

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-144598

(P2019-144598A)

(43) 公開日 令和1年8月29日(2019.8.29)

(5) Int.Cl.	F I	テーマコード (参考)
G10L 15/10 (2006.01)	G10L 15/10 500T	
G06F 3/16 (2006.01)	G10L 15/10 200W	
	G10L 15/10 500Z	
	G06F 3/16 630	

審査請求 有 請求項の数 14 O L 外国語出願 (全 30 頁)

(21) 出願番号 特願2019-101151 (P2019-101151)
 (22) 出願日 令和1年5月30日(2019.5.30)
 (62) 分割の表示 特願2017-550871 (P2017-550871) の分割
 原出願日 平成28年4月12日(2016.4.12)
 (31) 優先権主張番号 14/693,330
 (32) 優先日 平成27年4月22日(2015.4.22)
 (33) 優先権主張国・地域又は機関 米国 (US)

(71) 出願人 502208397
 グーグル エルエルシー
 アメリカ合衆国 カリフォルニア州 94043 マウンテン ビュー アンフィシアター パークウェイ 1600
 (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
 (72) 発明者 ボ・ワン
 アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パークウェイ・1600

最終頁に続く

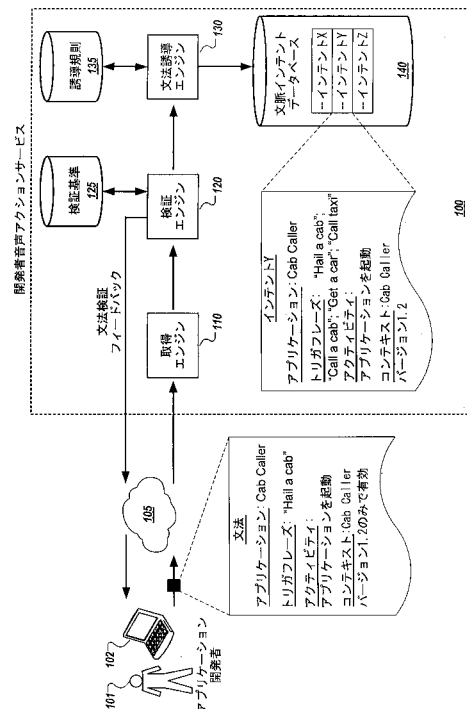
(54) 【発明の名称】 開発者音声アクションシステム

(57) 【要約】

【課題】 インテントによって指定されたアクティビティが、インテントによって指定されたターゲットアプリケーションにおいて、または、インテントによって指定されたターゲットアプリケーションによって実行される技法を提供する。

【解決手段】 アプリケーションおよび音声コマンドトリガ用語を識別するデータを受信し、受信したデータを検証し、受信したデータを、アプリケーション、音声コマンドトリガ用語、および、少なくとも音声コマンドトリガ用語に基づいて判定される1つまたは複数の他の音声コマンドトリガ用語を指定するパッシブデータ構造を生成するように誘導し、文脈インテントデータベースにおいてパッシブデータ構造を記憶し、文脈インテントデータベースは、1つまたは複数の他のパッシブデータ構造を備える、ための方法、システム、および装置。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

音声アクションサービスシステムによって、ユーザのコンピューティングデバイスにおいて提供される発話を受信するステップであって、前記発話は音声コマンドトリガフレーズを含む、ステップと、

前記音声アクションサービスシステムによって、前記音声コマンドトリガフレーズに関連するインテントを判定するために前記発話を処理するステップと、

前記音声アクションサービスシステムによって、各々が前記インテントを満足することができる少なくとも第1のアプリケーションおよび第2のアプリケーションを識別するステップであって、少なくとも前記第1のアプリケーションおよび前記第2のアプリケーションを識別するステップは、少なくとも前記第1のアプリケーションおよび前記第2のアプリケーションが1つまたは複数のデータベース中の前記インテントに関連していると判定するステップに基づく、ステップと、

前記音声アクションサービスシステムによって、前記第2のアプリケーションよりも前記第1のアプリケーションを選択するステップであって、前記第2のアプリケーションよりも前記第1のアプリケーションを選択するステップは、前記ユーザによる前記第1のアプリケーションの最近の使用状況に少なくとも部分的に基づく、ステップと、

前記音声アクションサービスシステムによって、前記ユーザの前記コンピューティングデバイスに、前記発話に応じて、前記選択した第1のアプリケーションの指示を提供するステップと

を含む、コンピュータ実施方法。

【請求項 2】

前記ユーザによる前記第1のアプリケーションの最近の使用状況に少なくとも部分的に基づいて、前記第2のアプリケーションよりも前記第1のアプリケーションを選択するステップは、前記第1のアプリケーションが、前記ユーザによって、前記音声コマンドトリガフレーズに応じて、最近選択されていたと判定するステップを含む、請求項1に記載のコンピュータ実施方法。

【請求項 3】

前記第2のアプリケーションよりも前記第1のアプリケーションを選択するステップはさらに、前記第1のアプリケーションと前記音声コマンドトリガフレーズまたは前記インテントのうちの少なくとも1つとの間の関係スコアの強度に少なくとも部分的に基づく、請求項1に記載のコンピュータ実施方法。

【請求項 4】

前記第2のアプリケーションよりも前記第1のアプリケーションを選択するステップはさらに、前記発話を受信した際に前記ユーザの前記コンピューティングデバイス上で前記第1のアプリケーションが実行されていることに少なくとも部分的に基づく、請求項1に記載のコンピュータ実施方法。

【請求項 5】

前記選択した第1のアプリケーションの前記指示を提供するステップは、前記選択した第1のアプリケーションの聴覚的指示を提供するステップを含む、請求項1に記載のコンピュータ実施方法。

【請求項 6】

前記音声アクションサービスシステムによって、前記ユーザの前記コンピューティングデバイスにおいて追加の発話を受信するステップであって、前記追加の発話は前記選択した第1のアプリケーションの確認を含む、ステップと、

前記追加の発話を受信するステップに応じて、前記インテントを満足するために前記第1のアプリケーションを実行するステップと

をさらに含む、請求項1に記載のコンピュータ実施方法。

【請求項 7】

前記インテントを判定するために前記発話を処理するステップは、

10

20

30

40

50

前記音声アクションサービスシステムによって、前記発話のトランスクリプションを取得するために前記発話に対して音声認識を行うステップと、

前記音声アクションサービスシステムによって、前記トランスクリプションの少なくとも一部が前記音声コマンドトリガフレーズを含むとともに前記音声コマンドトリガフレーズが前記_intentと一致していると判定するステップと

を含む、請求項1に記載のコンピュータ実施方法。

【請求項8】

少なくとも1つのプロセッサと、

命令を含む少なくとも1つのメモリとを含み、前記命令は、実行されると、前記少なくとも1つのプロセッサに、

ユーザのコンピューティングデバイスにおいて提供される発話を受信することであって、前記発話は音声コマンドトリガフレーズを含む、ことと、

前記発話が前記音声コマンドトリガフレーズを含んでいると判定するために前記発話を処理することと、

少なくとも第1のアプリケーションおよび第2のアプリケーションを識別することであって、少なくとも前記第1のアプリケーションおよび前記第2のアプリケーションを識別することは、少なくとも前記第1のアプリケーションおよび前記第2のアプリケーションが1つまたは複数のデータベース中の前記音声コマンドトリガフレーズにマッピングされていると判定することに基づく、ことと、

前記第2のアプリケーションよりも前記第1のアプリケーションを選択することであって、前記第2のアプリケーションよりも前記第1のアプリケーションを選択することは、前記ユーザによる前記第1のアプリケーションの最近の使用状況に少なくとも部分的に基づく、ことと、

前記ユーザの前記コンピューティングデバイスに、前記発話に応じて、前記選択した第1のアプリケーションの指示を提供することと

をさせる、システム。

【請求項9】

前記ユーザによる前記第1のアプリケーションの最近の使用状況に少なくとも部分的に基づいて、前記第2のアプリケーションよりも前記第1のアプリケーションを選択するための前記命令は、前記第1のアプリケーションが、前記ユーザによって、前記音声コマンドトリガフレーズに応じて、最近選択されていたと判定するための命令を含む、請求項8に記載のシステム。

【請求項10】

前記第2のアプリケーションよりも前記第1のアプリケーションを選択するための前記命令は、前記第1のアプリケーションと前記音声コマンドトリガフレーズとの間の関係スコアの強度に少なくとも部分的に基づいて、前記第2のアプリケーションよりも前記第1のアプリケーションを選択するための命令をさらに含む、請求項8に記載のシステム。

【請求項11】

前記第2のアプリケーションよりも前記第1のアプリケーションを選択するための前記命令は、前記発話を受信した際に前記ユーザの前記コンピューティングデバイス上で前記第1のアプリケーションが実行されていることに少なくとも部分的に基づいて、前記第2のアプリケーションよりも前記第1のアプリケーションを選択するための命令をさらに含む、請求項8に記載のシステム。

【請求項12】

前記選択した第1のアプリケーションの前記指示を提供するための前記命令は、前記選択した第1のアプリケーションの聴覚的指示を提供するための命令をさらに含む、請求項8に記載のシステム。

