憶するステップと、
 前記ユーザと前記自動化されたアシスタントとの間の前記第2の人間対コンピュータの対話において、前記ユーザから要求を受信することに応答して、前記第1の対話コンテキ

10

20

ストと関連する前記記憶された1つまたは複数の第1のパラメータと、前記第2の対話コンテキストと関連する前記記憶された1つまたは複数の第2のパラメータと、を取得し、いくつかの前記コンピューティングデバイスの1つまたは複数の出力構成要素または異なるコンピューティングデバイスに、前記取得したパラメータに基づく出力を生成させる、ステップと、を含み、前記出力は、前記第1および第2の対話コンテキストを含む、利用可能な過去の対話コンテキストの列挙リストを伝達する、

方法。

【請求項2】

前記1つまたは複数の第1のパラメータと前記1つまたは複数の第2のパラメータは、前記ユーザによって操作される前記コンピューティングデバイスに対してローカルである、コンピュータメモリ内に記憶される、

10

請求項1に記載の方法。

【請求項3】

前記1つまたは複数の第1のパラメータと前記1つまたは複数の第2のパラメータは、前記ユーザによって操作される前記コンピューティングデバイスに対してローカルである、コンピュータメモリから取得される、

請求項1に記載の方法。

【請求項4】

前記1つまたは複数の第1のパラメータを記憶するステップが、

前記1つまたは複数の第1のパラメータのための構成ファイルを作成するステップと、

前記構成ファイルを前記コンピュータメモリ内に記憶するステップとを含む、

請求項2に記載の方法。

20

【請求項5】

前記1つまたは複数の第1のパラメータが、前記第1の人間対コンピュータの対話内の前記ユーザの1つもしくは複数の意図、および/または、前記第1の人間対コンピュータの対話内の前記ユーザの前記1つもしくは複数の意図に関連付けられた1つもしくは複数のスロット値を含む、

請求項1に記載の方法。

【請求項6】

前記出力構成要素がディスプレイを含み、利用可能な過去の対話コンテキストの前記列挙リストが選択可能なリンクのリストを含み、前記選択可能なリンクのそれぞれが、前記自動化されたアシスタントを対応する対話コンテキストに戻すために選択可能である、

30

請求項1に記載の方法。

【請求項7】

前記出力構成要素がスピーカーを含み、利用可能な過去の対話コンテキストの前記列挙リストが自然言語スニペット (snippets) のリストを含み、前記自然言語スニペットのそれぞれが前記自動化されたアシスタントを対応する対話コンテキストに戻すために選択可能である、

請求項1に記載の方法。

【請求項8】

前記1つまたは複数の第2のパラメータを記憶するステップが、

前記1つまたは複数の第2のパラメータのための構成ファイルを作成するステップと、

前記構成ファイルを前記コンピュータメモリ内に記憶するステップとを含む、

請求項2に記載の方法。

40

【請求項9】

前記1つまたは複数の第2のパラメータが、前記第2の人間対コンピュータの対話内の前記ユーザの1つもしくは複数の意図、および/または前記第2の人間対コンピュータの対話内の前記ユーザの前記1つもしくは複数の意図に関連付けられた1つもしくは複数のスロット値を含む、

請求項1に記載の方法。

50

【請求項 10】

1つまたは複数のプロセッサと、前記1つまたは複数のプロセッサに動作可能に結合されたメモリとを備えたシステムであって、

前記メモリが、1つまたは複数のプロセッサによって命令が実行されることに応答して、前記1つまたは複数のプロセッサに

ユーザによって操作されるコンピューティングデバイスの1つまたは複数の入力構成要素において、前記ユーザと、前記プロセッサのうちの1つまたは複数によって少なくとも部分的に実装された自動化されたアシスタントとの間の第1の人間対コンピュータの対話中に前記ユーザから自由形式の自然言語入力の1つまたは複数のインスタンスを受信することと、

10

第1の対話コンテキストと関連する、前記ユーザと前記自動化されたアシスタントとの間の前記第1の人間対コンピュータの対話の1つまたは複数の第1のパラメータを記憶することと、

前記コンピューティングデバイスの前記1つまたは複数の入力構成要素において、前記ユーザと前記自動化されたアシスタントとの間の第2の人間対コンピュータの対話中に、前記ユーザから自由形式の自然言語入力の1つまたは複数のさらなるインスタンスを受信することと、

第2の対話コンテキストと関連する、前記ユーザと前記自動化されたアシスタントとの間の前記第2の人間対コンピュータの対話の1つまたは複数の第2のパラメータを記憶することと、

20

前記ユーザと前記自動化されたアシスタントとの間の前記第2の人間対コンピュータの対話に続いて、前記ユーザから要求を受信することに応答して、前記第1の対話コンテキストと関連する前記記憶された1つまたは複数の第1のパラメータと、前記第2の対話コンテキストと関連する前記記憶された1つまたは複数の第2のパラメータと、を取得し、いくつかの前記コンピューティングデバイスの1つまたは複数の出力構成要素または異なるコンピューティングデバイスに、前記取得したパラメータに基づく出力を生成させる、ことと、
を実行させるための命令を記憶し、

前記出力は、前記第1および第2の対話コンテキストを含む、利用可能な過去の対話コンテキストの列挙リストを伝達する、

システム。

30

【請求項 11】

1つまたは複数のプロセッサを使用して実行される方法であって、

ユーザによって操作されるコンピューティングデバイスの1つまたは複数の入力構成要素において、前記ユーザと、前記プロセッサのうちの1つまたは複数によって少なくとも部分的に実装された自動化されたアシスタントとの間の人間対コンピュータの対話セッション中に、前記ユーザから自由形式の自然言語入力の1つまたは複数のインスタンスを受信するステップと、

自由形式の自然言語入力の前記1つまたは複数のインスタンスに基づいて第1の対話コンテキストを生成するステップであって、前記第1の対話コンテキストが、前記ユーザの1つまたは複数の意図と、前記1つまたは複数の意図に関連付けられた1つまたは複数のスロット値とを含む、ステップと、

40

前記第1の対話コンテキストを生成した後、前記入力構成要素のうちの1つまたは複数において、同じ人間対コンピュータの対話セッションまたは前記ユーザと前記自動化されたアシスタントとの間の異なる人間対コンピュータの対話セッション中に、前記ユーザから自由形式の自然言語入力の1つまたは複数の追加インスタンスを受信するステップと、

自由形式の自然言語入力の前記1つまたは複数の追加インスタンスに基づいて第2の対話コンテキストを生成するステップであって、前記第2の対話コンテキストが、前記ユーザの1つまたは複数の追加の意図を含み、前記第2の対話コンテキストが、前記第1の対話コンテキストと意味論的に異なる、ステップと、

前記入力構成要素のうちの1つまたは複数において、前記ユーザから過去の対話コンテキ

50

ストの列挙リストについての要求を受信するステップと、

前記第1の対話コンテキストに基づいて、過去の対話コンテキストの前記列挙リストを伝達する出力を生成するステップであって、前記列挙リストは、前記第1の対話コンテキストの前記意図のうちの少なくとも1つまたは複数と、前記第1の対話コンテキストの前記スロット値のうちの1つまたは複数とを伝達するとともに、前記第2の対話コンテキストの前記意図のうちの少なくとも1つまたは複数と、前記第2の対話コンテキストの前記スロット値のうちの1つまたは複数とを伝達する、生成するステップと、

前記同じコンピューティングデバイスまたは前記ユーザによって操作される異なるコンピューティングデバイスにおいて前記出力を提示させるステップとを含む、

方法。

10

【請求項12】

前記1つまたは複数の第1のパラメータと前記1つまたは複数の第2のパラメータは、前記ユーザによって操作される前記コンピューティングデバイスに対してローカルである、コンピュータメモリ内に記憶される、

請求項10に記載のシステム。

【請求項13】

前記1つまたは複数の第1のパラメータと前記1つまたは複数の第2のパラメータは、前記ユーザによって操作される前記コンピューティングデバイスに対してローカルである、コンピュータメモリから取得される、

請求項10に記載のシステム。

20

【請求項14】

前記出力構成要素がディスプレイを含み、利用可能な過去の対話コンテキストの前記列挙リストが選択可能なリンクのリストを含み、前記選択可能なリンクのそれぞれが、前記自動化されたアシスタントを対応する対話コンテキストに戻すために選択可能である、

請求項10に記載のシステム。

【請求項15】

前記出力構成要素がスピーカーを含み、利用可能な過去の対話コンテキストの前記列挙リストが自然言語スニペット (snippets) のリストを含み、前記自然言語スニペットのそれぞれが前記自動化されたアシスタントを対応する対話コンテキストに戻すために選択可能である、

請求項10に記載のシステム。

30

【請求項16】

利用可能な過去の対話コンテキストの前記列挙リストが選択可能なリンクのリストを含み、前記選択可能なリンクのそれぞれが、前記自動化されたアシスタントを対応する対話コンテキストに戻すために選択可能である、

請求項11に記載の方法。

【請求項17】

利用可能な過去の対話コンテキストの前記列挙リストが自然言語スニペット (snippets) のリストを含み、前記自然言語スニペットのそれぞれが前記自動化されたアシスタントを対応する対話コンテキストに戻すために選択可能である、

請求項11に記載の方法。

40

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、自動化されたアシスタントによる以前の対話コンテキスト間の遷移に関する。

【背景技術】

【0002】

人間は、本明細書で「自動化されたアシスタント」(「チャットボット」、「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」、「パーソナル音声アシスタント」、「会話エージェント」などとも呼ばれる)と呼ばれる対話型ソフトウェア

50

アプリケーションを用いて人間対コンピュータの対話に関与する場合がある。例えば、自動化されたアシスタントと対話するとき、「ユーザ」と呼ばれる場合もある人間は、自由形式の自然言語入力を使用してコマンド、クエリ、および/または要求を提供する場合がある。自由形式の自然言語は、スピーチ認識を使用してテキストに変換された音声発声および/またはタイピングされた自由形式の自然言語入力を含む場合がある。

【0003】

自動化されたアシスタントは、ユーザが情報を取得する、サービスにアクセスする、および/または様々なタスクを実行することを可能にする。例えば、ユーザは、検索を実行する、行き方を知る、場合によっては、サードパーティアプリケーションと対話することができる。自動化されたアシスタントおよび/またはサードパーティアプリケーションによって提供されるアプリケーションプログラミングインターフェース(「API」)を使用してアクセスが容易になり得るこれらのサードパーティアプリケーションの対話は、ライドシェアアプリケーションから車を呼び出す、商品またはサービス(例えば、ピザ)を注文する、予約を行うなどの様々なアクションをユーザが実行することを可能にする。

10

【0004】

自動化されたアシスタントは、例えば、ユーザの意図を予測するために、一部は機械学習および他の人工知能技術も利用して、音声認識および自然言語処理を使用してユーザと会話する。自動化されたアシスタントは、一つには対話コンテキストを理解しているので、自然で直感的な方法でユーザと会話するのが上手い場合がある。対話コンテキストを利用するために、自動化されたアシスタントは、ユーザからの最新の入力、ユーザから来た質問、および/または、自動化されたアシスタントによって提供された応答/質問を保存する場合がある。例えば、ユーザは、「最寄りのコーヒーショップはどこですか」と尋ねる場合があり、それに対して、自動化されたアシスタントは、「東に2ブロックです」と応答する場合がある。ユーザは、次いで、「そこはどれくらい遅くまで開いてますか」と尋ねる場合がある。少なくともなんらかの形の対話コンテキストを保存することによって、自動化されたアシスタントは、「そこ」という代名詞が「コーヒーショップ」を指すことを決定することができる(すなわち、共参照解析)。

20

【0005】

しかしながら、従来の自動化されたアシスタントでは、ユーザが自動化されたアシスタントを別の異なるトピックに関与させるために人間対コンピュータの対話の1つのトピックを放棄する場合、第1のトピックに関連付けられた対話コンテキストは、失われる場合がある。ユーザが、中止した第1のトピックに関する以前の会話をまた始めたい場合、ユーザは、放棄された対話コンテキストが再構築され得るように、最初から会話を再開することを強制される場合がある。関連する会話が再開されるべき場合、ユーザおよびアシスタントが最初からコンテキストを再構築する必要がある場合があるので、対話コンテキストのこの損失は、ユーザの多くの一般的な使用例において技術的な問題を含む。これは、アシスタントが特定のユーザ入力を二度目に受信し、そのような再受信された入力をコンピュータメモリ内に再記憶させる必要があることを伴う場合がある。さらに、そのような再入力の受信は、入力を再処理するためにコンピュータ処理能力を必要とする場合がある。加えて、ユーザが対話コンテキストを再構築するための関連情報を再入力するときにアシスタントがユーザに促すかまたは他の方法でユーザと会話することを可能にするために、デバイス上のスピーカおよび/またはマイクロフォンに電力を供給するために、例えば、デバイス上のバッテリーからの電力が必要とされる場合がある。そのような一般的な使用例の1つの例は、自動化されたアシスタントとサードパーティアプリケーションとの間の対話を伴う。自動化されたアシスタントは、サードパーティアプリケーションと対話することができるが、しばしば、いくつかのサードパーティアプリケーションの全機能を複製することができず、したがって、例えば、人間対コンピュータの対話の複写物において表示されるディープリンクを使用して、ユーザをサードパーティアプリケーションにしばしば転送しなければならない。ユーザがサードパーティアプリケーションから自動化されたアシスタントに戻ると、特に、以前の間対コンピュータの対話からかなりの時間が経過し

30

40

50

た場合、以前の対話コンテキストは、失われる場合がある。

【0006】

これは、ユーザが、他にほとんど何もせず、サードパーティアプリケーションに行き、タスクを達成し、次いで、例えば、ユーザのモバイルデバイス上で実行されるオペレーティングシステムによって提供される戻るボタンを使用して、自動化されたアシスタントにすぐに戻る場合、問題を含まない場合がある。そのようなシナリオにおいて、自動化されたアシスタントは、会話がシームレスに再開され得るように、最近の対話コンテキストを単に保持することができる。しかしながら、ユーザが一般にマルチタスクを行い、無数の注意を乱すことに対処することで、アシスタントへの復帰は、多くの異なる経路を通る場合がある。いくらか類似した状況において、ウェブブラウザのユーザは、複数の研究のスレッド、一連の思考などを保持するために、しばしば複数のタブを使用する。

10

【発明の概要】

【課題を解決するための手段】

【0007】

本明細書では、ユーザと自動化されたアシスタントとの間の人間対コンピュータの対話中に生成された複数の意味論的に異なる対話コンテキストを保存するための技法について説明する。加えて、本明細書では、例えば、特定の対話コンテキストに関連付けられたサードパーティアプリケーションとの対話中に遷移コマンドを提供することによって、自動化されたアシスタントとの以前の会話を再開するために、ユーザが複数の対話コンテキストにおよび/またはそれらの間で遷移することを可能にするための技法について説明する。対話コンテキストが「意味論的に異なる」とみなされるために、2つの異なる対話コンテキストの基礎となる主題が完全に無関係である必要はない。むしろ、2つの対話コンテキストは、例えば、それらが、一致しないおよび/または重複しない意図を有する場合、予め定義されたトピックのライブラリなどからの異なる予め定義されたトピックに関連する場合、などの場合、「意味論的に異なる」場合がある。

20

【0008】

様々な実装形態において、人間対コンピュータの対話中に生成された対話コンテキストは、様々な情報を含み得る。いくつかの実装形態において、対話コンテキストは、ユーザによって提供された自由形式の自然言語入力および/または自動化されたアシスタントによって提供された自然言語出力などの様々な信号に基づいて検出された1つまたは複数のユーザ意図を含み得る。いくつかの実装形態において、対話コンテキストは、人間対コンピュータの対話セッション中に(ユーザおよび/または自動化されたアシスタントによって)言及された既知のエンティティ(例えば、人、場所、物)への1つまたは複数の参照も含み得る。

30

【0009】

いくつかの実装形態において、対話コンテキストは、ユーザの意図を履行するために、すなわち、ユーザの意図に関連付けられた1つまたは複数の「スロット」を満たすことによって使用可能な1つまたは複数のスロット値をさらに含み得る。スロット値は、自動化されたアシスタントによって実行されるべきタスクのためのパラメータとして使用可能な任意の値であり得る。スロット値は、様々なソースから取得され得る。ユーザは、積極的に、および/または、自動化されたアシスタントからの要請に回答して、スロット値を提供し得る。例えば、特定のサードパーティアプリケーションが、要求を履行するために1つまたは複数のスロットを必要とする場合がある。場合によっては、ユーザは、必要なスロットのための値を積極的に提供し得る。また、サードパーティアプリケーションは、ユーザが値を積極的に提供していない任意の必要なスロットのための値を自動化されたアシスタントに要請させ得る。加えて、他のスロット値は、例えば、ユーザによって操作されるコンピューティングデバイスによって提供される位置座標、ユーザの好み、過去の同様な状況下で使用されたスロット値、ユーザおよび/または自動化されたアシスタントによる最近の声明から抽出されたスロット値を使用して、自動的に決定され得る。

40

【0010】

50

対話コンテキストは、様々な方法において保存され得る。いくつかの実装形態において、例えば、対話コンテキストが、保存された複写物内の意図、スロット値、エンティティなどを検出することによって、例えば、オンザフライでおよび/または必要に応じて再構築または再開され得るように、人間対コンピュータの対話の複写物全体が保存され得る。いくつかの実装形態、特に、ユーザがグラフィカルユーザインターフェース(「GUI」)を使用して自動化されたアシスタントと対話する実装形態において、これは、ユーザが複写物を再検討することができ得るという追加の利点を提供し得る。場合によっては、複写物は、中断点、意図、言及されたエンティティなどのデータで注釈付けされ得る。加えて、または代替的に、いくつかの実装形態において、検出された意図、スロット値、言及されたエンティティなどの、対話コンテキストのコア要素のみが、JavaScript(登録商標) Object Notation(「JSON」)または他の同様のフォーマットなどの様々なフォーマットにおいて保存され得る。

10

【0011】

対話コンテキストは、様々な場所に保存され得る。いくつかの実装形態において、対話コンテキストは、自動化されたアシスタントと関わり合うためにユーザによって操作されるコンピューティングデバイスに対してローカルなメモリ内に保持され得る。これは、自動化されたアシスタントとの以前の会話に対してユーザのプライバシーを保護する、および/または、オフラインの間であっても本明細書で説明する技法を使用してユーザが以前の会話に再訪することを可能にするなどの様々な技術的利益を生じさせ得る。加えて、または代替的に、いくつかの実装形態において、対話コンテキストは、例えば、しばしば「クラウドベースサービス」と呼ばれる集合的に動作する1つまたは複数のコンピューティングシステムのメモリ内に、ユーザのコンピューティングデバイスからリモートで保持され得る。様々な実装形態において、対話コンテキストは、様々な時間量の間、および/または様々なイベントにわたって保存され得る。例えば、いくつかの実装形態において、対話コンテキストは、複数の論理的に異なる対話セッションにわたって保持され得る。加えて、または代替的には、いくつかの実装形態において、対話コンテキストは、自動的に選択され得るおよび/またはユーザによって手動で設定され得る数日、数週、または他の時間などの様々な時間量の間保持され得る。いくつかの実装形態において、対話コンテキストは、スタックとしてメモリ内に維持され得、古い対話コンテキストにアクセスするために最近の対話コンテキストがスタックの一番上から「ポップ可能」である。

20

30

【0012】

保存された対話コンテキストは、自動化されたアシスタントとの以前の会話を再開するためにユーザによって使用され得る。このようにして、ユーザは、以前のスレッドおよび/または思考プロセスを失うことなく、かつ、アシスタントに以前に提供された情報を再入力する必要なく、直感的な方法で、自動化されたアシスタントとの以前の会話を通して遷移または「トグル」することができる。タッチスクリーン、静電容量式タッチパッド、または他の視覚的入力/出力(「I/O」)メカニズムが利用可能ないくつかの実装形態において、ユーザは、ユーザが以前の対話コンテキストに遷移するために操作することができる、「戻るボタン」などの選択可能要素を提示され得る。この選択可能要素は、例えば、メモリ内で以前の対話コンテキストに関連付けられている特定のサードパーティアプリケーションと同時に提供され得る。加えて、または代替的には、ユーザが自動化されたアシスタントとの人間対コンピュータの対話において音声で関与するシナリオにおいて、ユーザは、自動化されたアシスタントを以前の対話コンテキストに戻すために、「Hey assistant, let's go back to the prior conversation (ねえアシスタント、以前の会話に戻りましょう)」などの1つまたは複数のコマンドを発声することができ得る。

40

【0013】

いくつかの実装形態において、自動化されたアシスタントは、ユーザが、ユーザが再訪したい対話コンテキストを選択することができるように、利用可能な過去の対話コンテキストの列挙されたリストをユーザに提示するように構成され得る。例えば、ユーザによって操作されるデバイスが表示能力を有する実装形態において、自動化されたアシスタント

50

は、例えば、リスト項目をタッチするか、またはリスト項目に関連してフレーズを話すことによってユーザが選択することができる過去の対話コンテキストに関連付けられた過去の会話の視覚的リストを提示し得る。いくつかの実装形態において、利用可能な以前の対話コンテキストの列挙されたリストは、「Hey @Assistant, what were we talking about before?(ねえ@アシスタント、私達は前に何について話しましたか)」などの自由形式の自然言語入力として提供されるユーザコマンドに応答して提示され得る。

【0014】

いくつかの実装形態において、自動化されたアシスタントは、例えば、ユーザが以前の会話について必要な情報を迅速に与えられ得るように、以前の対話コンテキストの要約を生成するように構成され得る。例えば、いくつかの実装形態において、自動化されたアシスタントは、対話コンテキストに基づいて自然言語出力を生成し得る。自然言語出力は、1つまたは複数の意図、スロット値、言及されたエンティティなどを識別するなど、対話コンテキストの生成につながる会話の様々な側面を伝達し得る。例えば、自動化されたアシスタントは、ユーザが話し合っていたことをユーザに思い出させる自然言語出力、ならびに、すでに満たされている任意のスロット値(例えば、ピザのトッピング)を生成し得る。

【0015】

例示的な例として、ユーザが、サイズ7の女性用ドレスシューズを買うために、自動化されたアシスタントと以前に関わり合っていたとする。さらに、ユーザが、夕食の予約をするための友人からのテキストメッセージによって買い物を中断され、ユーザが、夕食の予約をするためにサードパーティアプリケーションを開くために自動化されたアシスタントと関わり合った(すなわち、主題を変更する)とする。ユーザが靴の買い物を続けるために自動化されたアシスタントに戻るとき、最後の対話コンテキストは、夕食の予約に関連している。靴の買い物の会話を最初から再構築するのではなく、ユーザは、ユーザが友人によって中断されたときにやめた靴の買い物の会話を再開するために、本明細書で説明する技法を用い得る。具体的には、ユーザは、例えば、戻るボタンを押すことによって、または1つもしくは複数の音声コマンドを発行することによって、ユーザが靴を買い物している間に生成された以前の対話コンテキストに遷移するために、1つまたは複数のコマンドを発行することができ得る。これは、以前の対話コンテキストと、サードパーティの靴の買い物アプリケーションなどの、そのコンテキストに関連する特定のサードパーティアプリケーションとの間の関連付けをコンピュータメモリ内に形成することによって達成され得る。靴の買い物アプリケーションに戻ると、ユーザは、以前の対話コンテキスト、すなわち、サイズ7の女性用ドレスシューズを買うことにおけるアシスタントに戻るために遷移コマンドを発行することができ得る。これは、靴の買い物アプリケーションと同時に表示または他の方法で提供される戻るボタンを押すことを含み得る。遷移コマンドは、靴の買い物アプリケーションと靴の買い物コンテキストとの間の記憶された関連付けがアシスタントにおいて靴の買い物の会話を再開するために使用されることを引き起こし得る。同様に、ユーザがその後夕食の予約の会話に戻りたい場合、ユーザは、サードパーティの夕食の予約アプリケーションに戻ると、夕食の予約コンテキストにおけるアシスタントに戻るために遷移コマンドを発行し得る。これは、夕食の予約アプリケーションと同時に表示または他の方法で提供される戻るボタンを押すことを含み得る。

【0016】

いくつかの実装形態において、以前の対話コンテキストを再開するための音声コマンドは、ユーザによって発声された言語に基づいて、以前の対話コンテキストに一致させられ得る。上記の例を続けると、ユーザは、「ねえアシスタント、靴の買い物に戻りましょう」のようなことを言う場合がある。自動化されたアシスタントは、「靴の買い物」に最も類似する以前の対話コンテキストを識別し得、適切な意図/スロット値が入力された状態でその会話を再開し得る。

【0017】

いくつかの実装形態において、例えば、対話コンテキストにおける変化に関連して、1つまたは複数の中断点が会話の複写物に挿入され得る。例えば、ユーザが靴の買い物を話

10

20

30

40

50

し合い始めたが、次いで主題を天候に変更した場合、対話コンテキストごとに1つ、2つの中断点が複写物に挿入され得る。いくつかのそのような実装形態において、ユーザが以前の対話コンテキスト間で(例えば、戻るボタンを押すことによって)遷移すると、選択された対話コンテキストに関連付けられた複写物の一部は、例えば、その対話コンテキストに関連付けられた中断点を使用することによって、画面上にロードされ得る。このようにして、ユーザは、ユーザが再訪したいトピックに関連する以前の会話の全体を見ることができ

【0018】

様々な実装形態において、方法は、以下の、ユーザによって操作されるコンピューティングデバイスの1つまたは複数の入力構成要素において、ユーザとプロセッサのうちの1つまたは複数によって少なくとも部分的に実装された自動化されたアシスタントとの間の人間対コンピュータの対話セッション中に、ユーザから自由形式の自然言語入力の1つまたは複数のインスタンスを受信するステップと、自由形式の自然言語の1つまたは複数のインスタンスに基づいて第1の対話コンテキストを生成するステップであって、第1の対話コンテキストが、ユーザの1つまたは複数の意図と、1つまたは複数の意図に関連付けられた1つまたは複数のスロット値とを含む、ステップと、第1の対話コンテキストを生成した後、入力構成要素のうちの1つまたは複数において、同じ人間対コンピュータの対話セッションまたはユーザと自動化されたアシスタントとの間の異なる人間対コンピュータの対話セッション中にユーザから自由形式の自然言語入力の1つまたは複数の追加インスタンスを受信するステップと、自由形式の自然言語入力の1つまたは複数の追加インスタンスに基づいて第2の対話コンテキストを生成するステップであって、第2の対話コンテキストがユーザの1つまたは複数の追加の意図を含み、第2の対話コンテキストが第1の対話コンテキストと意味論的に異なる、ステップと、入力構成要素のうちの1つまたは複数において、自動化されたアシスタントを第1の対話コンテキストに戻すためのユーザからの遷移コマンドを受信するステップと、第1の対話コンテキストに基づいて、第1の対話コンテキストの意図のうちの少なくとも1つまたは複数と第1の対話コンテキストのスロット値のうちの1つまたは複数とを伝達する自然言語出力を生成するステップと、同じコンピューティングデバイスまたはユーザによって操作される異なるコンピューティングデバイスにおいて自然言語出力を提示させるステップとを含み得る。

【0019】

様々な実装形態において、遷移コマンドは、ユーザによって発声またはタイピングされた自由形式の自然言語入力を含み得る。様々な実装形態において、遷移コマンドは、選択可能なユーザインターフェース要素のユーザによる選択を含み得る。様々な実装形態において、選択可能なユーザインターフェース要素は、戻るボタンを含み得る。様々な実装形態において、戻るボタンは、ユーザによって操作されるコンピューティングデバイスのタッチスクリーン上に表示され得る。様々な実装形態において、ユーザによって操作されるコンピューティングデバイスは、独立型の対話型スピーカであり得、戻るボタンは、独立型の対話型スピーカの静電容量式タッチ表面上に表示される。

【0020】

様々な実装形態において、自動化されたアシスタントは、グラフィカルユーザインターフェース(「GUI」)をユーザのコンピューティングデバイス上にレンダリングさせ得、GUIは、人間対コンピュータの対話中にユーザと自動化されたアシスタントとの間で交換されるメッセージの複写物を提示する。様々な実装形態において、方法は、第1の対話コンテキストの意図の1つまたは複数に回答して、ユーザのコンピューティングデバイス上で別個のソフトウェアアプリケーションを起動させるステップをさらに含み得、別個のソフトウェアアプリケーションを起動させるステップは、自動化されたアシスタントによって提示されるGUIの代わりに、別個のソフトウェアアプリケーションに関連付けられた別のGUIをユーザに提示させるステップを含み、遷移コマンドは、別のGUIの選択可能な要素を介して受信される。様々な実装形態において、メッセージの複写物は、第1の対話コンテキストに関連付けられた第1の中断点と、第2の対話コンテキストに関連付けられた第2の中

10

20

30

40

50

断点とを含み得、遷移コマンドは、第1の中断点に関連付けられた複写物の一部を、自動化されたアシスタントによって提示されるGUI内にレンダリングさせる。

【0021】

様々な実装形態において、方法は、入力構成要素のうちの一つまたは複数において、自動化されたアシスタントが保存された対話コンテキストのリストを提供するためのリスト対話コンテキストコマンドをユーザから受信するステップと、リスト対話コンテキストコマンドに回答して、少なくとも第1および第2の対話コンテキストを識別するステップと、同じコンピューティングデバイスまたはユーザによって操作される異なるコンピューティングデバイスにおいて、少なくとも第1および第2の対話コンテキストを示す情報を提示させるステップとをさらに含む得る。様々な実装形態において、遷移コマンドは、第1の対話コンテキストに対応する少なくとも第1および第2の対話コンテキストを示す情報の要素の、ユーザによる選択を含む得る。

10

【0022】

別の態様において、方法は、ユーザによって操作されるコンピューティングデバイスの一つまたは複数の入力構成要素において、ユーザとプロセッサのうちの一つまたは複数によって少なくとも部分的に実装された自動化されたアシスタントとの間の第1の人間対コンピュータの対話中にユーザから自由形式の自然言語入力の一つまたは複数のインスタンスを受信するステップと、ユーザと自動化されたアシスタントとの間の第1の人間対コンピュータの対話の一つまたは複数の記憶された第1のパラメータを、ユーザと自動化されたアシスタントとの間の第1の人間対コンピュータの対話からユーザによってアクセス可能な第1のサードパーティアプリケーションに関連付けるステップと、コンピューティングデバイスの一つまたは複数の入力構成要素において、ユーザと自動化されたアシスタントとの間の第2の人間対コンピュータの対話中にユーザから自由形式の自然言語入力の一つまたは複数のさらなるインスタンスを受信するステップと、ユーザと自動化されたアシスタントとの間の第2の人間対コンピュータの対話に続いて、第1のサードパーティアプリケーションから自動化されたアシスタントに遷移するためのユーザからの遷移コマンドの受信に回答して、第1のサードパーティアプリケーションに関連付けられた記憶された一つまたは複数の第1のパラメータを取得するステップと、ユーザと自動化されたアシスタントとの間の第1の人間対コンピュータの対話を再開するために、取得された第1のパラメータを使用するステップとを含む得る。