【請求項13】

前記ユーザの前記コンピューティングデバイスにおいて追加の発話を受信することであって、前記追加の発話は前記選択した第1のアプリケーションの確認を含む、ことと、

10

20

30

40

50

前記追加の発話を受信することに応じて、前記第1のアプリケーションを実行すること

と
 をするための命令をさらに含む、請求項8に記載のシステム。

【請求項14】

命令を含むコンピュータ可読記憶媒体であって、前記命令は、実行されると、少なくとも1つのプロセッサに、

ユーザのコンピューティングデバイスにおいて提供される発話を受信することであって、前記発話は音声コマンドトリガフレーズを含む、ことと、

前記音声コマンドトリガフレーズに関連する_intentを判定するために前記発話を処理することと、

各々が前記_intentを満足することができる少なくとも第1のアプリケーションおよび第2のアプリケーションを識別することであって、少なくとも前記第1のアプリケーションおよび前記第2のアプリケーションを識別することは、少なくとも前記第1のアプリケーションおよび前記第2のアプリケーションが1つまたは複数のデータベース中の前記_intentに関連していると判定することに基づく、ことと、

前記第2のアプリケーションよりも前記第1のアプリケーションを選択することであって、前記第2のアプリケーションよりも前記第1のアプリケーションを選択することは、前記ユーザによる前記第1のアプリケーションの最近の使用状況に少なくとも部分的に基づく、ことと、

前記ユーザの前記コンピューティングデバイスに、前記発話に応じて、前記選択した第1のアプリケーションの指示を提供することと

をさせる、コンピュータ可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

この明細書は一般に、限定的ではなく、音声コマンドに関し、1つの特定の実施は、音声コマンドを設定するステップに関する。

【背景技術】

【0002】

ソフトウェアで定義されたアクティビティは、タスクを指定し得る。アクティビティは、タスクの実行のライフサイクルを制御するクラスである。intentは、特定のアクティビティ、および、特定のアクティビティに関連付けられたアプリケーションを指定し得るパッシブデータ構造である。intentは、アプリケーションによってトリガされ得、intentによって指定されたアクティビティが、intentによって指定されたターゲットアプリケーションにおいて、または、intentによって指定されたターゲットアプリケーションによって実行されるようにさせ得る。

【発明の概要】

【課題を解決するための手段】

【0003】

intentと称されるパッシブデータ構造は、アプリケーションと、アプリケーションにおいて、または、アプリケーションによって実行されるべきアクティビティとを指定し得る。アクティビティは、アプリケーションにおいて、または、アプリケーションによって実行されるタスクである。intentは、intentをトリガするために使用され得る1つまたは複数のトリガフレーズを指定し得、これによって、intentのトリガは、指定されたアプリケーションによるアクティビティの実行を引き起こす。たとえば、intentは、タクシーサービスのためのアプリケーションに関連付けられ得、ユーザが発話すると、トリガフレーズのうちの1つの検出が、タクシーサービスアプリケーションの起動をトリガできるように、音声アクションとして動作する1つまたは複数のトリガフレーズを指定し得る。intentは、アプリケーションによって、または、intentによって指定されたアプリケーション以外のオペレーティングシステムによってトリガされ得、こ

10

20

30

40

50

れによって、アプリケーションのトリガは、指定されたアクティビティが、特定のアプリケーションにおいて、または、特定のアプリケーションによって実行されるようにする。

【0004】

ホストによって提供されるサービスまたはツールは、アプリケーション開発者が、自分たちが開発している特定のアプリケーションを指定する新たなインテントを要求することを可能にすることができる。新たなインテントを求める発行は、文法の形式であり得、開発者は、サービスまたはツールへ文法を発行し得、サービスまたはツールは、文法に基づいて、特定のアプリケーションのためのインテントを生成し得る。

【0005】

新たなインテントの生成は、特定の結果が達成されるようにするためにアプリケーションへ提供され得る入力の数および範囲を増加させることによって、開発中の特定のアプリケーションの柔軟性を高め得る。結果的に、開発中のアプリケーションの信頼性が高められ得る。なぜなら、特定のアプリケーションは、ユーザの入力に対して正しい方式で応答することが保証され得、ユーザの入力の処理が向上されるからである。

10

【0006】

いくつかの実施では、開発者によって発行される文法は、アプリケーション、トリガフレーズ、アクティビティまたは複合アクティビティ、コンテキスト、または他の情報のうちの1つまたは複数指定し得る。文法は、開発者によって発行され、開発者音声アクションサービスまたはツールによって、ネットワークを介して受信される。開発者音声アクションサービスまたはツールは、文法を検証し得る。たとえば、文法の検証は、トリガフレーズが、オペレーティングシステムまたは別のアプリケーションのためのインテントに既に関連付けられているか否かを判定するステップを含み得、文法によって指定されたアクティビティが、指定されたアプリケーションによって実行され得るか否かを判定し得、トリガフレーズが、ある基準を満たすか否かを判定し得、発行された文法のフォーマットが、構文上有効であるか否かを判定し得るかまたは、さもなければ、開発者によって発行された文法が、新たなインテントを作成するために使用され得る実現可能な文法であるか否かを判定し得る。

20

【0007】

有効な文法は、文法をインテントへ変換するように、また、文法において指定されたトリガフレーズを拡張するように、開発者音声アクションサービスまたはツールによって誘導され得る。たとえば、文法の誘導は、第1のフォーマットで発行された文法を、第2のフォーマットのインテントへ変換するステップを含み得る。文法をインテントへ変換するステップに加えて、開発者によって発行されたトリガフレーズは、他の関連するトリガフレーズを含めるように拡張され得る。たとえば、インテントに関連付けられた1つまたは複数の関連するトリガフレーズを生成するために、1つまたは複数のトリガフレーズ拡張規則、類義語規則、用語オプション化規則、または他の規則が、トリガフレーズへ適用され得る。生成されたインテントは、おのおののアプリケーション、1つまたは複数のトリガフレーズ、およびオプションとして、アクティビティ、複合アクティビティ、文脈、または、インテントまたはインテントのトリガに関連する他の情報に関連付けられた1つまたは複数の他のインテントとともに、文脈インテントデータベースにおいて記憶され得る。

30

40

【0008】

文法の検証に基づく新たなインテントの生成は、開発者が、インテント自体の直接的な入力によって新たなインテントを生成することは必要ではないことを意味する。上述したように、文法は、誘導処理において変換され得る特定のフォーマットを有し得、この変換処理は、新たなインテントのための適切な候補であり得る文法の数の拡大を有効にする。なぜなら、入力文法フォーマットにおける要件が緩和されるからである。結果的に、アプリケーションは、より少ないユーザ入力で、新たなインテントを含めるように開発され得る。

【0009】

開発者によって発行された文法がインテントとして確立された後、ユーザは、クライア

50

ントデバイスへ音声入力を提供し得る。ユーザの音声入力は、音声入力のトランスクリプションを生成するように転記され得、トランスクリプションは、文脈インテントデータベース、または、クライアントデバイスのオペレーティングシステムに関連付けられたインテントを含むデータベースのように、インテントを指定する1つまたは複数の他のデータベースにおいて記憶されたインテントに関連付けられた1つまたは複数のトリガフレーズと照合され得る。トランスクリプションの一部が、特定のインテントに関連付けられたトリガフレーズに一致するとの判定に基づいて、インテントは、インテントに関連付けられたアクティビティが、特定のインテントによって指定されたアプリケーションにおいて、または、特定のインテントによって指定されたアプリケーションによって実行されるように、トリガされ得る。

10

【0010】

たとえば、タクシーサービスアプリケーションを起動させることに関連付けられたインテントは、「Call a cab」というトリガフレーズに関連付けられ得る。クライアントデバイスが、「Call a cab」というフレーズを含むユーザの音声入力を受信したに基づいて、クライアントデバイスまたは別のシステムは、タクシーサービスアプリケーションの起動をトリガし得る。

【0011】

この明細書において説明される主題の革新的な態様は、アプリケーションおよび音声コマンドトリガ用語を識別するデータを受信するステップと、受信したデータを検証するステップと、受信したデータを、アプリケーション、音声コマンドトリガ用語、および、少なくとも音声コマンドトリガ用語に基づいて判定される1つまたは複数の他の音声コマンドトリガ用語を指定するインテントを生成するように誘導するステップと、文脈インテントデータベースにおいてインテントを記憶するステップであって、文脈インテントデータベースは、1つまたは複数の他のインテントを備える、ステップと、からなるアクションを含む方法において具体化され得る。これら態様の他の実施形態は、コンピュータ記憶デバイスにおいてエンコードされた方法のアクションを実行するように構成された対応するシステム、装置、およびコンピュータプログラムを含む。

20

【0012】

これらおよび他の実施形態はおのおの、オプションとして、以下の特徴のうちの1つまたは複数を含み得る。様々な例において、受信したデータを検証するステップは、音声コマンドトリガ用語が、識別されたアプリケーションとは異なるアプリケーションに関連付けられていないことを判定するステップを備える。これらアクションは、アクションを識別するデータを受信するステップを備え、受信したデータを検証するステップは、識別されたアクションが、識別されたアプリケーションと互換的であることを判定するステップを備える。受信したデータを、アプリケーション、音声コマンドトリガ用語、および、少なくとも音声コマンドトリガ用語に基づいて判定される1つまたは複数の他の音声コマンドトリガ用語を指定するインテントを生成するように誘導するステップは、アプリケーションおよび音声コマンドトリガ用語を識別する、受信したデータを、異なるデータフォーマットへ変換するステップを備える。受信したデータを、アプリケーション、音声コマンドトリガ用語、および、少なくとも音声コマンドトリガ用語に基づいて判定される1つまたは複数の他の音声コマンドトリガ用語を指定するインテントを生成するように誘導するステップは、1つまたは複数の類義語規則、オプション化規則、または拡張規則を音声コマンドトリガ用語へ適用することによって、少なくとも音声コマンドトリガ用語に基づいて判定される1つまたは複数の他の音声コマンドトリガ用語を生成するステップを備える。

30

40

【0013】

これらアクションは、ユーザの発話を含むオーディオデータを取得するステップと、少なくともオーディオデータに基づいて、ユーザの発話のトランスクリプションを生成するステップと、トランスクリプションの少なくとも一部が、文脈インテントデータベースに記憶されたインテントによって指定された特定の音声コマンドトリガ用語に一致すること

50

を判定するステップであって、インテントは、特定のアプリケーションと、特定の音声コマンドトリガ用語とを指定する、ステップと、トランスクリプションの少なくとも一部が、インテントによって指定された特定の音声コマンドトリガ用語に一致することを判定するステップに基づいて、タスクが実行されるようにするステップとを備える。タスクが実行されるようにするステップは、特定のアプリケーションにタスクを実行させるステップを備える。これらアクションはさらに、おのものが(i)アプリケーションと、(ii)音声コマンドトリガ用語とを指定する1つまたは複数のインテントのおののについて、トランスクリプションの少なくとも一部が、インテントによって指定される音声コマンドトリガ用語と一致していることを判定するステップと、1つまたは複数のインテントのおののによって指定されるアプリケーションを含むアプリケーションのセットからの特定のアプリケーションの選択を要求するステップと、1つまたは複数のインテントのおののによって指定されるアプリケーションを含むアプリケーションのセットからの特定のアプリケーションの選択を示すデータを受信するステップと、特定のアプリケーションの選択を示すデータを受信するステップに応じて、特定の音声コマンドトリガ用語と、特定のアプリケーションを指定するインテントとの間の関係の強度を調節するステップと、を備え得る。

【0014】

これらアクションは、おのものが(i)アプリケーションと(ii)音声コマンドトリガ用語とを指定する1つまたは複数のインテントのおののについて、トランスクリプションの少なくとも一部が、インテントによって指定される音声コマンドトリガ用語と一致していることを判定するステップと、1つまたは複数のインテントのおののについて、インテントによって指定される音声コマンドトリガ用語の、インテントによって指定されるアプリケーションへの親和性を示すインテントに関する親和性尺度を判定するステップと、1つまたは複数のインテントのおののに関する親和性尺度に少なくとも基づいて、特定のインテントを選択するステップと、特定のインテントによって指定されるタスクが実行されるようにするステップと、を備える。これらアクションは、おのものが(i)アプリケーションと、(ii)音声コマンドトリガ用語とを指定する1つまたは複数のインテントのおののについて、トランスクリプションの少なくとも一部が、インテントによって指定される音声コマンドトリガ用語と一致していることを判定するステップと、1つまたは複数のインテントをランク付けするステップと、1つまたは複数のインテントのランク付けに少なくとも基づいて、特定のインテントを選択するステップと、特定のインテントによって指定されるタスクが実行されるようにするステップと、を備える。

【0015】

この明細書において説明される主題の1つまたは複数の実施態様の詳細は、添付図面および以下の詳細説明に記述される。主題の他の潜在的な特徴、態様、および利点は、詳細説明、図面、および特許請求の範囲から明らかになるであろう。

【図面の簡単な説明】

【0016】

【図1】開発者音声アクションシステムおよびサービスのための例示的なシステムを描写する図である。

【図2】音声アクションシステムおよびサービスを使用するための例示的なシステムを描写する図である。

【図3】開発者音声アクションシステムおよびサービスに関連付けられた例示的な処理のフローチャートである。

【発明を実施するための形態】

【0017】

様々な図面における同様の参照符号は、同様の要素を示す。

【0018】

図1は、開発者音声アクションサービスシステム100の例を示す。一般に、開発者音声アクションサービスは、アプリケーション開発者101が、開発者音声アクションサービスシ

10

20

30

40

50

ステム100によって、新たなインテントのための候補として取り扱われる文法を発行することを有効にする。

【0019】

開発者音声アクションサービスシステム100は、アプリケーション開発者101によって発行された文法を検証し、有効であれば、文法を、発行されたフォーマットから、インテントフォーマットへ変換し、文法によって指定されたトリガフレーズを、指定されたトリガフレーズに関連する追加のトリガフレーズを含むように拡張することによって、文法を誘導する。開発者音声アクションサービスシステム100は、様々なアプリケーションを指定する他のインテントを含む文脈インテントデータベース140に、この新たなインテントを記憶する。

10

【0020】

簡潔に述べると、開発者音声アクションサービスシステム100は、取得エンジン110、検証エンジン120、文法誘導エンジン130、および文脈インテントデータベース140を含む。取得エンジン110、検証エンジン120、文法誘導エンジン130、または文脈インテントデータベース140のいずれかは、別の音声アクションサービスシステムのような開発者音声アクションサービスシステム100と通信し得る他のシステムとは独立した開発者音声アクションサービスシステム100の構成要素として実施され得るか、または、これら他のシステムの構成要素として実施され得る。検証エンジン120は、検証基準データベース125において利用可能な検証基準へのアクセスを有し、文法誘導エンジン130は、誘導規則データベース135において利用可能な誘導規則へのアクセスを有する。アプリケーション開発者101は、1つまたは複数のネットワーク105を介して、開発者音声アクションサービスシステム100と通信しているコンピューティングデバイス102を使用して、開発者音声アクションサービスシステム100へ文法を発行し得る。

20

【0021】

アプリケーション開発者101は、特定のアプリケーションのための新たなインテントを要求するために、開発者音声アクションサービスシステム100へ文法を発行し得る。具体的には、アプリケーション開発者101は、1つまたは複数のネットワーク105を介して開発者音声アクションサービスシステム100と通信し得るコンピューティングデバイス102を使用して、文法を発行し得る。

【0022】

アプリケーション開発者101によって発行された文法は、新たなインテントの生成に関連する様々な情報を指定し得る。具体的には、文法は、アプリケーション開発者101が、新たなインテントを生成するために試みるために開発している特定のアプリケーションのような、特定のアプリケーションを指定し得る。あるいは、文法は、開発中のアプリケーションに関連する、または、開発中のアプリケーションと通信している特定のアプリケーションを指定する文法のような、アプリケーション開発者101が開発しているアプリケーションとは異なる特定のアプリケーションを指定し得る。さらに他の例において、文法は、特定の部分処理、スレッド、または、アプリケーション開発者101が開発中のアプリケーションまたは別のアプリケーションに関連付けられた他の区分のような、アプリケーションの特定の態様を指定し得る。アプリケーション開発者101は、それに加えて、文法におけるトリガフレーズを指定し得る。トリガフレーズは、ユーザによって提供された音声入力から検出された場合、インテントをトリガする1つまたは複数の用語を含む。いくつかの事例では、トリガフレーズは、1つまたは複数の引数を含み得る。引数は、ユーザの音声入力の1つまたは複数の用語によって置換され得るトリガフレーズにおけるプレースホルダの役割をする。たとえば、「私を\$の場所へ連れて行って下さい」で始まるトリガフレーズは、「私をワシントンDCの1425K通り北西へ連れて行って下さい」のようなユーザ音声入力、トリガフレーズに一致するものとして識別されることを有効にし得る。場所、日、日付、または時間、人、ユーザのクライアントデバイスに記憶された連絡先、曲、映画、または他のコンテンツアイテムのようなコンテンツのアイテム、または他の引数のような他の引数が検出され得る。

30

40

50

【 0 0 2 3 】

文法はまた、オプションとして、文法によって指定された特定のアプリケーションにおいて、または、文法によって指定された特定のアプリケーションによって実行されるタスクとして定義される特定のアクティビティをも指定し得る。それに加えて、いくつかの事例では、アプリケーション開発者101は、文法に関連付けられたコンテキストを指定し得る。コンテキストは、文法によって指定されたインテントのトリガのために存在していなければならない、または、存在する場合、トリガされたインテントの選択を修正しなければならない追加条件を定義する。いくつかの実施では、アプリケーション開発者101によって発行された文法は、開発者が定義したタイプを指定し得る。タイプは、トリガフレーズとともに含まれている場合に検出されねばならない1つまたは複数の用語として定義される。たとえば、1つまたは複数の用語によって識別された開発者が定義したタイプは、たとえば、「スリープ」または「ロック」のように、アプリケーション指定モードの特定のタイプに関連付けられ得るか、または、たとえば、中古車探索アプリケーションのための車モデル名のように、特定のアプリケーションへ提供される入力と連携して頻繁に使用される特定の用語によって識別され得るかまたは、あるアプリケーションまたはトリガフレーズに関連する開発者が定義したタイプであり得る。音声入力がユーザから受信された場合、音声アクションサービスシステムは、他の用語以外のタイプの用語を検出する可能性がより高くなり得る。たとえば、音声アクションサービスシステムは、アプリケーション開発者によって発行された1つまたは複数の文法によって定義されたタイプである「スリープ」に基づいて、「スリップ」という用語とは異なる音声入力から、「スリープ」という用語を検出する可能性がより高くなり得る。