20

30

【0023】

様々な実装形態において、方法は、ユーザと自動化されたアシスタントとの間の第1の人間対コンピュータの対話の一つまたは複数の第1のパラメータをコンピュータメモリ内に記憶させるステップをさらに含む得る。様々な実装形態において、記憶された一つまたは複数の第1のパラメータを取得するステップは、コンピュータメモリから一つまたは複数の第1のパラメータを取得するステップを含む得る。様々な実装形態において、コンピュータメモリは、リモートサーバ内に配置される。

【0024】

様々な実装形態において、一つまたは複数の第1のパラメータを記憶するステップは、一つまたは複数の第1のパラメータのための構成ファイルを作成するステップと、構成ファイルをコンピュータメモリ内に記憶するステップとを含む得る。様々な実装形態において、一つまたは複数の第1のパラメータは、第1の人間対コンピュータの対話コンテキストを示す。様々な実装形態において、一つまたは複数の第1のパラメータは、第1の人間対コンピュータの対話におけるユーザの一つもしくは複数の意図、および/または、第1の人間対コンピュータの対話におけるユーザの一つもしくは複数の意図に関連付けられた一つもしくは複数のスロット値を含む得る。

40

【0025】

様々な実装形態において、第1の人間対コンピュータの対話の一つまたは複数の記憶された第1のパラメータを第1のサードパーティアプリケーションに関連付けるステップは、第1のサードパーティアプリケーションに関連付けられた記憶された一つまたは複数の第1の

50

パラメータを取得するために、第1のサードパーティアプリケーションから自動化されたアシスタントに遷移するための遷移コマンドの受信に回答して、1つまたは複数のプロセッサによって呼び出し可能な第1のトークンを作成するステップを含み得る。

【0026】

様々な実装形態において、第1のトークンは、記憶された1つまたは複数の第1のパラメータと第1のサードパーティアプリケーションとの間のリンクを含み得る。様々な実装形態において、方法は、ユーザと自動化されたアシスタントとの間の第2の人間対コンピュータの対話の1つまたは複数の第2のパラメータを、ユーザと自動化されたアシスタントとの間の第2の人間対コンピュータの対話からユーザによってアクセス可能な第2のサードパーティアプリケーションに関連付けるステップをさらに含み得る。様々な実装形態において、第2のサードパーティアプリケーションから自動化されたアシスタントへの遷移のためのユーザからの遷移コマンドの受信に回答して、方法は、第2のサードパーティアプリケーションに関連付けられた記憶された1つまたは複数の第2のパラメータを取得するステップと、ユーザと自動化されたアシスタントとの間の第2の人間対コンピュータの対話を再開するために、取得された第2のパラメータを使用するステップとを含み得る。

10

【0027】

様々な実装形態において、第2の人間対コンピュータの対話の1つまたは複数の記憶された第2のパラメータを第2のサードパーティアプリケーションに関連付けるステップは、第2のサードパーティアプリケーションに関連付けられた記憶された1つまたは複数の第2のパラメータを取得するために、1つまたは複数のプロセッサによって、第2のサードパーティアプリケーションから自動化されたアシスタントに遷移するための遷移コマンドの受信に回答して、呼び出し可能な第2のトークンを作成するステップを含み得る。様々な実装形態において、第2のトークンは、記憶された1つまたは複数の第2のパラメータと第2のサードパーティアプリケーションとの間のリンクを含み得る。

20

【0028】

加えて、いくつかの実装形態は、1つまたは複数のコンピューティングデバイスの1つまたは複数のプロセッサを含み、1つまたは複数のプロセッサは、関連するメモリ内に記憶された命令を実行するように動作可能であり、命令は、前述の方法のいずれかの実行を引き起こすように構成される。いくつかの実装形態は、前述の方法のいずれかを実行するために1つまたは複数のプロセッサによって実行可能なコンピュータ命令を記憶する1つまたは複数の非一時的なコンピュータ可読記憶媒体も含む。

30

【0029】

本明細書でより詳細に説明する前述の概念および追加の概念のすべての組合せは、本明細書で開示される主題の一部であると考えられることが理解されるべきである。例えば、本開示の最後に現れる特許請求された主題のすべての組合せは、本明細書で開示される主題の一部であると考えられる。

【図面の簡単な説明】

【0030】

【図1】本明細書で開示される実装形態が実装され得る例示的な環境のブロック図である。

【図2A】様々な実装形態による、ユーザと、別のユーザと、ユーザのコンピューティングデバイスを使用する自動化されたアシスタントとの間の例示的な対話を示す図である。

40

【図2B】様々な実装形態による、ユーザと、別のユーザと、ユーザのコンピューティングデバイスを使用する自動化されたアシスタントとの間の例示的な対話を示す図である。

【図2C】様々な実装形態による、ユーザと、別のユーザと、ユーザのコンピューティングデバイスを使用する自動化されたアシスタントとの間の例示的な対話を示す図である。

【図2D】様々な実装形態による、ユーザと、別のユーザと、ユーザのコンピューティングデバイスを使用する自動化されたアシスタントとの間の例示的な対話を示す図である。

【図3A】様々な実装形態による、ユーザと独立型の対話型スピーカを使用する自動化されたアシスタントとの間の例示的な対話を示す図である。

【図3B】様々な実装形態による、ユーザと独立型の対話型スピーカを使用する自動化さ

50

れたアシスタントとの間の例示的な対話を示す図である。

【図3C】様々な実装形態による、ユーザと独立型の対話型スピーカを使用する自動化されたアシスタントとの間の例示的な対話を示す図である。

【図3D】様々な実装形態による、ユーザと独立型の対話型スピーカを使用する自動化されたアシスタントとの間の例示的な対話を示す図である。

【図4】本明細書で開示される実装形態による例示的な方法を示すフローチャートである。

【図5】コンピューティングデバイスの例示的なアーキテクチャを示す図である。

【図6】例示的なコンピューティングデバイスのブロック図である。

【発明を実施するための形態】

【0031】

ここで図1に進むと、本明細書で開示される技法が実装され得る例示的な環境が示されている。例示的な環境は、複数のクライアントコンピューティングデバイス106₁~106_Nを含む。各クライアントデバイス106は、自動化されたアシスタントクライアント118のそれぞれのインスタンスを実行し得る。自然言語理解エンジン135などの1つまたは複数のクラウドベースの自動化されたアシスタント構成要素119は、110において全体として示す1つまたは複数のローカルおよび/またはワイドエリアネットワーク(例えば、インターネット)を介してクライアントデバイス106₁~106_Nに通信可能に結合される1つまたは複数のコンピューティングシステム(集合的に「クラウド」コンピューティングシステムと呼ばれる)上に実装され得る。

【0032】

いくつかの実装形態において、複数のクライアントコンピューティングデバイス106₁~106_N(本明細書では単に「クライアントデバイス」とも呼ばれる)は、本明細書で説明する技法の実行を容易にするために、様々な方法で互いに関連付けられ得る。例えば、いくつかの実装形態において、複数のクライアントコンピューティングデバイス106₁~106_Nは、それらが1人または複数のユーザ(例えば、個人、家族、組織の従業員、他の予め定義されたグループなど)によって操作されるクライアントデバイス106の協調エコシステムのメンバーであることによって互いに関連付けられ得る。

【0033】

いくつかの実装形態において、自動化されたアシスタントクライアント118のインスタンスは、1つまたは複数のクラウドベースの自動化されたアシスタント構成要素119との相互作用によって、ユーザの観点から、ユーザが人間対コンピュータの対話に参与し得る自動化されたアシスタント120の論理インスタンスであるように見えるものを形成し得る。そのような自動化されたアシスタント120の2つのインスタンスが図1に示されている。破線で囲まれた第1の自動化されたアシスタント120Aは、第1のクライアントデバイス106₁を操作する第1のユーザ(図示せず)にサービスを提供し、自動化されたアシスタントクライアント118₁と、1つまたは複数のクラウドベースの自動化されたアシスタント構成要素119とを含む。一点二鎖線で囲まれた第2の自動化されたアシスタント120Bは、別のクライアントデバイス106_Nを操作する第2のユーザ(図示せず)にサービスを提供し、自動化されたアシスタントクライアント118_Nと、1つまたは複数のクラウドベースの自動化されたアシスタント構成要素119とを含む。

【0034】

したがって、クライアントデバイス106上で実行される自動化されたアシスタントクライアント118に参与する各ユーザは、実際には、自動化されたアシスタント120のユーザ自身の論理インスタンスに参与し得ることが理解されるべきである。簡潔さおよび単純さのために、特定のユーザに「サービスを提供する」として本明細書で使用される「自動化されたアシスタント」という用語は、ユーザによって操作されるクライアントデバイス106上で実行される自動化されたアシスタントクライアント118と、(複数の自動化されたアシスタントクライアント118間で共有され得る)1つまたは複数のクラウドベースの自動化されたアシスタント構成要素119との組合せを指す。いくつかの実装形態において、自動化されたアシスタント120は、自動化されたアシスタント120のその特定のインスタンス

10

20

30

40

50

によってユーザが実際に「サービスを提供」されたかどうかに関係なく、任意のユーザからの要求に応答し得ることも理解されるべきである。

【0035】

クライアントデバイス106₁~106_Nは、例えば、デスクトップコンピューティングデバイス、ラップトップコンピューティングデバイス、タブレットコンピューティングデバイス、携帯電話コンピューティングデバイス、ユーザのビークルのコンピューティングデバイス(例えば、車載通信システム、車載エンターテイメントシステム、車載ナビゲーションシステム)、独立型の対話型スピーカ、スマートテレビなどのスマート家電、および/または、コンピューティングデバイスを含むユーザのウェアラブル装置(例えば、コンピューティングデバイスを有するユーザの腕時計、コンピューティングデバイスを有するユーザの眼鏡、仮想または拡張現実コンピューティングデバイス)のうちの1つまたは複数を含み得る。追加のおよび/または代替のクライアントコンピューティングデバイスが提供され得る。

10

【0036】

クライアントコンピューティングデバイス106₁~106_Nの各々は、複数のメッセージ交換クライアント107₁~107_Nのうちの対応する1つなどの様々な異なるアプリケーションを動作させ得る。メッセージ交換クライアント107₁~107_Nは、様々な形態であり得、形態は、クライアントコンピューティングデバイス106₁~106_Nにわたって異なり得、および/または、複数の形態が、クライアントコンピューティングデバイス106₁~106_Nのうちの単一のものにおいて動作され得る。いくつかの実装形態において、メッセージ交換クライアント107₁~107_Nのうちの1つまたは複数は、ショートメッセージングサービス(「SMS」)および/またはマルチメディアメッセージングサービス(「MMS」)クライアント、オンラインチャットクライアント(例えば、インスタントメッセージングサービス、インターネットリレーチャット、または「IRC」など)、ソーシャルネットワークに関連付けられたメッセージングアプリケーション、自動化されたアシスタント120との会話専用のパーソナルアシスタントメッセージングサービスなどの形態であり得る。いくつかの実装形態において、メッセージ交換クライアント107₁~107_Nのうちの1つまたは複数は、クライアントコンピューティングデバイス106のウェブブラウザ(図示せず)または他のアプリケーションによってレンダリングされたウェブページまたは他のリソースを介して実装され得る。

20

【0037】

本明細書でより詳細に説明するように、自動化されたアシスタント120は、1つまたは複数のクライアントデバイス106₁~106_Nのユーザインターフェース入力および出力デバイスを介して1人または複数のユーザとの人間対コンピュータの対話セッションに参与する。いくつかの実装形態において、自動化されたアシスタント120は、クライアントデバイス106₁~106_Nのうちの1つの1つまたは複数のユーザインターフェース入力デバイスを介してユーザによって提供されるユーザインターフェース入力に応答して、ユーザとの人間対コンピュータの対話セッションに参与し得る。それらの実装形態のうちのいくつかにおいて、ユーザインターフェース入力は、自動化されたアシスタント120に明示的に向けられる。例えば、メッセージ交換クライアント107₁~107_Nのうちの1つは、自動化されたアシスタント120との会話専用のパーソナルアシスタントメッセージングサービスであり得、そのパーソナルアシスタントメッセージングサービスを介して提供されるユーザインターフェース入力は、自動化されたアシスタント120に自動的に提供され得る。また、例えば、ユーザインターフェース入力は、自動化されたアシスタント120が呼び出されるべきであることを示す特定のユーザインターフェース入力に基づいて、メッセージ交換クライアント107₁~107_Nのうちの1つまたは複数において自動化されたアシスタント120に明示的に向けられ得る。例えば、特定のユーザインターフェース入力は、1つまたは複数のタイピングされた文字(例えば、(@AutomatedAssistant)、ハードウェアボタンおよび/もしくは仮想ボタンとのユーザ対話(例えば、タップ、ロングタップ)、口頭コマンド(例えば、「ねえ、自動化されたアシスタント」)、ならびに/または他の特定のユーザインターフェース入力であり得る。

30

40

50

【0038】

いくつかの実装形態において、自動化されたアシスタント120は、ユーザインターフェース入力に自動化されたアシスタント120に明示的に向けられていない場合でも、ユーザインターフェース入力に回答して人間対コンピュータの対話セッションに参与し得る。例えば、自動化されたアシスタント120は、ユーザインターフェース入力の内容を調べ、ユーザインターフェース入力内に特定の用語が存在することに回答して、および/または他の合図に基づいて、対話セッションに参与し得る。多くの実装形態において、ユーザは、コマンド、検索などを発声することができ、自動化されたアシスタント120は、発声をテキストに変換するためにスピーチ認識を利用し、それに応じて、例えば、検索結果、一般的な情報を提供することによって、および/または、1つもしくは複数の回答アクションをとる(例えば、メディアを再生する、ゲームを起動する、食べ物を注文するなど)ことによって、テキストに回答し得る。いくつかの実装形態において、自動化されたアシスタント120は、加えて、または代替的に、発声をテキストに変換することなく、発声に回答することができる。例えば、自動化されたアシスタント120は、音声入力を、埋め込み、(音声入力内に存在するエンティティを示す)エンティティ表現、および/または他の「非テキスト」表現に変換し、そのような非テキスト表現において動作することができる。したがって、音声入力から変換されたテキストに基づいて動作するものとして本明細書で説明する実装形態は、加えて、および/または代替的に、音声入力において直接、および/または音声入力の他の非テキスト表現において動作し得る。