10

20

【 0 0 2 4 】

いくつかの例において、タイプは、特定のエンティティに関連付けられ得るか、または、特定のエンティティを参照するように動作し得る。たとえば、「私の家」という用語によって定義されるタイプは、たとえば、ユーザが、プロパティに関連付けられたアドレスを、アプリケーションである音声アクションサービスシステムに登録することに基づいて、または、「私の家」という用語によって定義されたタイプが、ユーザのプロパティ、または、プロパティのアドレスに関連付けられるべきであることを、アプリケーション開発者101が文法において指定することに基づいて、ユーザに関連付けられたプロパティに関連付けられ得る。ユーザ音声入力におけるタイプを定義する用語を検出することは、トリガされたインテントによって指定されたアクティビティが、タイプに関連付けられたエンティティに関して実行されるようにし得る。

30

【 0 0 2 5 】

たとえば、ユーザが「私の家にタクシーを寄こして下さい」という音声入力を、音声アクションサービスシステムへ提供することに基づいて、音声アクションサービスシステムは、音声入力が、タクシーサービスアプリケーションのためのインテントによって指定されたトリガフレーズに一致していることを判定し得、さらに、音声入力に含まれる「私の家」という用語が、ユーザの自宅の住所に関連付けられたタイプに対応することを判定し得る。したがって、インテントのトリガは、タクシーサービスアプリケーションに対して、タクシーに、ユーザを、現在の場所から、自宅住所へ送るように要求させ得る。

40

【 0 0 2 6 】

そのような例では、ユーザの自宅住所を指定する情報が、ユーザのクライアントデバイスにおいて実行しているタクシーサービスアプリケーションの事例に関連付けて記憶され得、開発者音声アクションサービスシステム100において記憶され得、文脈インテントデータベース140において記憶されたインテントによって指定され得るか、または、別の方式で記憶され得る。いくつかの例では、タイプに関連付けられたエンティティは、特定のユーザに対して特定であり、これによって、「私の家」という用語は、異なるユーザに関連付けられた異なるプロパティのアドレスを参照するようになるであろう。そして、タイプに関連付けられた正しいエンティティが、タイプを含む音声入力を提供するユーザのために選択され得る。

50

【 0 0 2 7 】

例として、アプリケーション開発者101は、図1に図示される文法を発行し得る。文法は、タクシーサービスアプリケーションである「Cab Caller」と呼ばれるアプリケーションを指定し得る。それに加えて、文法は、アプリケーション開発者101が、文法から生成されるインテントのトリガを引き起こすべきトリガフレーズとして示した「Hail a cab」というトリガフレーズを指定し得る。文法はまた、特定のアクティビティ、すなわち、指定されたアプリケーションの起動を指定する。したがって、文法から生成されたインテントのトリガは、起動されている「Cab Caller」というアプリケーションに帰着するであろう。繰り返すが、文法から生成されるインテントのトリガは、サードパーティのアプリケーションを起動することに限定されないが、他のアプリケーションまたは機能を、開発者によって指定され得るようにトリガし得る。アプリケーション開発者101は、それに加えて、文法のためのコンテキストを指定した。具体的には、コンテキストは、文法から生成されたインテントが、インテントをトリガする音声入力を提供するクライアントデバイスが、「Cab Caller」というアプリケーションの「バージョン1.2」をホストする場合にのみトリガされるべきであることを示す。日時、場所、デバイス方位、天候、ユーザ認証、アイテムへの近接度またはアイテムを搬送する人、ユーザのクライアントデバイス上の特定のソフトウェアまたは構成要素の存在(または不在)等のような他の多くのコンテキストが、この開示によって考慮される。

10

【 0 0 2 8 】

アプリケーション開発者101によって発行された文法は、取得エンジン110において、1つまたは複数のネットワーク105を介して開発者音声アクションサービスシステム100によって受信される。取得エンジン110は、開発者音声アクションサービスシステム100が、アプリケーション開発者によって発行された文法を受信することを有効にするために、コンピューティングデバイス102のようなコンピューティングデバイスのためのインターフェースを提供する。

20

【 0 0 2 9 】

取得エンジン110は、アプリケーション開発者101によって発行された文法を受信し、この文法を、検証エンジン120へ提供する。検証エンジン120は文法を受信し、新たなインテントが生成され得るか否かを、発行された文法に基づいて判定するために、文法を検証する。文法を検証するために、検証エンジン120は、検証基準データベース125においてアクセス可能な検証基準にアクセスし得る。検証基準データベース125は、あらかじめ決定されたか、または、受信した文法に応じて生成された検証基準を含み得る。

30

【 0 0 3 0 】

たとえば、検証エンジン120は、文法によって指定されたトリガフレーズが既にインテントに関連付けられているか否かを判定することによって、文法を検証し得る。それに加えて、または、その代わりに、文法を検証するステップは、文法によって指定されたアプリケーションが、文法によって指定された特定のアクティビティ、または、別のデフォルトアクティビティを実行することができるかを判定するステップを含み得る。たとえば、検証エンジン120は、どのアクティビティが、アプリケーションにおいて、または、アプリケーションによって実行され得、どのデータが、アプリケーションへ、または、アプリケーション内で利用可能であることを示す情報、または他の情報のような、アプリケーションに関する情報を示す文法において指定されたアプリケーションに関連付けられたマニフェストまたは他の情報へのアクセスを有し得る。検証エンジン120は、文法を検証するために、識別されたアプリケーションに関連付けられたマニフェストまたは他の情報を調査し得る。

40

【 0 0 3 1 】

文法を検証するステップはさらに、文法において指定されたコンテキストが、適切なコンテキストであるか否かを、たとえば、コンテキストが満足され得ること、たとえば、「Cab Caller」というアプリケーションの「バージョン1.2」が存在することを確認することによって判定するステップを含み得る。文法を検証するステップはさらに、文法がイン

50

テントのフォーマットへ変換され得るように、適切なフォーマットで提供されていることを判定するステップをも含み得る。それに加えて、または、その代わりに、文法を検証するステップは、文法のために発行されたトリガフレーズが、適切なフォーマットであることを、たとえば、トリガフレーズに含まれる引数が正しくフォーマットされていること、または、サポートされた引数タイプであることを判定することによって判定するステップを含み得る。文法を検証するステップはまた、アプリケーション開発者101から受信した文法において指定されるアクティビティを実行することに関連しているか、または実行するように要求された1つまたは複数のアクティビティを識別するステップを含み得る。たとえば、アプリケーション開発者101が、タクシーサービスアプリケーションを使用して、タクシーを頼むためのアクティビティを文法において指定しているのであれば、検証エンジン120は、タクシーアプリケーションサービスを使用してタクシーを頼むステップはまた、ユーザのクライアントデバイスにおいてタクシーサービスアプリケーションを起動するステップと、ユーザのクライアントデバイスにおいて実行している位置決めシステムアプリケーションを使用してユーザの位置を決定するステップとからなるアクティビティを含んでいると判定し得る。

10

20

30

40

50

【0032】

それに加えて、検証エンジン120は、文法のフォーマットまたは文法の一部が、構文上有効であるか否かを判定することによって、受信した文法を検証し得る。いくつかの事例では、文法誘導エンジン130は、受信した文法が構文上有効であれば、すなわち、そのデータ構造が、期待される文法フォーマットに従って適切にフォーマットされており、文法誘導エンジン130によって処理または変換され得る用語のみを含んでいる場合にのみ、受信した文法を誘導でき得る。

【0033】

いくつかの実施では、開発者音声アクションサービスシステム100によって、アプリケーション開発者101へフィードバックが提供され得る。たとえば、検証エンジン120は、アプリケーション開発者101のコンピューティングデバイス102へ、アプリケーション開発者101によって発行された文法が検証されたか否かを示す情報を提供し得る。インテントが文法を使用して成功裡に生成されているか否かを示す情報、文法において指定されたトリガフレーズに基づいて生成された1つまたは複数の他の関連するトリガフレーズを指定する情報、または他の情報のような、追加のまたは異なるフィードバックもまたアプリケーション開発者101へ提供され得る。

【0034】

検証エンジン120が、アプリケーション開発者101によって発行された文法を検証することに基づいて、検証エンジン120は、文法を、文法誘導エンジン130へ発行し得る。文法誘導エンジン130は、文法を受信し、文法を、文法に基づくインテントを生成するように誘導し得る。文法を誘導するステップは、文法を、開発者音声アクションサービスシステム100によって受信されたフォーマットから、インテントフォーマットへ変換するステップを含み得る。文法を誘導するステップはさらに、文法において指定されたトリガフレーズに基づいて1つまたは複数の他のトリガフレーズを生成することによって文法において指定され、かつ、指定されたトリガフレーズに関連しているトリガフレーズを拡張するステップを含み得る。

【0035】

たとえば、開発者音声アクションサービスシステム100は、アプリケーション開発者が、アプリケーション開発者によってより容易に理解されるフォーマットで文法を、たとえば、アプリケーション開発者が、テキスト形式引数を含むフォーマットで文法を発行することを有効にすることによって、他のユーザへ発行することを有効にし得る。文法誘導エンジン130は、このフォーマットで文法を受信し得、この文法を、インテントのためのフォーマットへ変換し得る。インテントフォーマットは、アプリケーション開発者または他のユーザにとって、解釈または理解がより困難なフォーマットであり得るか、または、音声アクションサービスに特有のフォーマットであり得る。