10

【0039】

クライアントコンピューティングデバイス106₁~106_N、および、クラウドベースの自動化されたアシスタント構成要素119を動作させるコンピューティングデバイスの各々は、データおよびソフトウェアアプリケーションの記憶のための1つまたは複数のメモリと、データにアクセスし、アプリケーションを実行するための1つまたは複数のプロセッサと、ネットワークを介する通信を容易にする他の構成要素とを含み得る。クライアントコンピューティングデバイス106₁~106_Nのうちの1つもしくは複数によって、および/または自動化されたアシスタント120によって実行される動作は、複数のコンピュータシステムにわたって分散され得る。自動化されたアシスタント120は、例えば、ネットワークを介して互いに結合された1つまたは複数の場所における1つまたは複数のコンピュータ上で実行されるコンピュータプログラムとして実装され得る。

20

30

【0040】

上述のように、様々な実装形態において、クライアントコンピューティングデバイス106₁~106_Nの各々は、自動化されたアシスタントクライアント118を動作させ得る。様々な実施形態において、各自動化されたアシスタントクライアント118は、対応するスピーチ取り込み/テキストからスピーチへ(「TTS」)/スピーチからテキストへ(「STT」)モジュール114を含み得る。他の実装形態において、スピーチ取り込み/TTS/STTモジュール114の1つまたは複数の態様は、自動化されたアシスタント118とは別に実装され得る。様々な実装形態において、スピーチ取り込み/TTS/STTモジュール114は、音声クエリに基づいてスピーチ認識出力を生成し得る。

【0041】

各スピーチ取り込み/TTS/STTモジュール114は、1つまたは複数の機能、すなわち、例えばマイクロフォンを介して、ユーザのスピーチを取り込むこと、その取り込んだ音声をテキストに(および/または他の表現または埋め込みに)変換すること、および/または、テキストをスピーチに変換することを実行するように構成され得る。例えば、いくつかの実装形態において、クライアントデバイス106は、コンピューティングリソース(例えば、プロセッササイクル、メモリ、バッテリーなど)に関して比較的制約される場合があるので、各クライアントデバイス106にローカルであるスピーチ取り込み/TTS/STTモジュール114は、有限数の異なる口頭のフレーズ、特に、自動化されたアシスタント120を呼び出す特定のフレーズをテキストに(または、より低次元の埋め込みなどの他の形態に)変換するように構成され得る。他のスピーチ入力は、クラウドベースのTTSモジュール116および/ま

40

50

たはクラウドベースのSTTモジュール117を含み得るクラウドベースの自動化されたアシスタント構成要素119に送信され得る。

【0042】

クラウドベースのSTTモジュール117は、スピーチ取り込み/TTS/STTモジュール114によって取り込まれた音声データを(次いで、自然言語プロセッサ122に提供され得る)テキストに変換するために、クラウドの実質的に無制限のリソースを活用するように構成され得る。クラウドベースのTTSモジュールモジュール116は、テキストデータ(例えば、自動化されたアシスタント120によって定式化された自然言語応答)をコンピュータ生成スピーチ出力に変換するために、クラウドの実質的に無制限のリソースを利用するように構成され得る。いくつかの実装形態において、TTSモジュール116は、例えば1つまたは複数のスピーカを使用して直接出力されるように、コンピュータ生成スピーチ出力をクライアントデバイス106に提供し得る。他の実装形態において、自動化されたアシスタント120によって生成されたテキストデータ(例えば、自然言語応答)は、スピーチ取り込み/TTS/STTモジュール114に提供され得、スピーチ取り込み/TTS/STTモジュール114は、次いで、テキストデータを、ローカルに出力されるコンピュータ生成スピーチに変換し得る。

10

【0043】

自動化されたアシスタント120(および、特に、クラウドベースの自動化されたアシスタント構成要素119)は、自然言語理解エンジン135と、前述のTTSモジュール116と、前述のSTTモジュール117と、以下でさらに詳細に説明する他の構成要素とを含み得る。いくつかの実装形態において、自動化されたアシスタント120のエンジンおよび/またはモジュールのうちの1つまたは複数は、省略され得、結合され得、および/または、自動化されたアシスタント120とは別の構成要素において実装され得る。いくつかの実装形態において、プライバシーを保護するために、自然言語理解エンジン135、スピーチ取り込み/TTS/STTモジュール114などの、自動化されたアシスタント120の構成要素のうちの1つまたは複数は、少なくとも部分的にクライアントデバイス106において(例えば、クラウドを除外して)実装され得る。

20

【0044】

いくつかの実装形態において、自動化されたアシスタント120は、自動化されたアシスタント120との人間対コンピュータの対話セッション中にクライアントデバイス106₁~106_Nのうちの1つのユーザによって生成された様々な入力に回答して回答コンテンツを生成する。自動化されたアシスタント120は、対話セッションの一部としてユーザに提示するための回答コンテンツを(例えば、ユーザのクライアントデバイスから分離されている場合、1つまたは複数のネットワークを介して)提供し得る。例えば、自動化されたアシスタント120は、クライアントデバイス106₁~106_Nのうちの1つを介して提供される自由形式の自然言語入力に回答して回答コンテンツを生成し得る。本明細書で使用される自由形式の自然言語入力は、ユーザによって定式化され、ユーザによる選択のために提示されるオプションのグループに制約されない入力である。

30

【0045】

本明細書で使用される「対話セッション」は、ユーザと自動化されたアシスタント120(および、場合によっては、他の人間の参加者)との間の1つまたは複数のメッセージの論理的に自己充足的な交換を含み得る。自動化されたアシスタント120は、セッション間の時間経過、セッション間のユーザコンテキスト(例えば、場所、予定された会議の前/最中/後など)の変化、ユーザと、ユーザと自動化されたアシスタントとの間の対話以外のクライアントデバイスとの間の1つまたは複数の介在する対話(例えば、ユーザがしばらくの間アプリケーションを切り替え、ユーザが後にそこから離れて独立型の音声起動製品に戻る)の検出、セッション間のクライアントデバイスのロック/スリープ、自動化されたアシスタント120の1つまたは複数のインスタンスとのインターフェースとして使用されるクライアントデバイスの変更、ユーザによって用いられる入力/出力(「I/O」)モダリティの変更などの、様々な信号に基づいて、ユーザとの複数の対話セッション間を区別し得る。

40

【0046】

50

自然言語理解エンジン135の自然言語プロセッサ122は、クライアントデバイス106₁~106_Nを介してユーザによって生成された自然言語入力を処理し、自動化されたアシスタント120の1つまたは複数の他の構成要素によって使用するための注釈付きの出力(例えば、テキスト形式における)を生成し得る。例えば、自然言語プロセッサ122は、クライアントデバイス106₁の1つまたは複数のユーザインターフェース入力デバイスを介してユーザによって生成された自然言語の自由形式の入力を処理し得る。生成された注釈付きの出力は、自然言語入力の1つまたは複数の注釈と、オプションで自然言語入力の用語のうちの1つまたは複数(例えば、すべて)とを含む。

【0047】

いくつかの実装形態において、自然言語プロセッサ122は、自然言語入力内の様々なタイプの文法情報を識別し、注釈を付けるように構成される。例えば、自然言語プロセッサ122は、個々の単語を形態素に分割し得、および/または、形態素に、例えばそれらのクラスで注釈を付け得る形態学的エンジンを含み得る。自然言語プロセッサ122は、用語にそれらの文法的役割で注釈を付けるように構成された品詞タガーも含み得る。例えば、品詞タガーは、「名詞」、「動詞」、「形容詞」、「代名詞」などのその品詞で各用語にタグ付けし得る。また、例えば、いくつかの実装形態において、自然言語プロセッサ122は、加えて、および/または代替的に、自然言語入力内の用語間の構文関係を決定するように構成された依存関係パーサ(図示せず)を含み得る。例えば、依存関係パーサは、どの用語が他の用語を修飾するのか、文の主語および動詞など(例えば、解析木)を決定し得、そのような依存関係の注釈を付け得る。

【0048】

いくつかの実装形態において、自然言語プロセッサ122は、加えて、および/または代替的に、人(例えば、文学の登場人物、有名人、著名人などを含む)、組織、場所(実在および架空)などへの参照などの、1つまたは複数のセグメント内のエンティティ参照に注釈を付けるように構成されたエンティティタガー(図示せず)を含み得る。いくつかの実装形態において、エンティティに関するデータは、ナレッジグラフ(図示せず)などの1つまたは複数のデータベース内に記憶され得る。いくつかの実装形態において、ナレッジグラフは、既知のエンティティ(および、場合によっては、エンティティ属性)を表すノード、ならびに、ノードを接続してエンティティ間の関係を表すエッジを含み得る。例えば、「バナナ」ノードは、「果物」ノードに(例えば、子供として)接続され得、「果物」ノードは、「農産物」および/または「食品」ノードに(例えば、子供として)接続され得る。別の例として、「Hypothetical Cafe(仮想のカフェ)」と呼ばれるレストランは、その住所、提供される食品のタイプ、営業時間、連絡先情報などの属性も含むノードによって表され得る。いくつかの実装形態において、「Hypothetical Cafe」ノードは、「レストラン」ノード、「ビジネス」ノード、レストランが位置する都市および/または州を表すノードなどの1つまたは複数の他のノードに、(例えば、子供と親の関係を表す)エッジによって接続され得る。

【0049】

自然言語プロセッサ122のエンティティタガーは、高レベルの粒度(例えば、人などのエンティティクラスへのすべての参照の識別を可能にする)および/または低レベルの粒度(例えば、特定の人などの特定のエンティティへのすべての参照の識別を可能にする)においてエンティティへの参照に注釈を付け得る。エンティティタガーは、特定のエンティティを解決するために自然言語入力の内容に依存し得、および/または、特定のエンティティを解決するためにナレッジグラフもしくは他のエンティティデータベースとオプションで通信し得る。

【0050】

いくつかの実装形態において、自然言語プロセッサ122は、追加で、および/または代替的に、1つまたは複数のコンテキストキューに基づいて同じエンティティへの参照をグループ化または「クラスタ化」するように構成された共参照リゾルバ(図示せず)を含み得る。例えば、共参照リゾルバは、「I liked Hypothetical Cafe last time we ate there.(

私は、最後にそこで食べた仮想カフェが好きでした。)という自然言語入力内の「there(そこで)」という用語を「Hypothetical Cafe(仮想カフェ)」に解決するために利用され得る。

【0051】

いくつかの実装形態において、自然言語プロセッサ122の1つまたは複数の構成要素は、自然言語プロセッサ122の1つまたは複数の他の構成要素からの注釈に依存し得る。例えば、いくつかの実装形態において、名前付きエンティタガーは、特定のエンティティへのすべての言及に注釈を付ける際に、共参照リゾルバおよび/または依存関係パーサからの注釈に依存し得る。また、例えば、いくつかの実装形態において、共参照リゾルバは、同じエンティティへの参照をクラスタ化する際に、依存関係パーサからの注釈に依存し得る。いくつかの実装形態において、特定の自然言語入力を処理する際に、自然言語プロセッサ122の1つまたは複数の構成要素は、1つまたは複数の注釈を決定するために、特定の自然言語入力以外の関連する以前の入力および/または他の関連するデータを使用し得る。

10

【0052】

自然言語理解エンジン135は、自然言語プロセッサ122の注釈付き出力に基づいて、自動化されたアシスタント120との人間対コンピュータの対話セッションに参与するユーザの意図を決定するように構成された意図マッチャ136も含み得る。図1において自然言語プロセッサ122とは別に示されているが、他の実装形態において、意図マッチャ136は、自然言語プロセッサ122の(または、より一般的には、自然言語プロセッサ122を含むパイプラインの)不可欠な部分であり得る。いくつかの実装形態において、自然言語プロセッサ122および意図マッチャ136は、前述の「自然言語理解」エンジン135を集合的に形成し得る。

20

【0053】

意図マッチャ136は、ユーザの意図を決定するために様々な技法を使用し得る。いくつかの実装形態において、意図マッチャ136は、例えば、文法と応答アクション(または、より一般的には、意図)との間の複数のマッピングを含む1つまたは複数のデータベース137にアクセスし得る。加えて、または代替的には、いくつかの実装形態において、1つまたは複数のデータベース137は、ユーザの入力に基づいてユーザの意図を示す出力を生成するように訓練された1つまたは複数の機械学習モデルを記憶し得る。

【0054】

文法は、例えば、ユーザの最も一般的な意図を表すために、選択、(例えば、手動で)定式化、および/または経時的に学習され得る。例えば、「play artist (再生するアーティスト)」という1つの文法は、ユーザによって操作されるクライアントデバイス106において artist による音楽を再生させる応答アクションを呼び出す意図にマッピングされ得る。「[weather | forecast]today([天気 | 予報]今日)」という別の文法は、「what's the weather today(今日の天気はなんですか)」および「what's the forecast for today?(今日の天気予報はなんですか)」などのユーザクエリに一致し得る場合がある。「play artist」の例示的な文法に見られるように、いくつかの文法は、スロット値(または「パラメータ」)で満たされ得るスロット(例えば、artist)を有する。スロット値は、様々な方法において決定され得る。しばしば、ユーザは、スロット値を積極的に提供する。例えば、「Order me a topping pizza(トッピングピザを注文してください)」という文法について、ユーザは、「order me a sausage pizza(ソーセージピザを注文してください)」というフレーズを話す可能性があり得、その場合、スロット topping は、自動的に満たされる。加えて、または代替的には、ユーザが、スロット値を積極的に提供せずに、スロット値で満たされるべきスロットを含む文法を呼び出す場合、自動化されたアシスタント120は、ユーザからのそれらのスロット値を要請し得る(例えば、「what type of crust do you want on your pizza?(ピザにどのタイプのクラストが欲しいですか)」)。

30

40

【0055】

(手動で作成され得る)多くの文法とは対照的に、機械学習モデルは、例えば、ユーザと

50

自動化されたアシスタントとの間の対話のログを使用して、自動的に訓練され得る。機械学習モデルは、ニューラルネットワークなどの様々な形態をとり得る。それは、ユーザの入力からユーザの意図を予測するために様々な方法において訓練され得る。例えば、いくつかの実装形態において、個々の訓練例を含む訓練データが提供され得る。各訓練例は、例えば、ユーザからの(例えば、テキストまたは非テキスト形式における)自由形式の入力を含み得、(例えば、手動で)意図でラベル付けされ得る。訓練例は、出力を生成するために、機械学習モデル(例えば、ニューラルネットワーク)にわたる入力として適用され得る。出力は、エラーを決定するためにラベルと比較され得る。このエラーは、モデルの隠れ層に関連付けられた重みを調整するために勾配降下(例えば、確率論的、バッチなど)および/またはバックプロパゲーションなどの技法を使用して、モデルを訓練するために使用され得る。そのようなモデルがいくつかの(通常は多数の)訓練例を用いて訓練されると、それは、ラベル付けされていない自由形式の自然言語入力から意図を予測する出力を生成するために使用され得る。

10

【 0 0 5 6 】

いくつかの実装形態において、自動化されたアシスタント120は、ユーザとサードパーティアプリケーションとの間のトランザクションを容易に(または「仲介」)し得る。これらのサードパーティアプリケーションは、例えばクラウドベースの自動化されたアシスタント構成要素119を動作させるものとは別のコンピューティングシステム上で動作する場合と動作しない場合がある。したがって、意図マッチャ136によって識別され得る1種類のユーザ意図は、サードパーティアプリケーションに関与することである。例えば、自動化されたアシスタント120は、ピザ配達サービスへのアプリケーションプログラミングインターフェース(「API」)へのアクセスを提供し得る。ユーザは、自動化されたアシスタント120を呼び出し、「I'd like to order a pizza.(ピザを注文したいです)」などのコマンドを提供し得る。意図マッチャ136は、このコマンドを、自動化されたアシスタント120がサードパーティのピザ配達サービスに関与するようにトリガする(場合によってはサードパーティによってデータベース137に追加され得る)文法にマッピングし得る。サードパーティのピザ配達サービスは、自動化されたアシスタント120に、ピザ配達注文を履行するために満たされる必要があるスロットの最小リストを提供し得る。自動化されたアシスタント120は、スロットのためのパラメータを要請する自然言語出力を生成し、(クライアントデバイス106を介して)ユーザに提供し得る。