【0036】

開発者音声アクションサービスシステム100へ発行された文法において指定されるトリガフレーズを拡張することによって、開発者音声アクションサービスシステム100が、指定されたトリガフレーズに基づいている1つまたは複数の関連するトリガフレーズを生成することを可能とする。関連するトリガフレーズを生成するために、様々な拡張規則が、トリガフレーズへ適用され得る。たとえば、文法において指定されたトリガフレーズに関連し、文法において指定されたトリガフレーズに少なくとも部分的に基づいている1つまたは複数の他のトリガフレーズを生成するために、1つまたは複数の用語オプション化規則、類義語規則、拡張規則、または、他の問合せ用語改訂規則のような他の改訂規則が、文法において指定されたトリガフレーズへ適用され得る。

10

【0037】

オプション化規則は、トリガフレーズにおける1つまたは複数の特定の用語が、文法において指定されたトリガフレーズに基づいて生成された関連するトリガフレーズにおいてオプションであるか否か指定するトリガフレーズへ適用され得る。類義語規則は、文法において指定されたトリガフレーズに基づいて生成された関連するトリガフレーズに含まれる文法の1つまたは複数の用語のための1つまたは複数の類義語を識別し得る。拡張規則は、文法において指定されたトリガフレーズが、文法において指定されたトリガフレーズにおける用語に類似または関連しているとして識別される1つまたは複数の用語のような、1つまたは複数の追加の規則を含むように拡張されることを有効にし得る。文法において指定されるトリガフレーズに基づいて、関連するトリガフレーズを生成するための他の規則が適用され得る。

20

【0038】

それに加えて、いくつかの例では、文法誘導エンジン130は、受信したトリガフレーズ、または、指定されたトリガフレーズのために生成された1つまたは複数の関連するトリガフレーズを1つまたは複数の他の言語へ変換することによって、受信した文法において指定されるトリガフレーズに関連する1つまたは複数のトリガフレーズを生成し得る。たとえば、「Hail a cab」というトリガフレーズが、英語から、標準中国語、スペイン語、アラビア語、フランス語、ヒンディー語、ドイツ語、ロシア語、または他の言語のうちの1つまたは複数へ変換され得、1つまたは複数の変換は、文法において指定された「Hail a cab」というトリガフレーズに関連するトリガフレーズとして識別され得る。

30

【0039】

それに加えて、いくつかの事例では、文法誘導エンジン130は、受信した文法において指定されたトリガフレーズに含まれる用語または引数の変形および置換を判定することができ得る。たとえば、買物アプリケーションに関連付けられた文法において指定されるトリガフレーズが、たとえば米国サイズにおけるスーツのサイズのように、第1の単位における洋服サイズパラメータのための引数を指定するのであれば、文法誘導エンジン130は、たとえば欧州サイズにおけるスーツのサイズのように、第2の単位における洋服サイズパラメータのための引数を、別のトリガフレーズの中にも含むことによって、この引数の変形または置換を判定することができ得る。

40

【0040】

いくつかの事例では、文法誘導エンジン130は、関係スコアの強度を判定し得、関係スコアの強度を、文法において指定されたトリガフレーズ、および文法において指定されたトリガフレーズに基づいて判定されたトリガフレーズのうちの1つまたは複数へ関連付け得る。関係スコアの強度は、文法において指定されたアプリケーションに対する、トリガフレーズの適用可能性または独自性の推定値を示し得る。たとえば、アプリケーション開発者101から受信された、「Hail a cab」というトリガフレーズを指定する文法に基づいて、文法誘導エンジン130は、「Hail a cab」というトリガフレーズと、関連するトリガフレーズとして識別された「Call a taxi」というトリガフレーズとの間の関係の強度を、受信した文法において指定されたトリガフレーズに少なくとも基づいて判定し得る。

【0041】

50

いくつかの事例では、検証エンジン120および/または文法誘導エンジン130は、要求された文法がスパムであるか否かを判定するために、トリガフレーズとアプリケーションとの間の関係の強度を判定し得る。たとえば、文法誘導エンジン130は、「Hail a cab」というトリガフレーズと、「Cab Caller」というタクシーサービスアプリケーションとの間の関係の強度、または、タクシー呼び出しアプリケーションを指定するインテントを判定し得、この判定に基づいて、関係の強度のスコアを、「Hail a cab」というトリガフレーズへ割り当て得る。それに加えて、いくつかの例では、広告主またはアプリケーション開発者は、推奨を試みているアプリケーション、および、トリガフレーズを指定する開発者音声アクションサービスへ文法を発行するように促され得る。検証エンジン120および/または文法誘導エンジン130は、トリガフレーズが、識別されたアプリケーションに対して十分に強い関係を有していないことを判定し得、したがって、受信した文法に基づいて新たなインテントを生成することを拒否し得、および/または、アプリケーションへより強く関連しているトリガフレーズを提供しなければならないことを示すフィードバックを、文法を発行したユーザへ提供し得る。

10

20

30

40

50

【0042】

たとえば、スポーツニュースアプリケーションのための広告主は、スポーツニュースアプリケーションと「Call a taxi」というトリガフレーズとを指定する文法を発行し得る。検証エンジン120および/または文法誘導エンジン130は、「Call a taxi」というトリガフレーズが、スポーツニュースアプリケーションへ十分強い関係を有していないことを判定し得、この文法に基づいて新たなインテントが生成されることを拒否し得る。

【0043】

いくつかの実施では、文法誘導エンジン130はまた、コンピューティングデバイス102において、アプリケーション開発者101へフィードバックを提供し得る。たとえば、受信した文法を、アプリケーション開発者101によって発行された文法によって識別されたトリガフレーズに少なくとも基づいて判定された1つまたは複数の他のトリガフレーズを生成するように誘導することに基づいて、アプリケーション開発者101が、他のトリガフレーズが、識別されたトリガフレーズのための代替トリガフレーズを受け入れ可能であるか否かを判定するために、1つまたは複数の他のトリガフレーズを評価することを可能にするために、1つまたは複数の他のトリガフレーズを識別するフィードバックが、アプリケーション開発者101へ提示され得る。いくつかの実施では、アプリケーション開発者101は、たとえば、文法誘導エンジン130によって生成された1つまたは複数の他のトリガフレーズのうちのどれが、新たなインテントにおいて指定されるべきであることを示すために、フィードバックへ応答を提供することができ得る。

【0044】

それに加えて、いくつかの事例では、文法誘導エンジン130または開発者音声アクションサービスシステム100の別の構成要素は、受信したインテントに基づく1つまたは複数の例示的な問合せまたは音声入力、受信した文法によって指定されたトリガフレーズに基づいて判定された1つまたは複数の他のトリガフレーズ、または、受信したインテントから生成されたインテントを生成し得る。例示的な問合せまたは音声入力は、文脈インテントデータベース140において記憶されたインテントへのアクセスを有する音声アクションサービスをどのように使用するかをユーザへ教示するための例として使用され得る。たとえば、例示的な問合せまたはトリガフレーズは、アプリケーション開発者101によって発行された文法から生成されたインテントをトリガする音声入力をユーザが最初に提供した場合、ユーザへの出力のために提供され得る。

【0045】

文法誘導エンジン130が、1つまたは複数の関連するトリガフレーズを生成し、1つまたは複数の関連するトリガフレーズを含む文法をインテントフォーマットへ変換することによって、文法において指定されたトリガフレーズを拡張したことに基づいて、開発者音声アクションサービスシステム100は、文脈インテントデータベース140において新たなインテントを記憶し得る。たとえば、アプリケーション開発者101によって発行された文法に

基づいて生成された新たなインテントは、少なくとも「インテントX」および「インテントZ」をも含む文脈インテントデータベース140において記憶されたインテントのセットへ、「インテントY」として追加され得る。たとえば、「インテントY」および「インテントZ」は、アプリケーション開発者101によって発行された文法において指定されたアプリケーションとは異なり得る、または、同じであり得る、1つまたは複数のアプリケーションに関連付けられ得る1つまたは複数の他のアプリケーション開発者によって発行された文法に基づいて生成されたインテントであり得る。

【0046】

文脈インテントデータベースにおいて記憶されたインテントは、インテントフォーマットでアプリケーション開発者101によって発行された文法によって指定されるパラメータを指定し得る。たとえば、アプリケーション開発者101によって発行された文法に基づいて生成された「インテントY」は、「Cab Caller」というアプリケーションを指定し得、インテントがトリガされた場合、「Cab Caller」というアプリケーションにおいて、または「Cab Caller」というアプリケーションによって実行されるべき起動アプリケーションアクティビティを指定し得、インテントのトリガに対して追加の条件、すなわち、インテントがトリガされ得るクライアントデバイスが、「Cab Caller」というアプリケーションの「バージョン1.2」を特別に扱う必要があること、を課すコンテキストを指定し得る。それに加えて、「インテントY」は、アプリケーション開発者101によって発行されたトリガフレーズのみならず、アプリケーション開発者101によって発行されたトリガフレーズに少なくとも基づいて、開発者音声アクションサービスシステム100が判定した1つまたは複数の追加のトリガフレーズをも指定し得る。たとえば、「Hail a cab」というトリガフレーズは、「Call a cab」、「Get a car」、および「Call taxi」というトリガフレーズをも含むように拡張され得、それらのいずれかの検出は、追加のコンテキストが満たされている限り、「インテントY」のトリガに帰着するであろう。