20

30

【 0 0 5 7 】

履行エンジン124は、意図マッチャ136によって出力された意図、ならびに、任意の関連するスロット値(ユーザによって積極的に提供されたかユーザから要請されたかに関わらず)を受信し、意図を履行するように構成され得る。様々な実装形態において、ユーザの意図の履行は、例えば、履行エンジン124によって、様々な履行情報を生成/取得させ得る。以下に説明するように、履行情報は、いくつかの実装形態において、履行情報に基づいて自然言語出力を生成し得る自然言語生成器(いくつかの図における「NLG」)126に提供され得る。

【 0 0 5 8 】

意図は、様々な方法において履行され得るので、履行情報は、様々な形態をとり得る。ユーザが、「Where were the outdoor shots of 'The Shining' filmed?(シャイニングの野外ショットは、どこで撮影されたのですか)」などの純粋な情報を要求したとする。ユーザの意図は、例えば、意図マッチャ136によって、検索クエリであるとして決定され得る。検索クエリの意図および内容は、履行エンジン124に提供され得、履行エンジン124は、図1に示すように、応答情報のための文書および/または他のデータ源(例えば、ナレッジグラフなど)のコーパスを検索するように構成された1つまたは複数の検索エンジン150と通信し得る。履行エンジン124は、検索クエリを示すデータ(例えば、クエリのテキスト、次元減少した埋め込みなど)を検索エンジン150に提供し得る。検索エンジン150は、GPS座標などの応答情報、または、「Timberline Lodge, Mt. Hood, Oregon.(オレゴン州、フッド山、ティンバーラインロッジ)」などの、他のより明示的な情報を提供し得る。

40

50

この応答情報は、履行エンジン124によって生成された履行情報の一部を形成し得る。

【0059】

加えて、または代替的には、履行エンジン124は、例えば自然言語理解エンジン135から、ユーザによって提供された、または他の手段(例えば、ユーザのGPS座標、ユーザの好みなど)を使用して決定されたユーザの意図および任意のスロット値を受信し、応答アクションをトリガするように構成され得る。応答アクションは、例えば、商品/サービスを注文すること、タイマを開始すること、リマインダを設定すること、通話を開始すること、メディアを再生すること、メッセージを送信することなどを含み得る。いくつかのそのような実装形態において、履行情報は、履行に関連付けられたスロット値、(場合によっては所定の応答から選択され得る)確認応答などを含み得る。

10

【0060】

上記のように、自然言語生成器126は、様々なソースから取得されたデータに基づいて自然言語出力(例えば、人間のスピーチを模倣するように設計された話し言葉/フレーズ)を生成および/または選択するように構成され得る。いくつかの実装形態において、自然言語生成器126は、入力として、履行エンジン124による意図の履行に関連付けられた履行情報を受信し、履行情報に基づいて自然言語出力を生成するように構成され得る。加えて、または代替的には、自然言語生成器126は、(例えば、スロットを要求した)サードパーティアプリケーションなどの他のソースからの情報を受信し得、それは、ユーザのための自然言語出力を定式化するために使用され得る。

【0061】

対話コンテキストエンジン138は、ユーザと自動化されたアシスタント120との間の対話コンテキストを保持させ、利用可能にするように構成され得る。いくつかの実装形態において、対話コンテキストエンジン138は、クライアントデバイス106₁~106_Nからリモートにある1つまたは複数のデータベース139内に対話コンテキストを保持(例えば、記憶)し得る。加えて、または代替的には、対話コンテキストは、例えば、ユーザのプライバシーを維持するために、および/または、クライアントデバイス106がオフラインであるときであっても本明細書で説明する技法が実施され得るように、1つまたは複数のクライアントデバイス106に対してローカルであるメモリ内に記憶され得る。いくつかのそのような実装形態において、対話コンテキストエンジン138は、クライアントデバイス106上に全体的または部分的に実装され得る。

20

【0062】

様々な実装形態において、人間対コンピュータの対話中に生成され、対話コンテキストエンジン138によって管理される対話コンテキストは、様々な情報を含み得る。いくつかの実装形態において、対話コンテキストは、ユーザによって提供された自由形式の自然言語入力、および/または、例えば自然言語生成器126によって生成された自然言語出力などの様々な信号に基づいて、例えば意図マッチャ136によって検出された、1つまたは複数のユーザの意図を含み得る。いくつかの実装形態において、対話コンテキストは、人間対コンピュータの対話セッション中に(ユーザによっておよび/または自動化されたアシスタントによって)言及され、前述のエンティティタガーによって検出された既知のエンティティ(例えば、人、場所、物)への1つまたは複数の参照も含み得る。

30

【0063】

いくつかの実装形態において、対話コンテキストは、ユーザの意図を履行するために、すなわち、ユーザの意図に関連付けられた1つまたは複数のスロットを満たすことによって使用可能な1つまたは複数のスロット値をさらに含み得る。スロット値は、以前に説明したように、様々なソースから取得され得る。対話コンテキストは、対話コンテキストエンジン138によって様々な方法において保存され得る。いくつかの実装形態において、例えば、対話コンテキストが、保存された複写物内の意図、スロット値、エンティティなどを検出することによって、例えば、オンザフライでおよび/または必要に応じて再構築され得るように、データベース139内またはクライアントデバイス106に対してローカルなメモリ内に、人間対コンピュータの対話の複写物全体が保存される。いくつかの実装形態、

40

50

特に、ユーザがメッセージ交換クライアント107を使用して自動化されたアシスタント120と対話する実装形態において、これは、ユーザが複写物を視覚的に再検討することができ得るという追加の利点を提供し得る。加えて、または代替的には、いくつかの実装形態において、検出された意図、スロット値、言及されたエンティティなどの、対話コンテキストのコア要素のみが、JavaScript(登録商標) Object Notation(「JSON」)または他の同様のフォーマットなどの様々なフォーマットにおいて保存され得る。

【0064】

対話コンテキストエンジン138などの図1の様々な構成要素は、ユーザと自動化されたアシスタント120との間の人間対コンピュータの対話中に生成された複数の意味論的に異なる対話コンテキストを保存するための本開示の選択された態様を実行するように構成され得る。加えて、自動化されたアシスタントとの以前の会話を再開するために、ユーザが複数の対話コンテキストにおよび/またはそれらの間で遷移することを可能にするための技法について本明細書で説明する。例えば、いくつかの実装形態において、ユーザは、以前のスレッドおよび/または思考プロセスを失うことなく、直感的な方法で、自動化されたアシスタント120との以前の会話を通して遷移または「トグル」することができる。タッチスクリーン、静電容量式タッチパッド、または他の視覚的入力/出力(「I/O」)メカニズムがクライアントデバイス106において利用可能ないくつかの実装形態において、ユーザは、ユーザが以前の対話コンテキストに遷移するために操作することができる、「戻るボタン」などの選択可能要素を提示され得る。加えて、または代替的には、ユーザが自動化されたアシスタントとの人間対コンピュータの対話において音声で関与するためにクライアントデバイス106を操作するシナリオにおいて、ユーザは、自動化されたアシスタント120を以前の対話コンテキストに戻すために、「ねえアシスタント、以前の会話に戻りましょう」などの1つまたは複数のコマンドを発声することができ得る。

【0065】

図2A、図2B、図2C、および図2Dは、ユーザと、別のユーザと、ユーザによって操作されるコンピューティングデバイス206上に少なくとも部分的に実装された自動化されたアシスタント120との間の例示的な対話を示す。これらの図は、本開示の様々な態様、特に、過去の会話を再開するために複数の異なる過去の対話コンテキスト間を遷移するユーザの能力を示す。図2A~図2Dにおいて、クライアントデバイス206は、タッチスクリーン240と、様々なユーザインターフェースボタン281、282、283と、テキスト入力フィールド288と、マイクロフォンアイコン289とを含む。マイクロフォンアイコン289は、ユーザが音声入力を提供することを可能にするように動作可能であり得、音声入力は、例えば、スピーチ取り込み/TTS/STTモジュール114によってスピーチ認識され得、例えば、テキスト入力フィールド288に入力するために使用され得る。

【0066】

図2Aにおいて、クライアントデバイス206を操作するユーザ(「You(あなた)」)は、(例えば、入力フィールド288を使用して)以下の自由形式の自然言語入力、「SHOP FOR WOMEN'S DRESS SHOES IN SIZE 7(サイズ7の女性用ドレスシューズを買い物する)」を提供する。いくつかの実装形態において、この声明は、会話の新しいトピックを開始するので、靴の買い物のトピックに関連付けられた中断点が作成され得、ラベル「BP1」によって示されるように、複写物に関連付けられ得、ラベル「BP1」は、例示の目的のために図中に含まれ、タッチスクリーン240上に実際にレンダリングされる場合とされない場合がある。自動化されたアシスタントは、SHOE_SHOPPINGの予測された意図のためのスロット値を探す方法として、「DO YOU WANT TO SHOP FOR A SPECIFIC COLOR SHOE OR SPECIFIC BRAND?(特定の色の靴または特定のブランドを購入したいですか)」と尋ねる。ユーザは、「BLACK COLOR SHOES(黒い色の靴)」と返信することによってスロット値を提供する。

【0067】

この時点で、ジョーという名前の別のユーザが、複写物の一部として表示され、「WANT TO GO TO DINNER TONIGHT AT 7:00 AT VINCENZO'S?(今夜7時にVINCENZO'S

10

20

30

40

50

に夕食に行きたいですか)」と読み上げるメッセージをユーザに送信する。この時点で、新しい会話コンテキスト、今夜の夕食の開始を示すために、別の中断点、図2AにおけるBP2が複写物に挿入され得る。ユーザは、「YES(はい)」と応答し、それに応じて、自動化されたアシスタントは、「WANT TO BOOK A RESERVATION FOR 2 FOR DINNER TONIGHT AT 7:00 AT VINCENZO'S?(VINCENZO'Sでの今夜7時の夕食2人の予約をとりたいですか)」と割り込む。

【0068】

会話は、図2Bに続く。ユーザは、予約をするという自動化されたアシスタントの申し出に対して肯定して(「YES(はい)」)応答する。その意図(午後7時および「Vincenzo's」のパラメータで夕食の予約を行う)が履行されると、自動化されたアシスタントは、例えば、図1における自然言語出力生成器126として、「I HAVE BOOKED A DINNER RESERVATION FOR YOU AND JOE FOR TONIGHT AT 7:00 AT VINCENZO'S.(VINCENZO'Sでの今夜7時のあなたとジョーのための夕食の予約をとりました)」というテキストを生成して出力する。ジョーは、次いで、「THANKS. SEE YOU THEN.(ありがとう。またね。)」と再び割り込む。

【0069】

この時点で、ユーザと自動化されたアシスタントとの間のこの会話中にアクティブな2つの対話コンテキスト、すなわち、靴の買い物に関連するBP1に関連付けられた一方と、他方は、夕食の予約をすることに関連するBP2に関連付けられた他方が存在する。これらの対話コンテキストを示すデータは、例えば、データベース139内に対話コンテキストエンジン138によって保持される。その後、ユーザは、ユーザが最初から始めることなくそれらの過去の会話を再開することができるように、これらの対話コンテキスト(および、おそらくBP1よりも前に発生した他の対話コンテキスト)を通して戻ることができ得る。

【0070】

ユーザは、様々な方法においてこれらの対話コンテキストを通して戻り得る。いくつかの実装形態において、ユーザは、クライアントデバイス206のオペレーティングシステムによって提供される場合もされない場合もあるグラフィカル要素281などの戻るボタンを操作することによって、過去の対話コンテキストを逆向きにトグルし得る。図2Bにおいて、ユーザは、靴の買い物の対話コンテキストに戻るために、自由形式の自然言語入力、すなわち、「HEY ASSISTANT - BACK TO SHOE SHOPPING.(ねえアシスタント、靴の買い物に戻って)」を代わりに使用する。いくつかの実装形態において、同じ中断点BP1は、例えば、この中断点に関連付けられた複写物の異なる部分が単一の対話コンテキストに関連するものとして後に一緒に再構成され得るように、複写物のこの部分にもう一度挿入され得る。

【0071】

以前の対話コンテキストに戻るとき、自動化されたアシスタントと話し合っていたことをユーザに思い出させることが有益である場合がある。したがって、様々な実装形態において、以前の対話コンテキストに戻った後、自動化されたアシスタント120は、遷移された対話コンテキストの意図のうち少なくとも1つまたは複数と、遷移された対話コンテキストの1つまたは複数のスロット値とを伝達する自然言語出力を生成し得る。自動化されたアシスタント120は、次いで、この自然言語出力を、同じコンピューティングデバイスにおいて、またはユーザによって操作される異なるコンピューティングデバイスにおいて提示させ得る。例えば、図2Bにおいて、自動化されたアシスタント120は、「YOU WERE INTERESTED IN BLACK COLOR WOMEN'S DRESS SHOES IN SIZE 7(あなたは、サイズ7の黒色の婦人用ドレスシューズに興味がありました)」とすることによって、過去の対話コンテキストを要約する。

【0072】

図2Cは、ジョーが「THANKS. SEE YOU THEN.(ありがとう。またね。)」と言った後に発生し得る代替の対話を示す。ユーザが靴の買い物に積極的に戻る代わりに、ユーザは、代わりに、自動化されたアシスタント120に、「HEY ASSISTANT - WHAT WERE WE

10

20

30

40

50

TALKING ABOUT?(ねえアシスタント、私達は、何について話していましたか)」と尋ねる。自動化されたアシスタント120は、過去の対話コンテキストの列挙リスト「WE HAVE TALKED ABOUT: THE WEATHER FORECAST FOR TONIGHT; SHOE SHOPPING; AND DINNER RESERVATIONS FOR TONIGHT(私達は、今夜の天気予報、靴の買い物、および今夜の夕食の予約について話しました)」を出力することによって応答する。図2Cに示すように、これらの過去のトピックの各々は、中断点に関連付けられ得る。いくつかの実装形態において、中断点は、例えばユーザがそれぞれの対話コンテキストに戻るために1つを選択することができるように、タッチスクリーン240上に実際にレンダリングされ得る。

【0073】

しかしながら、図2Dにおいて、ユーザは、代わりに、「BACK TO SHOE SHOPPING. WHERE DID WE LEAVE OFF?(靴の買い物に戻って。私達は、どこでやめましたか。)」と言う。この時点で、自動化されたアシスタント120は、「YOU WERE INTERESTED IN BLACK COLOR WOMEN'S DRESS SHOES IN SIZE 7. DID YOU WANT TO SHOP FOR A SPECIFIC BRAND?(あなたは、サイズ7の黒色の婦人用ドレスシューズに興味がありました。特定のブランドを購入したかったですか)」と過去の会話に関する要約の自然言語出力を提供する。その後、ユーザは、「ALL BRANDS THAT ARE ON SALE(販売中のすべてのブランド)」と返答し、靴の買い物の会話が再開する。

【0074】

図3A、図3B、図3C、および図3Dは、様々な実装形態による、ユーザ301と、ユーザ301によって操作される独立型の対話型スピーカ306としての自動化されたアシスタント120との間の例示的な対話を示す。対話は、ユーザ301が独立型の対話型スピーカ306の自動化されたアシスタント120に、「SHOP FOR WOMEN'S DRESS SHOES IN SIZE 7.(サイズ7の女性用ドレスシューズを買い物する)」ように求めることで開始する。「SHOP(買い物する)」および「SHOES(靴)」への参照は、自動化されたアシスタント120が「SHOE SHOPPING(靴の買い物)」の第1の対話コンテキストを識別することを可能にする。第1の対話コンテキストを識別することに加えて、自動化されたアシスタント120は、第1の中断点BP1を第1の対話コンテキストに関連付け得る。中断点は、図3Aにおいて「BP1」として表されているが、これは、例示のみを目的とし、ユーザ301によって実際に話される可能性は低い。

【0075】

自動化されたアシスタント120は、ユーザ301に「DO YOU WANT TO SHOP FOR A SPECIFIC COLOR SHOE OR SPECIFIC BRAND?(特定の色の靴または特定のブランドを購入したいですか)」と尋ね得る。ユーザ301は、検索を「BLACK COLOR SHOES(黒色の靴)」に限定することによって応答し得る。この時点で、自動化されたアシスタント120は、以下の声明「REMINDER: DINNER TONIGHT AT 7:00 AT VINCENZO'S WITH JOE. WANT TO MAKE A RESERVATION?(リマインダ:今夜7時にVINCENZO'Sでジョーと夕食。予約をとりたいですか。)」を提供するために会話を中断し得る。これは、この特定の時間に配信されるようにユーザ301が自分自身に対して以前に設定したリマインダ、または、ジョーなどの他の誰かによってトリガされたリマインダであった場合がある。とにかく、「DINNER(夕食)」および「RESERVATION(予約)」への参照は、自動化されたアシスタント120が「DINNER RESERVATION(夕食の予約)」の第2の対話コンテキストを識別することを可能にする。第2の対話コンテキストを識別することに加えて、自動化されたアシスタント120は、第2の中断点BP2を第2の対話コンテキストに関連付け得る。

【0076】

リマインダに直接対処する代わりに、ユーザ301は、「WHAT IS THE WEATHER FORECAST FOR TONIGHT?(今夜の天気予報はなんですか)」と応答して、主題をもう一度変更する。これは、例えば、ユーザが、天気がいい場合、外に座ることを好むためであり得る。「THE WEATHER FORECAST FOR TONIGHT(今夜の天気予報)」への参照は、自

10

20

30

40

50

動化されたアシスタント120が「THE WEATHER FORECAST FOR TONIGHT(今夜の天気予報)」の第3の対話コンテキストを識別することを可能にする。第3の対話コンテキストを識別することに加えて、自動化されたアシスタント120は、第3の中断点BP3を第3の対話コンテキストに関連付け得る。自動化されたアシスタント120は、「WARM AND CLEAR SKIES.(暖かくて晴れ)」と述べることによってその天気予報に回答し得る。

【0077】

この好都合の天気予報は、夕食の予約をする際のユーザの決定に影響を与える場合がある。例えば、ユーザ301は、屋内で食べることを選択するのではなく、「MAKE THE DINNER RESERVATION FOR TONIGHT AT VINCENZO'S WITH PATIO SEATING(パティオ席でVINCENZO'Sでの今夜の夕食の予約をしてください)」で回答し得る。ユーザ301に回答して、自動化されたアシスタント120は、夕食の予約を行うためにサードパーティアプリケーションと対話し得、「DINNER RESERVATION BOOKED FOR TONIGHT AT 7:00 AT VINCENZO'S.(VINCENZO'Sでの今夜7時に予約された夕食の予約)」があることをユーザに警告し得る。

10

【0078】

会話は、図3Bに続く。ユーザ301は、その自動化されたアシスタント120に「MESSAGE JOE: I BOOKED A DINNER RESERVATION FOR TONIGHT AT 7:00 AT VINCENZO'S WITH PATIO SEATING.(ジョーにメッセージ:VINCENZO'Sのパティオ席での今夜7時の夕食の予約をしました)」を要求し得る。さらに、自動化されたアシスタント120は、「MESSAGE SENT TO JOE.(メッセージがジョーに送信されました)」と回答することによって要求を確認し得る。ユーザ301が以前の対話コンテキストを忘れたと仮定する。上記のように、ユーザ301が以前の対話コンテキストに戻るために用い得る様々な技法が存在する。

20

【0079】

いくつかの実装形態において、図3Bに示すように、ユーザ301は、「HEY ASSISTANT - WHAT WERE WE TALKING ABOUT?(ねえアシスタント、私達は、何について話していましたか)」と言うことによって、以前の対話コンテキストに戻り得る。それに回答して、自動化されたアシスタント120は、「WE HAVE TALKED ABOUT: 1) SHOE SHOPPING; 2) DINNER RESERVATION; AND 3) THE WEATHER FORECAST FOR TONIGHT.(私達は、1)靴の買い物、2)夕食の予約、および3)今夜の天気予報について話していました)」などの以前の対話コンテキストの列挙リストを提供し得る。ユーザ301が「SHOE SHOPPING(靴の買い物)」を指定した場合、自動化されたアシスタント120は、第1の中断点BP1に関連付けられた対話に戻ることによって、第1の対話コンテキストに戻り得る。ユーザ301が「DINNER RESERVATION(夕食の予約)」を指定した場合、自動化されたアシスタント120は、第2の中断点BP2に関連付けられた対話に戻ることによって、第2の対話コンテキストに戻り得る。ユーザ301が「THE WEATHER FORECAST FOR TONIGHT(今夜の天気予報)」を指定した場合、自動化されたアシスタント120は、第3の中断点BP3に関連付けられた対話に戻ることによって、第3の対話コンテキストに戻り得る。

30

【0080】

ユーザ301が列挙された以前の対話コンテキストのうちの1つに戻ることを選択した場合、自動化されたアシスタント120は、特定の列挙された対話コンテキストの要約を提供し得る。例えば、ユーザ301は、「BACK TO SHOE SHOPPING. WHERE DID WE LEAVE OFF?(靴の買い物に戻って。私達は、どこでやめましたか。)」と言い得る。自動化されたアシスタント120は、「YOU WERE INTERESTED IN BLACK COLOR WOMEN'S DRESS SHOES IN SIZE 7. DID YOU WANT TO SHOP FOR A SPECIFIC BRAND?(あなたは、サイズ7の黒色の婦人用ドレスシューズに興味がありました。特定のブランドを購入したかったですか)」と回答し得る。ユーザ301は、「ALL BRANDS THAT ARE ON SALE(販売中のすべてのブランド)」のように言うことによって回答し得る。

40

【0081】

図3Cは、図3Bに示すものに対する代替対話を示す。いくつかの実装形態において、図3

50

Cに示すように、ユーザ301は、「HEY ASSISTANT - WHAT WERE WE TALKING ABOUT?(ねえアシスタント、私達は、何について話していましたか)」と言うことによって、直前の対話コンテキストに戻り得る。それに応答して、自動化されたアシスタント120は、直前の対話コンテキスト「WE WERE TALKING ABOUT THE WEATHER FORECAST FOR TONIGHT.(私達は、今夜の天気予報について話していました)」を提供し得る。しかしながら、ユーザ301は、より以前の対話コンテキストに戻りたい場合があり、「BEFORE THAT(その前)」と言うことによってこれを示すことができる。それに応答して、自動化されたアシスタント120は、より以前の対話コンテキスト「WE WERE TALKING ABOUT YOUR DINNER RESERVATION TONIGHT WITH JOE.(私達は、ジョーと今夜のあなたの夕食の予約について話していました)」を提供し得る。さらに、ユーザ301は、さらにより以前の対話コンテキストに戻りたい(例えば、さらにトグルしたい)場合があり、「BEFORE THAT」と言うことによってこれを示すことができる。それに応答して、自動化されたアシスタント120は、さらにより以前の対話コンテキスト「WE WERE TALKING ABOUT SHOE SHOPPING.(私達は、靴の買い物について話していました)」を提供し得る。このようにして、ユーザ301は、所望の対話コンテキストに到達し得る。

【0082】

いくつかの実装形態において、図3Dに示すように、ユーザ301は、「BACK TO SHOE SHOPPING. WHERE DID WE LEAVE OFF?(靴の買い物に戻って。私達は、どこでやめましたか。)」と指定することによって、以前の対話コンテキストに戻り得る。それに応答して、自動化されたアシスタント120は、以前の対話コンテキストの要約「YOU WERE INTERESTED IN BLACK COLOR WOMEN'S DRESS SHOES IN SIZE 7. DID YOU WANT TO SHOP FOR A SPECIFIC BRAND?(あなたは、サイズ7の黒色の婦人用ドレスシューズに興味がありました。特定のブランドを購入したかったですか)」を提供し得る。ユーザは、次いで、検索を「ALL BRANDS THAT ARE ON SALE(販売中のすべてのブランド)」に限定する口頭の出力を提供し得る。それに応答して、自動化されたアシスタント120は、「BRAND X AND BRAND Y ARE BOTH HAVING A SALE ON BLACK COLOR WOMEN'S DRESS SHOES IN SIZE 7.(ブランドXとブランドYの両方が、サイズ7の黒色の女性用ドレスシューズを販売している)」と決定し得る。以前のコンテキストに基づく対話は、続く場合があり、または、ユーザ301は、検索を十分に限定し得、「THANKS(ありがとう)」と言う場合がある。

【0083】

図3A~図3Dにおいて、ユーザ301は、音声コマンドを使用して対話状態間を遷移した。しかしながら、これは、限定ではない。代替として、ユーザ301は、独立型の対話型スピーカ306上の戻るボタン381と対話することができた。例えば、独立型の対話型スピーカ306は、選択可能な戻るボタンをレンダリングするように動作し得る、例えば、静電容量式タッチパッドの下に配置された1つまたは複数の1つまたは複数のLEDを有する、基本的なディスプレイを含み得る。加えて、または代替的には、いくつかの独立型の対話型スピーカは、例えば、戻るボタンをレンダリングするために、または、過去の対話コンテキストを要約するテキストをレンダリングする、過去の対話コンテキストをリストする(および、リストされたアイテムは、ユーザによって手動で選択され得る)などを行うために使用され得る、取り付けられたタッチスクリーンディスプレイを含み得る。

【0084】

図4は、本明細書で開示される実装形態による例示的な方法400を示すフローチャートである。便宜上、フローチャートの動作について、動作を実行するシステムを参照して説明する。このシステムは、自動化されたアシスタント120を実装するコンピューティングシステムの1つまたは複数の構成要素などの、様々なコンピュータシステムの様々な構成要素を含み得る。さらに、方法400の動作は、特定の順序において示されているが、これは、限定であることを意味しない。1つまたは複数の動作が、並べ替えられ得、省略され得、または追加され得る。

【0085】

10

20

30

40

50

ブロック452において、システムは、ユーザと自動化されたアシスタント120との間の人間対コンピュータの対話セッション中に、ユーザから自由形式の自然言語入力の一つまたは複数のインスタンスを受信し得る。この自由形式の入力は、(例えば、メッセージ交換クライアント107を使用して)タイピングおよび/または話され得、次いで、テキストにスピーチ認識され得る。ブロック454において、システムは、自由形式の自然言語入力の一つまたは複数のインスタンスに基づいて、第1の対話コンテキストを生成し得る。第1の対話コンテキストは、例えば、ユーザの一つまたは複数の意図、ならびに、一つまたは複数の意図に関連付けられた一つまたは複数のスロット値を含み得る。

【0086】

ブロック456において、システムは、同じ人間対コンピュータの対話セッションまたはユーザと自動化されたアシスタント120との間の異なる人間対コンピュータの対話セッション中に、ユーザから自由形式の自然言語入力の一つまたは複数の追加インスタンスを受信し得る。自由形式の自然言語入力のこれらの追加インスタンスは、ブロック452のものとは異なる主題に関連し得る。ブロック458において、システムは、自由形式の自然言語入力の一つまたは複数の追加インスタンスに基づいて、第2の対話コンテキストを生成し得る。第2の対話コンテキストは、ユーザの一つまたは複数の追加の意図を含み得、ブロック454において生成された第1の対話コンテキストとは意味論的に異なるとみなされ得る。

【0087】

ブロック460において、システムは、自動化されたアシスタント120を第2の対話コンテキストから第1の対話コンテキストに戻すためのユーザからの遷移コマンドを受信し得る。このコマンドは、話されるか、または別の入力モダリティを使用して、例えば、戻るボタンを使用して提供され得る。ブロック462において、システムは、第1の対話コンテキストに基づいて、第1の対話コンテキストの意図のうちの少なくとも一つまたは複数と、第1の対話コンテキストのスロット値のうちの一つまたは複数とを伝達する自然言語出力を生成し得る。ブロック464において、システムは、自然言語出力を、同じコンピューティングデバイスにおいて、またはユーザによって操作される異なるコンピューティングデバイスにおいて提示させ得る。例えば、クラウドベースの自動化されたアシスタント構成要素119は、自然言語出力のテキストをスピーチ取り込み/TTS/STTモジュール114に送信し得、スピーチ取り込み/TTS/STTモジュール114は、テキストをスピーチに変換し、クライアントデバイス106の一つまたは複数のスピーカを介して出力し得る。

【0088】

いくつかの実装形態において、ユーザは、場合によっては、自動化されたアシスタント120との人間対コンピュータの対話からユーザによってアクセス可能なサードパーティアプリケーションであり得る別個のソフトウェアアプリケーションに参与するために、自動化されたアシスタント120との会話から外され得る。これは、例えば、自動化されたアシスタント120が、別個のソフトウェアアプリケーションを起動するためにユーザによって選択可能ないわゆる「ディープリンク」をユーザに提示するときに発生し得る。そのようなディープリンクは、例えば人間対コンピュータの対話中に検出された一つまたは複数の意図またはスロット値に基づいて、自動化されたアシスタント120によって提示され得る。いくつかのそのような実装形態において、別個のソフトウェアアプリケーションは、ディープリンクの選択の前にアクティブであった対話コンテキストがもう一度アクティブになって、自動化されたアシスタント120と対話するために使用可能なインターフェースに戻るために選択可能な、戻るボタンなどの選択可能要素を含むグラフィカルユーザインターフェースを含み得る。他の実装形態において、ユーザは、戻るために、基礎となるオペレーティングシステムによってレンダリングされた戻るボタンなどの、別個のソフトウェアアプリケーションに特に関連しない戻るボタンを操作し得る。

【0089】

図5は、様々な実施形態による、本開示の選択された態様を実施するための例示的な方法500を示す。便宜上、フローチャートの動作について、動作を実行するシステムを参照して説明する。このシステムは、様々なコンピュータシステムの様々な構成要素を含み得

10

20

30

40

50

る。さらに、方法500の動作は、特定の順序において示されているが、これは、限定であることを意味しない。1つまたは複数の動作が、並べ替えられ得、省略され得、または追加され得る。