【0047】

一旦、インテントが、アプリケーション開発者101によって発行された文法に基づいて生成され、文脈インテントデータベース140において記憶されると、ユーザは、インテントに関連付けられたトリガフレーズを含む音声入力を提供することによって、インテントをトリガすることが可能であり得る。たとえば、クライアントデバイスへ「Call taxi」という音声入力を提供する「Cab Caller」というアプリケーションの「バージョン1.2」をホストするクライアントデバイスのユーザに基づいて、クライアントデバイスは、音声入力が「Call taxi」というトリガフレーズを含んでいることを判定し得、モバイルデバイスが「Cab Caller」というアプリケーションの「バージョン1.2」をホストしていることを判定し得、それに応じて、「Cab Caller」というアプリケーションを、ユーザのクライアントデバイス上で起動させる「インテントY」をトリガし得る。

【0048】

いくつかの実施では、文脈インテントデータベース140において新たに生成されたインテントを記憶する処理は、迅速な更新プッシュまたは展開サポートでもサポートされている処理であり得る。たとえば、音声開発者音声アクションサービスシステム100は、受信した文法を一旦誘導すると、新たに生成されたインテントを、文脈インテントデータベース140において迅速に記憶すること、および、新たに生成されたインテントを、文脈インテントデータベース140に記憶されたインテントへのアクセスを有する音声アクションサービスのユーザへ利用可能とすることが可能であり得る。

【0049】

図2は、ユーザが、文脈インテントデータベース240によって記憶されたインテントへ一致された音声入力を発行することを有効にする音声アクションサービスシステム200の例示的な使用の場合を示す。音声入力をインテントへ一致させることは、インテントによって指定されたアクティビティが、ユーザのクライアントデバイス202におけるアプリケーションにおいて、または、ユーザのクライアントデバイス202におけるアプリケーションによって実行されるようにする。アプリケーションはまた、インテントによって指定され

る。

【0050】

簡潔に述べると、図2に図示されるように、音声アクションサービスシステム200は、音声認識エンジン210、マッチャ220、明確化エンジン230、および実行エンジン250を含む。音声認識エンジン210は、クライアントデバイス202のユーザ201からの音声入力を含むオーディオデータを、1つまたは複数のワイヤまたはワイヤレスネットワーク205を介して受信する。たとえば、ユーザ201は、クライアントデバイス202にホストされた音声インターフェースアプリケーションへ音声入力を提供し得、クライアントデバイス202にホストされた音声インターフェースアプリケーションは、音声入力を含むオーディオデータを、1つまたは複数のネットワーク205を介して、音声アクションサービスシステム200へ送信し得る。図2においてクライアントデバイス202から分離しているとして図示されているが、いくつかの実施形態では、音声認識エンジン210、マッチャ220、明確化エンジン230、および実行エンジン250のうちの1つまたは複数は、本明細書において議論されている他のシステムおよびサブシステムと同様に代替的に実施され得るか、または、ユーザ201に関連付けられたクライアントデバイス202の一部を形成し得る。

10

【0051】

いくつかの事例では、音声認識エンジン210は、音声アクションサービスシステム200によって定義される1つまたは複数のタイプへのアクセスを有し得る。たとえば、音声アクションサービスシステム200は、人々、場所、映画または曲のようなコンテンツのアイテム、または他の定義されたタイプのようなエンティティを識別する1つまたは複数の知識ベースへのアクセスを有し得る。オーディオデータ204において音声認識を実行する場合、または、ユーザの音声入力に基づいて、インテントの選択および実行に関連する他の動作を実行する場合、音声認識エンジン210、または、音声アクションサービスシステム200の他の構成要素は、1つまたは複数の知識ベースへアクセスし得る。

20

【0052】

マッチャ220は、図1のアプリケーション開発者101のようなアプリケーション開発者によって発行された文法に基づいて生成された1つまたは複数のインテントを含む文脈インテントデータベース240へのアクセスを有する。マッチャ220は、それに加えて、1つまたは複数の特有のオペレーティングシステムに特有か、または、特定のアプリケーションに関するとしてオペレーティングシステムによって指定された、1つまたは複数のインテントを含むオペレーティングシステム(OS)インテントデータベース245へのアクセスを有する。たとえば、OSインテントデータベース245において記憶されたインテントは、デバイスをリスタートまたはロックする、ワイヤレスネットワークへ参加する等のためのアクティビティのような、オペレーティングシステムによって実行されるアクティビティ、または、メッセージング、電話、または電子メールアプリケーションを起動するためのアクティビティのような、1つまたは複数のアプリケーションに関するオペレーティングシステムによって実行されるアクティビティのためのインテントを含み得る。それに加えて、明確化エンジン230は、特定のユーザまたはクライアントデバイスに特有の情報を含むユーザパーソナル化データベース235へのアクセスを有する。ユーザパーソナル化データは、クライアントデバイス202から受信した音声入力に一致すると判定されたインテントを明確化するために使用され得る。

30

40

【0053】

図2に図示されるように、ユーザ201は、1つまたは複数のアクションを実行するためにユーザ201によって提供される音声コマンドのような音声入力を受信することが可能なクライアントデバイス202を有し得る。いくつかの事例では、クライアントデバイス202は、マイクロホン、または、ユーザから音声入力を取得するための他のデバイスを含み、クライアントデバイス202において実行するオペレーティングシステムまたはアプリケーションは、音声入力を受信するためのソフトウェアを含み得る。クライアントデバイス202は、「Call taxi」という音声入力のような音声入力を、ユーザ201から受信し得る。クライアントデバイス202が音声入力を受信することに基づいて、クライアントデバイスは、音

50

声入力を含むオーディオデータ204を、1つまたは複数のネットワーク205を介して、音声アクションサービスシステム200へ送信し得る。

【0054】

音声アクションサービスシステム200の音声認識エンジン210は、オーディオデータ204を受信し、オーディオデータ204について音声認識処理を実行し、オーディオデータ204に含まれる音声入力のトランスクリプションを取得する。たとえば、音声認識エンジン210は、オーディオデータ204を受信し得、オプションとして、オーディオデータ204をフィルタし、オーディオデータ204からノイズおよびバックグラウンドオーディオを除去し、オーディオデータ204のクリーンバージョンを生成し得る。音声認識エンジン210は、その後、音声入力のトランスクリプションを取得するために、オーディオデータ204のクリーンバージョンに対して音声認識処理を実行し得る。たとえば、音声認識エンジン210は、クライアントデバイス202から受信したオーディオデータ204から、「Call taxi」というトランスクリプションを生成し得る。

10

【0055】

音声認識エンジン210は、オーディオデータ204のトランスクリプションを、マッチャ220へ提供する。マッチャ220は、このトランスクリプションを、文脈インテントデータベース240またはOSインテントデータベース245において記憶されたインテントによって指定された1つまたは複数のトリガフレーズと照合する。いくつかの事例では、トランスクリプションが、特定のインテントに関連付けられた特定のトリガフレーズと一致すると判定するステップは、トランスクリプションの1つまたは複数の用語が、特定のトリガフレーズの1つまたは複数の用語と一致すると判定することのように、トランスクリプションの少なくとも一部が、特定のトリガフレーズと一致すると判定するステップを含み得る。したがって、特定のトランスクリプションは、潜在的に、多数のインテントを求めることとして判定され得、これによって、音声アクションサービスシステム200は、候補インテントから、トリガするための特定のインテントを判定しなければならない。トランスクリプションが、1つまたは複数の候補インテントによって指定された1つまたは複数のトリガフレーズに一致すると判定することに基づいて、マッチャ220は、1つまたは複数の候補トリガフレーズを指定する明確化エンジン230へ情報を提供し得る。

20

【0056】

いくつかの事例では、ユーザ201は、クライアントデバイス202へテキスト形式入力を提供し得、これによって、テキスト形式入力は、音声認識エンジン210へ提供されることなく、マッチャ220へダイレクトに提供されるようになり得る。マッチャ220は、テキスト形式入力を受信し得、テキスト形式入力に基づいて1つまたは複数の候補インテントを識別するために、上述されたマッチング動作を実行し得る。いくつかの事例では、マッチャ220は、テキスト形式入力に含まれる引数のようなタイプまたは他の特徴を識別することが可能であり得る。テキスト形式入力に含まれる。

30

【0057】

いくつかの例では、マッチャ220は、オーディオデータ204の1つまたは複数の関連するトランスクリプションを生成するために、オーディオデータ204のトランスクリプションを拡張し得る。たとえば、「Call taxi」というトランスクリプションは、1つまたは複数のクエリ用語拡張規則、類義語規則、用語オプション化規則、または、「Call a taxi」、「Request a cab」、「Hail a car」等の関連するトランスクリプションを生成するために他の規則を使用して拡張され得る。マッチャ220は、トランスクリプションまたは関連するトランスクリプションが一致するトリガフレーズを判定するために、関連するトランスクリプションのうちの1つまたは複数、文脈インテントデータベース240またはOSインテントデータベース245において記憶されたインテントによって指定されたトリガフレーズへ一致させ得る。