【0090】

ブロック552において、システムは、ユーザによって操作されるコンピューティングデバイスの1つまたは複数の入力構成要素において、ユーザと1つまたは複数のプロセッサによって少なくとも部分的に実装された自動化されたアシスタントとの間の第1の人間対コンピュータの対話中にユーザから自由形式の自然言語入力の1つまたは複数のインスタンスを受信し得る。この動作は、図4のブロック452と同様であり得る。

【0091】

ブロック554において、システムは、ユーザと自動化されたアシスタントとの間の第1の人間対コンピュータの対話の1つまたは複数の記憶された第1のパラメータを、ユーザと自動化されたアシスタントとの間の第1の人間対コンピュータの対話からユーザによってアクセス可能な第1のサードパーティアプリケーションに関連付け得る。これらのパラメータは、例えばリモートサーバにおいておよび/またはクライアントデバイスにおいてローカルに配置されたコンピュータメモリ内に記憶され得る。いくつかの実装形態において、1つまたは複数の第1のパラメータを記憶することは、1つまたは複数の第1のパラメータのための構成ファイルを作成することと、構成ファイルをコンピュータメモリ内に記憶することとを含み得る。場合によっては、構成ファイルは、JSONなどのフォーマットにおいて作成され得、実際には、対話コンテキストを表し得る。例示的な使用例として、第1のサードパーティアプリケーション(例えば、靴の買い物アプリケーション)へのディープリンクが第1の人間対コンピュータの対話の一部として提示される場合、1つまたは複数の意図、スロット値、言及されたエンティティなど(例えば、所望の靴のサイズ、スタイル、色など)を含み得る対話コンテキストが、第1のサードパーティアプリケーションに関連付けられ得る。

【0092】

ブロック556において、システムは、コンピューティングデバイスの1つまたは複数の入力構成要素において、ユーザと自動化されたアシスタントとの間の第2の人間対コンピュータの対話中に、ユーザから自由形式の自然言語入力の1つまたは複数のさらなるインスタンスを受信し得る。この動作は、図4のブロック456と同様であり得る。

【0093】

ユーザと自動化されたアシスタントとの間の第2の人間対コンピュータの対話に続いて、第1のサードパーティアプリケーションから自動化されたアシスタントに遷移するためのユーザからの遷移コマンドの受信に応答して、ブロック558において、システムは、例えばコンピュータメモリから、第1のサードパーティアプリケーションに関連付けられた記憶された1つまたは複数の第1のパラメータを取得し得る。ブロック560において、システムは、ユーザと自動化されたアシスタントとの間の第1の人間対コンピュータの対話を再開するために、取得された第1のパラメータを使用し得る。このようにして、ユーザは、ユーザがサードパーティアプリケーションへのディープリンクを選択する前に、自動化されたアシスタントとの会話を再開することができる。

【0094】

いくつかの実装形態において、第1の人間対コンピュータの対話の1つまたは複数の記憶された第1のパラメータを第1のサードパーティアプリケーションに関連付けることは、第1のサードパーティアプリケーションに関連付けられた記憶された1つまたは複数の第1のパラメータを取得するために、例えば第1のサードパーティアプリケーションから自動化されたアシスタントに遷移するための遷移コマンドの受信に応答して、呼び出し可能なトークンを作成することを含み得る。いくつかのそのような実装形態において、トークンは、記憶された1つまたは複数の第1のパラメータと第1のサードパーティアプリケーションとの間のリンクの形態をとり得る。

【0095】

10

20

30

40

50

図6は、本明細書で説明する技法の1つまたは複数の態様を実行するためにオプションで利用され得る例示的なコンピューティングデバイス610のブロック図である。いくつかの実装形態において、クライアントコンピューティングデバイス、ユーザ制御のリソースエンジン130、および/または他の構成要素のうちの1つまたは複数の、例示的なコンピューティングデバイス610の1つまたは複数の構成要素を備え得る。

【0096】

コンピューティングデバイス610は、典型的には、バスサブシステム612を介していくつかの周辺デバイスと通信する少なくとも1つのプロセッサ614を含む。これらの周辺デバイスは、例えばメモリサブシステム625とファイル記憶サブシステム626とを含む記憶サブシステム624と、ユーザインターフェース出力デバイス620と、ユーザインターフェース入力デバイス622と、ネットワークインターフェースサブシステム616とを含み得る。入力および出力デバイスは、コンピューティングデバイス610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616は、外部ネットワークへのインターフェースを提供し、他のコンピューティングデバイス内の対応するインターフェースデバイスに結合される。

10

【0097】

ユーザインターフェース入力デバイス622は、キーボード、マウス、トラックボール、タッチパッド、もしくはグラフィックタブレットなどのポインティングデバイス、スキャナ、ディスプレイに組み込まれたタッチスクリーン、音声認識システムなどのオーディオ入力デバイス、マイクロフォン、および/または他のタイプの入力デバイスを含み得る。一般に、「入力デバイス」という用語の使用は、情報をコンピューティングデバイス610内または通信ネットワーク上に入力するためのすべての可能なタイプのデバイスおよび方法を含むことを意図している。

20

【0098】

ユーザインターフェース出力デバイス620は、ディスプレイサブシステム、プリンタ、ファックスマシン、または、オーディオ出力デバイスなどの他の非視覚的ディスプレイを含み得る。ディスプレイサブシステムは、陰極線管(CRT)、液晶ディスプレイ(LCD)などのフラットパネルディスプレイ、投影デバイス、または、可視画像を作成するための何か他のメカニズムを含み得る。ディスプレイサブシステムは、オーディオ出力デバイスを介するなどの非視覚的ディスプレイも提供し得る。一般に、「出力デバイス」という用語の使用は、コンピューティングデバイス610からユーザまたは別のマシンもしくはコンピューティングデバイスに情報を出力するためのすべての可能なタイプのデバイスおよび方法を含むことを意図している。

30

【0099】

記憶サブシステム624は、本明細書で説明する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およびデータ構成物を記憶する。例えば、記憶サブシステム624は、図4および図5の方法の選択された態様を実行し、同様に図1に示す様々な構成要素を実装するためのロジックを含み得る。

【0100】

これらのソフトウェアモジュールは、一般に、プロセッサ614単独によって、または他のプロセッサと組み合わせて実行される。記憶サブシステム624内で使用されるメモリ625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)630と、固定された命令が記憶される読み出し専用メモリ(ROM)632とを含むいくつかのメモリを含むことができる。ファイル記憶サブシステム626は、プログラムおよびデータファイルのための永続的な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、関連するリムーバブル媒体を有するフロッピーディスクドライブ、CD-ROMドライブ、光学ドライブ、またはリムーバブル媒体カートリッジを含み得る。特定の実装形態の機能を実装するモジュールは、ファイル記憶サブシステム626によって記憶サブシステム624内またはプロセッサ614によってアクセス可能な他のマシン内に記憶され得る。

40

【0101】

50

バスサブシステム612は、コンピューティングデバイス610の様々な構成要素およびサブシステムを意図したよう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612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実装形態は、複数のバスを使用し得る。

【0102】

コンピューティングデバイス610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、または任意の他のデータ処理システムもしくはコンピューティングデバイスを含む様々なタイプのものであり得る。コンピュータおよびネットワークの絶えず変化する性質のため、図6に示すコンピューティングデバイス610の説明は、いくつかの実装形態を説明する目的のための特定の例としてのみ意図されている。コンピューティングデバイス610の多くの他の構成は、図6に示すコンピューティングデバイスよりも多いまたは少ない構成要素を有することが可能である。

10

【0103】

本明細書で論じる特定の实装形態がユーザに関する個人情報(例えば、他の電子通信から抽出されたユーザデータ、ユーザのソーシャルネットワークに関する情報、ユーザの位置、ユーザの時間、ユーザの生体情報、ならびにユーザの活動および人口統計情報、ユーザ間の関係など)を収集または使用し得る状況において、ユーザには、情報が収集されるかどうか、個人情報が記憶されるかどうか、個人情報が使用されるかどうか、および、どのようにユーザに関する情報が収集され、記憶され、使用されるかを制御する1つまたは複数の機会が提供される。すなわち、本明細書で論じるシステムおよび方法は、関連するユーザからそうするための明示的な承認を受け取った場合のみ、ユーザの個人情報を収集、記憶、および/または使用する。

20

【0104】

例えば、ユーザには、プログラムまたは機能がその特定のユーザまたはプログラムもしくは機能に関連する他のユーザに関するユーザ情報を収集するかどうかについての制御を提供される。個人情報が収集される各ユーザは、情報が収集されるかどうかについてまたは情報のどの部分が収集されるかについて許可または承認を提供するために、そのユーザに関連する情報収集について制御することを可能にするために、1つまたは複数のオプションを提供される。例えば、ユーザは、通信ネットワークを介して1つまたは複数のそのような制御オプションを提供され得る。加えて、特定のデータは、個人を特定できる情報が削除されるように、記憶または使用される前に1つまたは複数の方法において処理され得る。一例として、ユーザの識別情報は、個人を特定できる情報を決定することができないように処理され得る。別の例として、ユーザの地理的位置は、ユーザの特定の位置を決定することができないように、より大きい領域に一般化され得る。

30

【0105】

いくつかの実装形態について本明細書で説明および例示してきたが、機能を実行するためのならびに/または結果および/もしくは本明細書で説明される利点のうちの1つまたは複数を得るための様々な他の手段および/または構造が利用可能であり得、そのような変形および/または変更の各々は、本明細書で説明される実装形態の範囲内にあるとみなされる。より一般的には、本明細書で説明するすべてのパラメータ、寸法、材料、および構成は、例示であり、実際のパラメータ、寸法、材料、および/または構成は、教示が使用される特定の用途に依存することを意味する。当業者は、日常的な実験のみを使用して、本明細書で説明する特定の实装形態に対する多くの均等物を認識するか、または確認することができる。したがって、前述の実装形態は、単なる例として提示されており、添付した特許請求の範囲およびその均等物の範囲内で、実装形態は、具体的に説明および特許請求されている以外の方法で実施され得ることが理解されるべきである。本開示の実装形態は、本明細書で説明する個々の特徴、システム、物品、材料、キット、および/または方法の各々に向けられている。加えて、2つ以上のそのような特徴、システム、物品、材料、キット、および/または方法の任意の組合せは、そのような特徴、システム、物品、材料、キット、および/または方法が相互に矛盾しない場合、本開示の範囲内に含まれる。

40

50

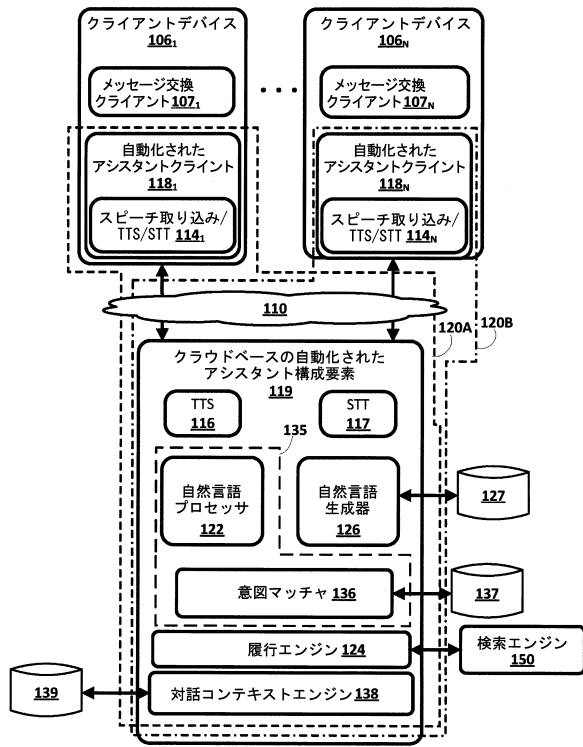
【符号の説明】

【0106】

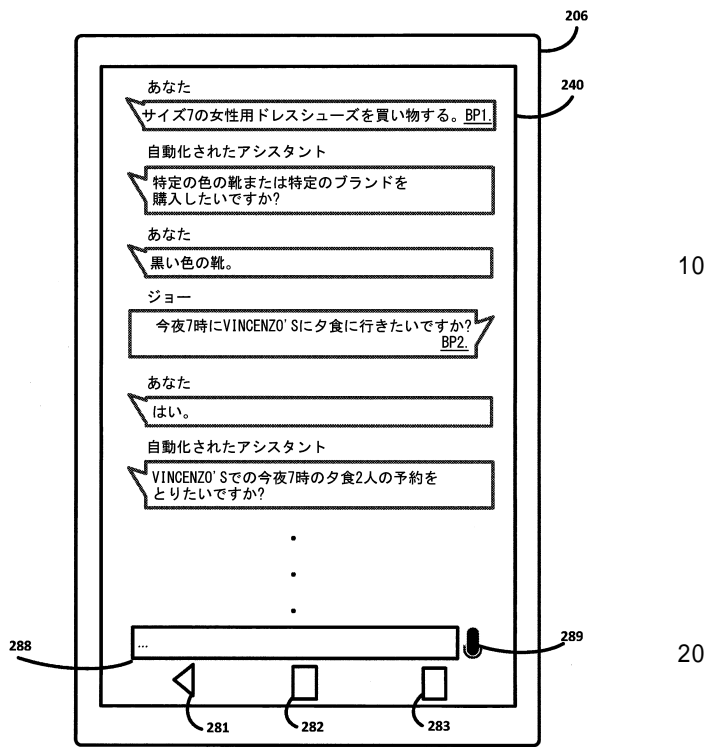
106	クライアントデバイス	
106 ₁ ~ 106 _N	コンピューティングデバイス、クライアントデバイス、クライアントコンピューティングデバイス	
107	メッセージ交換クライアント	
107 ₁ ~ 107 _N	メッセージ交換クライアント	
110	ローカルおよび/またはワイドエリアネットワーク	
114	スピーチ取り込み/テキストからスピーチへ(「TTS」)/スピーチからテキストへ(「STT」)モジュール、スピーチ取り込み/TTS/STT	10
116	クラウドベースのTTSモジュール	
117	クラウドベースのSTTモジュール	
118、118 ₁ 、118 _N	自動化されたアシスタントクライアント	
119	クラウドベースの自動化されたアシスタント構成要素	
120A	第1の自動化されたアシスタント	
120B	第2の自動化されたアシスタント	
122	自然言語プロセッサ	
124	履行エンジン	
126	自然言語生成器	
130	ユーザ制御のリソースエンジン	20
135	自然言語理解エンジン	
136	意図マッチャ	
137	データベース	
138	対話コンテキストエンジン	
139	データベース	
150	検索エンジン	
206	コンピューティングデバイス、クライアントデバイス	
240	タッチスクリーン	
281	ユーザインターフェースボタン、グラフィカル要素	
282	ユーザインターフェースボタン	30
283	ユーザインターフェースボタン	
288	テキスト入力フィールド、入力フィールド	
289	マイクロフォンアイコン	
306	独立型の対話型スピーカ	
381	戻るボタン	
610	コンピューティングデバイス	
612	バスサブシステム	
614	プロセッサ	
616	ネットワークインターフェースサブシステム	
620	ユーザインターフェース出力デバイス	40
622	ユーザインターフェース入力デバイス	
624	記憶サブシステム	
625	メモリサブシステム、メモリ	
630	メインランダムアクセスメモリ(RAM)	
632	読み出し専用メモリ(ROM)	