40

【0058】

いくつかの事例では、オーディオデータ204のトランスクリプションを拡張することは、マッチャ220に対して、関連するトランスクリプションが、既に存在する、または、異

50

なるアプリケーションに関連付けられているトリガフレーズとコンフリクトするか否かを判定することを要求する。たとえば、「Call taxi」というトリガフレーズは、タクシーサービスアプリケーションである「Cab Caller」に関連付けられたインテントによって指定され得、「Request a cab」という関連するトリガフレーズは、「TaxiNow」というタクシーサービスアプリケーションに関連付けられたインテントによって指定され得る。マッチャ220は、トランスクリプションおよび関連するトランスクリプションが、異なるアプリケーションに関連するインテントによって指定されたトリガフレーズに一致することを判定し得、したがって、いくつかの例では、一致処理から、関連するトランスクリプションである「Request a cab」を無視または削除し得る。したがって、マッチャ220、または、音声アクションサービスシステム200の別の構成要素は、一致が検出される機会を増加させるために、ユーザ入力のトランスクリプションを調節することが可能であり得る一方、オーディオデータ204の拡張されたトランスクリプションによって生成される任意のコンフリクトも取り扱う。

10

20

30

40

50

【0059】

明確化エンジン230は、1つまたは複数の候補インテントを指定する情報を受信し得、候補インテントから、コールをトリガする特定のインテントを識別するための動作を実行し得る。明確化エンジン230は、ユーザパーソナリゼーションデータベース235においてアクセスされた情報に基づいて、候補インテントのおおのに関連付けられた信頼度スコアに基づいて、トリガするための特定の候補インテントを選択するユーザフィードバックを要求または受信することに基づいて、または他の情報に基づいて、候補インテントからトリガするための特定のインテントを識別し得る。

【0060】

たとえば、音声アクションサービスシステム200は、たとえば、ユーザ201が、クライアントデバイス202によってホストされている特定のアプリケーションへログインされることに基づいて、クライアントデバイス202が、ユーザ201またはクライアントデバイス202を識別する情報を、音声アクションサービスシステム200へ発行することに基づいて、ユーザ201、または、ユーザ201によって使用されているクライアントデバイス202を識別する情報を受信し得るか、または、受信可能であり得る。明確化エンジン230は、この識別情報に基づいて、ユーザ201またはクライアントデバイス202を識別し得、特定のユーザ201またはクライアントデバイス202のためのインテントに関連する情報を指定する情報に、ユーザパーソナリゼーションデータベース235においてアクセスし得る。たとえば、ユーザパーソナリゼーションデータベース235において明確化エンジン230によってアクセスされる情報は、ユーザ201は典型的には、特定のアプリケーションを開くことのように、特定のインテントをトリガするために、オーディオデータ204または特定のテキスト形式入力に含まれる音声入力を提供することを指定し得る。同様に、ユーザパーソナリゼーションデータベース235においてアクセスされる情報は、ユーザ201によって提供された、または、ユーザ201によって、過去のアクティビティに基づいて判定された嗜好を指定し得る。たとえば、音声アクションサービスシステム200は、過去のユーザアクティビティに基づいて、ユーザ201は、特定のタクシーサービスアプリケーションを別のものよりも好む可能性が高いと判定し得るか、または、ユーザ201は特定のアプリケーションを参照する場合、典型的に特定の用語を使用することを判定し得る。

【0061】

ユーザパーソナリゼーションデータベース235においてアクセスされた情報に基づいて、明確化エンジン230は、候補インテントから特定のインテントを選択し得、特定のインテントを識別する実行エンジン250へ情報を提供し得る。たとえば、明確化エンジン230は、「Call taxi」というトランスクリプションに基づいて識別された2つの候補インテントを識別する情報を受信し得る。1つの候補インテントは「Cab Called」というタクシーサービスアプリケーションを指定し得、第2のインテントは、「TaxiNow」と呼ばれるタクシーサービスアプリケーションを指定し得る。明確化エンジン230は、ユーザパーソナリゼーションデータベース235において情報へアクセスし得、ユーザ201またはクライアントデバイス202のユー

ザは典型的には、「TaxiNow」というタクシーサービスアプリケーションよりも「Cab Caller」というタクシーサービスアプリケーションを使用すると判定し得、よって、2つの候補_intentから「Cab Caller」というアプリケーションを指定する候補_intentを選択し得る。その後、明確化エンジン230は、「Cab Caller」というアプリケーションを指定する_intentを識別する情報を、実行エンジン250へ送信し得る。このように、明確化エンジン230は、ユーザ201からの追加の入力を必要とすることなく、特定の候補_intentを選択するために、ユーザ201のためのパーソナリゼーションデータに依存し得る。

【0062】

あるいは、明確化エンジン230は、候補_intentのおのおのに関連付けられた信頼度スコアに基づいて、候補_intentのセットから、特定の_intentを識別し得る。たとえば、候補_intentのおのおのは、候補_intentが、ユーザ201が音声アクションを使用してトリガすることを意図していた_intentである信頼性を近似する信頼度スコアに関連付けられ得る。いくつかの事例では、明確化エンジン230またはマッチャ220は、候補_intentのおのおののための信頼度スコアを判定し得る。

10

【0063】

候補_intentへ割り当てられる信頼度スコアは、1つまたは複数の要因に基づき得る。たとえば、信頼度スコアは、トランスクリプションと、intentによって指定されたトリガフレーズとの間の一致の正確性に少なくとも部分的に基づいて判定され得る。そのような例では、トランスクリプションに正確に一致するトリガフレーズを指定する候補_intentは、候補_intentがユーザ201の意図であったとの高い信頼度を示す候補スコアを割り当てられ得る。

20

【0064】

他の実施では、intentによって指定された異なるトリガフレーズが、トリガフレーズとintentとの間の関係の強度を示すスコアに関連付けられ得、信頼度スコアが、一致したトリガフレーズのための関係スコアの強度に少なくとも部分的に基づいて判定され得る。たとえば、「Hail a cab」というトリガフレーズのための関係スコアの強度は、同じintentによって指定された「Get a car」というトリガフレーズのための関係スコアの強度よりも、トリガフレーズと、「Cab Caller」というアプリケーションを指定するintentとの間のより強い関係を示し得る。

【0065】

他の要因は、ユーザの場所、または、ユーザが含まれているとして識別されるアクティビティに基づいて、または、他の情報に基づいて、どれくらい頻りに候補_intentがトリガされたのか、どれくらい最近、候補_intentがトリガされたのか、他のどのアプリケーションが、関連され得る音声アクションが受信された時間においてクライアントデバイス202において実行しているのか、または、候補_intentによって指定されたアプリケーションと同時に頻りに使用されるのかのように、候補_intentのための信頼度スコアを判定する際に考慮され得る。いくつかの実施では、これらの要因の組合せが、特定の候補_intentのための信頼度スコアを判定する際に考慮され得る。1つまたは複数の候補_intentの信頼度スコアに基づいて、特定の候補_intentが、明確化エンジン230によって選択され得、明確化エンジンは、選択された候補_intentを識別する情報を、実行エンジン250へ送信し得る。

30

40

【0066】

それに加えて、他のユーザのアクションは、何人の他のユーザが、候補_intentによって指定されるアプリケーションをアクティブに使用しているのか、ユーザの知り合いまたは友達のうちの何人が、候補_intentによって指定されるアプリケーションを使用しているのか等のように、候補_intentの信頼度スコアを判定する場合に考慮され得る。たとえば、ユーザが「Call taxi」という音声入力を提供することに基づいて、音声アクションサービスシステム200は、2つの候補タクシーサービスアプリケーションが、音声入力によってトリガされ得ることを判定し得、さらに、より多くのタクシー運転手が、現在、これらアプリケーションのうちの1つを他のものよりも使用していることを判定し得る

50

。その結果、アプリケーションをアクティブに使用しているより多くの運転手を有するタクシーサービスアプリケーションのためのインテントへ割り当てられた信頼度スコアは、アプリケーションをアクティブに使用しているより少ない運転手しか有さないタクシーサービスアプリケーションのためのインテントへ割り当てられた信頼度スコアよりも高くなり得る。

【0067】

さらに他の例では、明確化エンジン230は、候補インテントの1つまたは複数から特定の候補インテントを選択するための要求を、ユーザ201のためのクライアントデバイス202へ提供し得る。たとえば、マッチャ220からの1つまたは複数の候補インテントを指定する情報を受信することに基づいて、明確化エンジンは、特定の候補インテント、または、特定のインテントによって指定される特定のアプリケーションを選択するための要求が、ユーザ201のためのクライアントデバイス202において提供されるようにし得る。明確化エンジン230は、ユーザ201による選択を示す情報を受信し得、ユーザ201が選択した特定の候補インテントを識別する情報を実行エンジン250へ提供し得る。

10

【0068】

たとえば、「Call taxi」というトランスクリプションが、「Cab Caller」というアプリケーションを指定するインテントに関連付けられたトリガフレーズに一致し、また、「TaxiNow」というアプリケーションを指定するインテントに関連付けられたトリガフレーズに一致することを判定することに基づいて、明確化エンジン230は、ユーザ201に対して、「Cab Caller」というアプリケーションまたは「TaxiNow」というアプリケーションのいずれかを選択するように要求するユーザインターフェースを、クライアントデバイス202において、ユーザ201へ提供し得る。明確化エンジン230は、ユーザが「Cab Caller」というアプリケーションを選択したことを示す情報を受信し得、これに応じて、「Cab Caller」アプリケーションを指定する候補インテントを識別する情報を実行エンジン250へ提供し得る。