【 図 面 】

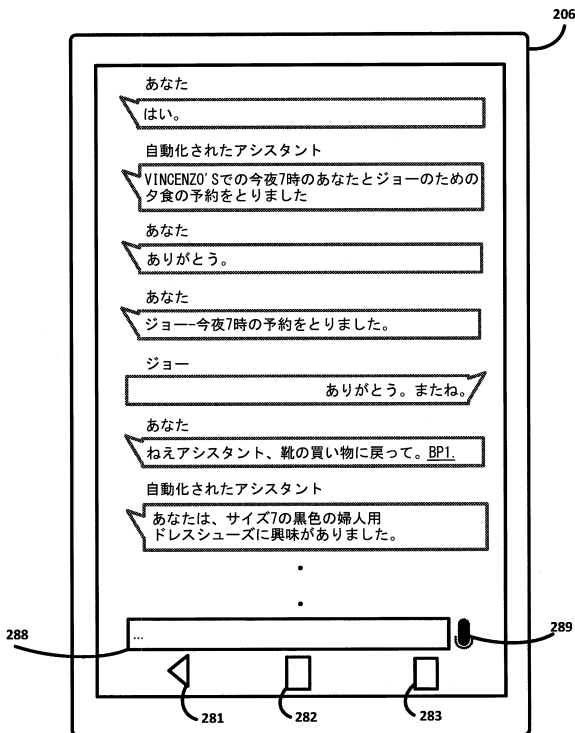
【 図 1 】



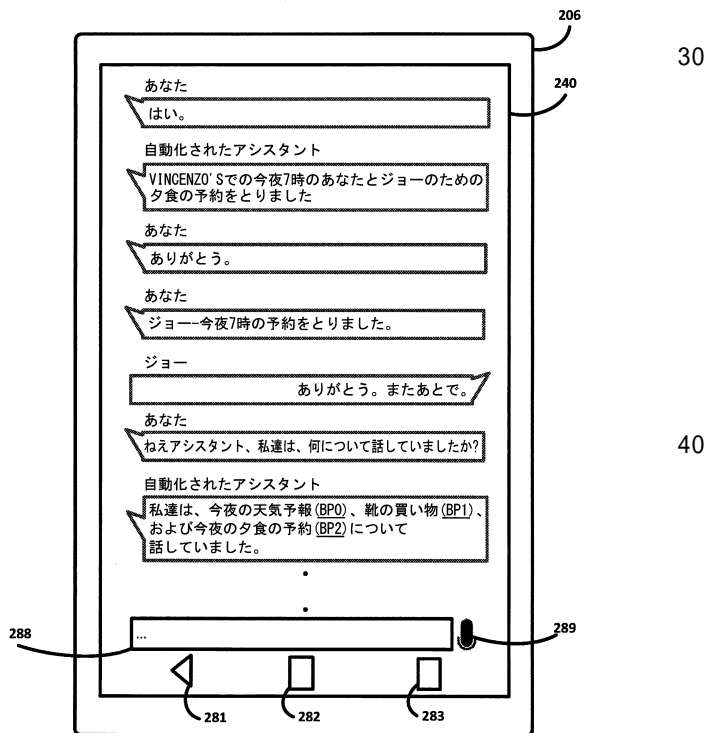
【 図 2 A 】



【 図 2 B 】



【 図 2 C 】



10

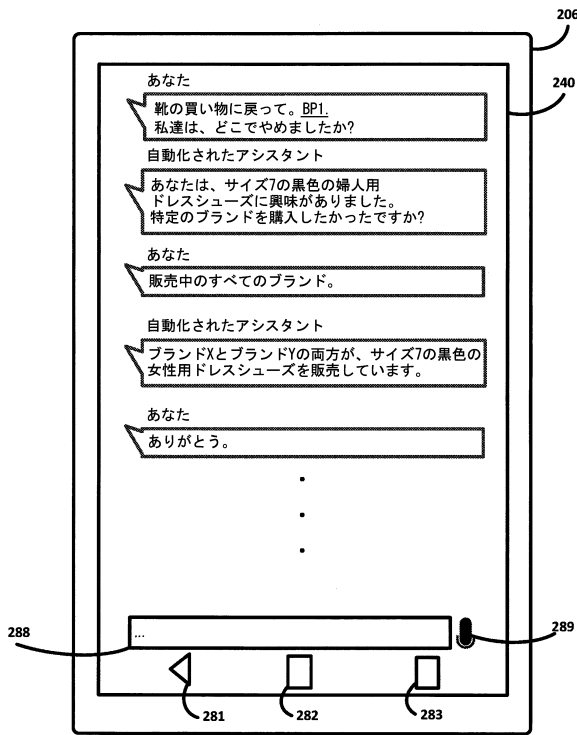
20

30

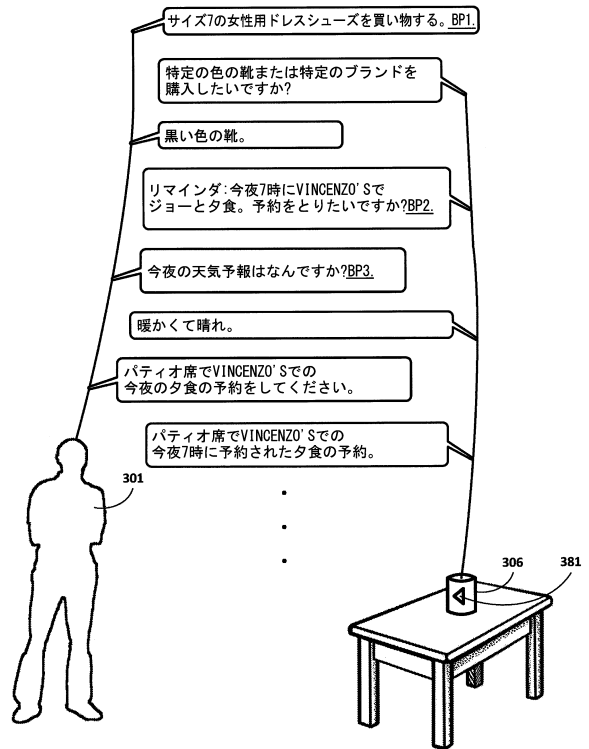
40

50

【図 2 D】



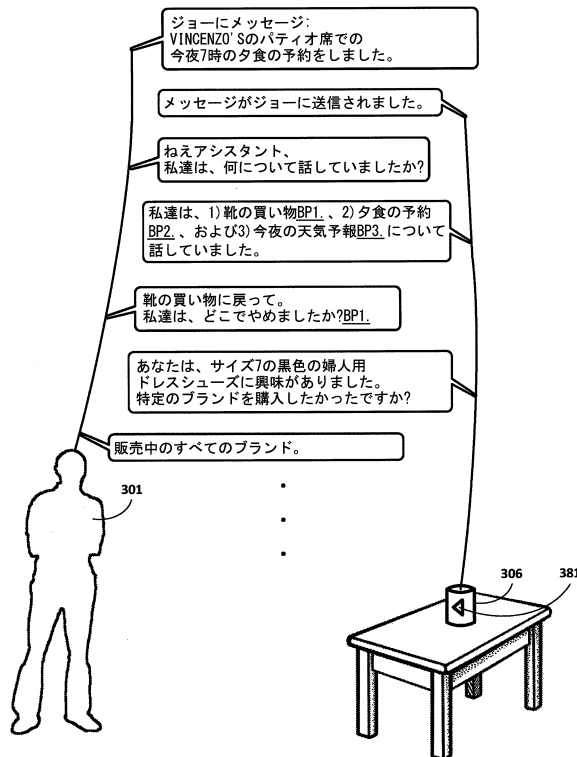
【図 3 A】



10

20

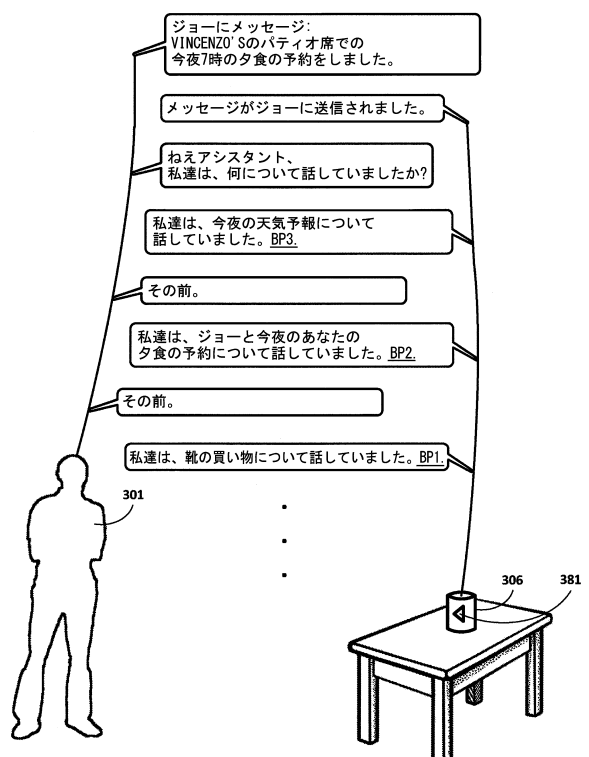
【図 3 B】



30

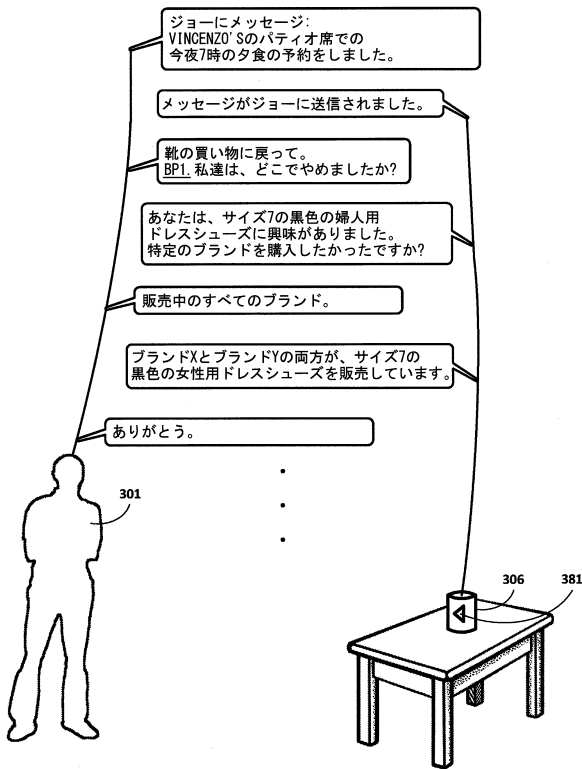
40

【図 3 C】

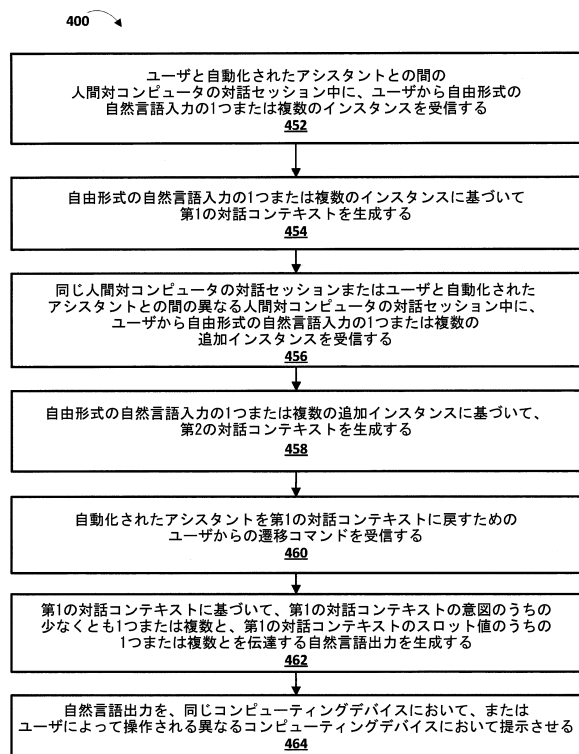


50

【図3D】



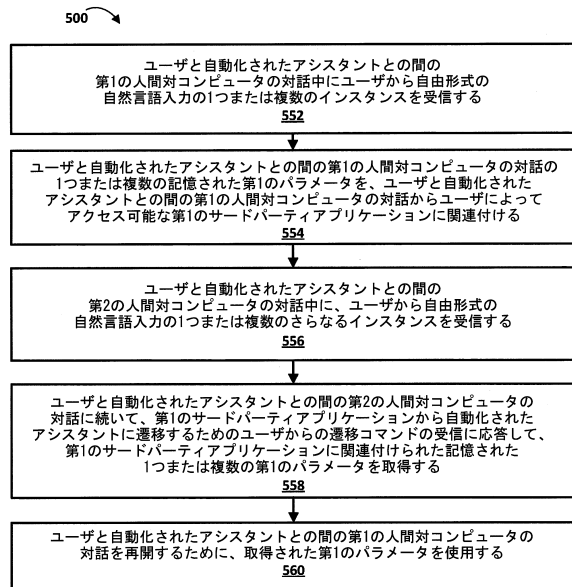
【図4】



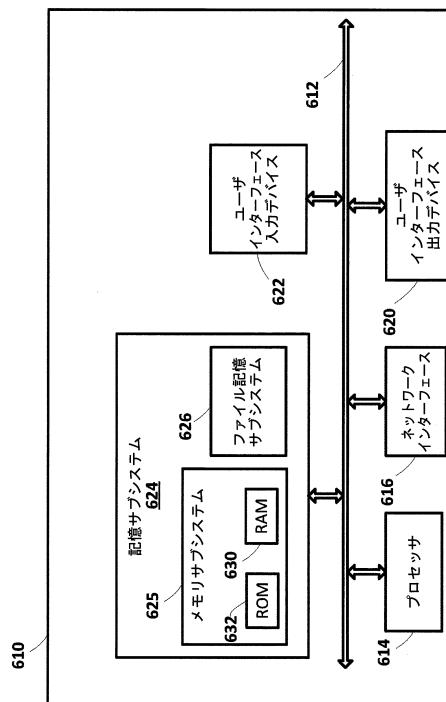
10

20

【図5】



【図6】



30

40

50

フロントページの続き

(51)国際特許分類

F I
G 0 6 F 3/16 6 3 0

(72)発明者

ジャスティン・ルイス

アメリカ合衆国・カリフォルニア・9 4 0 4 3・マウンテン・ビュー・アンフィシアター・パーク
ウェイ・1 6 0 0

(72)発明者

スコット・デイヴィース

アメリカ合衆国・カリフォルニア・9 4 0 4 3・マウンテン・ビュー・アンフィシアター・パーク
ウェイ・1 6 0 0

審査官 長 由紀子

(56)参考文献

特開 2 0 0 5 - 2 1 5 6 9 4 (J P , A)

米国特許出願公開第 2 0 1 7 / 0 3 0 0 8 3 1 (U S , A 1)

特表 2 0 1 9 - 5 2 0 6 6 1 (J P , A)

米国特許出願公開第 2 0 1 5 / 0 1 4 9 1 7 7 (U S , A 1)

特開 2 0 1 3 - 0 8 0 4 7 6 (J P , A)

(58)調査した分野 (Int.Cl., D B名)

G 0 6 F 4 0 / 0 0 - 5 8

G 1 0 L 1 5 / 2 2

G 0 6 F 1 6 / 9 0