20

【0069】

いくつかの実施では、明確化エンジン230は、信頼度スコアまたは他の情報に基づいて、候補インテントをランク付けし得、ランク付けに基づいて、特定の候補インテントを選択し得るか、または、ユーザ201へ提示されたユーザインターフェースに候補インテントを表示し得る。たとえば、明確化エンジン230は、候補インテントのおのおののために判定された信頼度スコアに基づいて、候補インテントをランク付けし得、これによって、たとえば、より高い信頼度スコアを有する候補インテントが、より低い信頼度スコアを有する候補インテントの上にランク付けされるようになる。他の要因は、どれくらい頻繁に、または、どれくらい最近に、候補インテントがトリガされたのか、どれくらい頻繁に、または、どれくらい最近に、候補インテントによって指定されたアプリケーションが、類似の音声入力に応じてユーザによって選択されたのかのように、候補インテントのランク付けに影響を与え得る。

30

【0070】

いくつかの事例では、多数の候補インテントからの特定の候補インテントのユーザ選択を要求するか、または、少なくとも1つの候補インテントによっておのおの指定されたアプリケーションのセットから、候補インテントによって指定された特定のアプリケーションの選択を要求するユーザインターフェースは、ランク付け、信頼度スコア、または、候補インテントまたはランク付けに関連付けられた他の情報または要因に基づいて、候補インテントまたはアプリケーションを別の方式で、ユーザへ提示し得る。たとえば、候補インテントによって指定された「Cab Caller」というアプリケーションが、「TaxiNow」というアプリケーションよりも、ユーザ201によってより頻繁に選択されたことを判定することに基づいて、「Cab Caller」というアプリケーションまたは「TaxiNow」というアプリケーションのうちの1つの選択を要求する、ユーザ201へ提示されるユーザインターフェースは、「Cab Caller」のためのオプションを、「TaxiNow」のためのオプションよりもより大きな選択可能なアイコンとして示し得、図2に図示されるように、「TaxiNow」とい

40

50

うアプリケーションのためのオプションよりも上方に、「Cab Caller」というアプリケーションのためのオプションを表示し得るか、または、さもなければ、これらオプションを、別の方式で、または、ユーザインターフェースの異なる場所に図示し得る。

【0071】

いくつかの事例では、明確化エンジン230はまた、1つまたは複数の候補インテントから特定の候補インテントを選択する情報を判定または受信する明確化に基づいて、ユーザパーソナル化データベース235において情報を更新し得る。たとえば、明確化エンジン230が、「Call taxi」という音声入力を提供することに応じて、「Cab Caller」というアプリケーションのユーザ201による選択を受信することに基づいて、明確化エンジン230は、ユーザパーソナル化データベースおよび/または文脈インテントデータベース240においてユーザパーソナル化情報を更新し得る。そのようなパーソナル化情報を更新するステップは、「Call taxi」というトリガフレーズと、「Cab Caller」というアプリケーションを指定する候補インテントとの間の関係スコアの強度を高めるステップを含み得、「Cab Caller」というアプリケーションを指定する候補インテントのための信頼度スコアを、より強い信頼度を示すように調節するステップと、「TaxiNow」というアプリケーションを指定する候補インテントのための信頼度スコアを、低減された信頼度を示すように調節するステップとを含み得るか、または、さもなければ、パーソナル化情報、または、ユーザパーソナル化データベース235または文脈インテントデータベース240において記憶されたインテントに関連付けられた情報を更新し得る。

10

【0072】

他の例では、明確化エンジンはまた、文脈インテントデータベース140またはOSインテントデータベースにおいて記憶されたインテントへアクセスでき得、インテントまたはインテントに関連する情報に関連付けられたパラメータを修正または調節でき得る。たとえば、「Call taxi」というユーザの音声入力に応じて、ユーザ201が、ユーザ201へ提供されたユーザインターフェースにおいて、「Cab Caller」というアプリケーションを選択し、「TaxiNow」というアプリケーションを選択しなかったことを判定することに基づいて、明確化エンジン230は、インテントのために指定されたトリガフレーズのセットから「Call taxi」というトリガフレーズを削除するために、「TaxiNow」というアプリケーションを指定し、文脈インテントデータベース140において記憶された候補インテントを更新し得るか、または、インテントによって指定された「Call taxi」というトリガフレーズのための関係スコアの強度を調節し得る。この処理および類似の処理は、音声アクションサービスシステム200が、適切なインテントをトリガしている際におけるサービスの撤回可能性を評価し、音声アクションサービスの精度または効率を高めるように、インテント、または、インテントに関連付けられた情報を調節することを有効にし得る。

20

30

【0073】

いくつかの例では、多数のユーザからのフィードバックが、インテント、または、インテントに関連する情報に関連付けられたパラメータをどのようにして調節するのかを判定する際に、アグリゲートおよび分析され得る。たとえば、ユーザの集合からアグリゲートされ、「Call taxi」というトリガフレーズが典型的には、ユーザが「TaxiNow」というアプリケーションよりも「Cab Caller」というアプリケーションを選択するという結果になることを示すフィードバックに基づいて、音声アクションサービスシステム200は、「Call taxi」というトリガフレーズと、「Cab Caller」というアプリケーションを指定するインテント、または「Cab Caller」というアプリケーション自体の、関係の強度または信頼度スコアを高めることを判定し得る。

40

【0074】

上記では個別に議論されているが、明確化エンジン230へ利用可能な方法のうちの1つまたは複数は、候補インテントから特定のインテントを選択するために使用され得る。たとえば、明確化エンジン230は、マッチャ220から、5つの候補インテントのセットを識別するデータを受信し得、5つの候補インテントに関連付けられた、または、5つの候補インテントのために判定された信頼度スコアに基づいて、候補インテントのうちの3つを削除し

50

得る。明確化エンジン230はその後、特定の候補_intentを判定するために、残りの2つの候補_intentによって指定されたアプリケーションから選択することをユーザ201へ求める要求を、クライアントデバイス202へ出力し得る。その後、明確化エンジン230は、ユーザが選択した_intentを指定する候補_intentを識別する情報を実行エンジン250へ送信し得る。

【0075】

それに加えて、いくつかの事例では、音声アクションサービスシステム200は、明確化エンジン230によって使用される方法のうちの1つまたは複数による候補_intentの分析に基づいて、異なるユーザインターフェースをユーザへ提示し得る。たとえば、明確化エンジン230が、候補_intentのための信頼度スコアのみに基づいて、特定の候補_intentを選択するのであれば、明確化エンジン230は、特定の_intent、または、特定の_intentによって指定されたアプリケーションを識別するユーザインターフェースを、クライアントデバイス202においてユーザ201へ提供し得る。あるいは、明確化エンジン230が、候補_intentの信頼度スコアに基づいて、明確化エンジン230が、候補_intent、または、候補_intentによって指定されたアプリケーションのユーザ選択を要求すべきことを判定するのであれば、明確化エンジン230は、ユーザ選択を要求する、出力のための異なるユーザインターフェースを、クライアントデバイス202において提供し得る。

【0076】

実行エンジン250は、明確化エンジン230から、特定の_intentを識別する情報を受信し得、intentによって指定されたアクティビティが、intentによって指定されたアプリケーションにおいて、または、intentによって指定されたアプリケーションによって実行されるように、intentをトリガし得る。たとえば、実行エンジン250は、文脈intentデータベース240において、識別されたintentにアクセスし得るか、または、さもなければ、識別されたintentにアクセスし得るか、または、識別されたintentを受信し得る。実行エンジン250は、intentによって指定されたアクティビティが、intentによって指定されたアプリケーションにおいて、または、intentによって指定されたアプリケーションによって実行されるようにするようintentをトリガし得る。たとえば、実行エンジン250は、アクティビティが、指定されたアプリケーションにおいて実行されるようにし得、アプリケーションにおいてアクティビティを実行するように、オペレーティングシステムまたは他のアプリケーションへ制御を提供し得、または、アプリケーションに対して、指定されたアクティビティを実行させ得る。他の例では、実行エンジン250は、トリガすべきintent、または、intentによって指定されたトリガすべきアクティビティを識別する情報を、クライアントデバイス202へ提供し得、クライアントデバイス202は、この情報を受信し、受信した情報に基づいて、intentまたはアクティビティをトリガし得る。いくつかの事例では、実行エンジン250は、クライアントデバイス202において実施され得る。

【0077】

例として、「Cab Caller」というアプリケーションを指定する候補_intentを識別する情報を明確化エンジン230から受信することに基づいて、実行エンジン250は、文脈intentデータベース240において「Cab Caller」というアプリケーションを指定するintentにアクセスし得、「Cab Caller」というアプリケーションを指定するintentをトリガし得る。「Cab Caller」というアプリケーションを指定するintentをトリガするステップは、「Cab Caller」というアプリケーションを起動するアクティビティのように、intentによって指定されたアクティビティを判定するステップを含み得る。intentによって指定されたアクティビティが、「Cab Caller」というアプリケーションを起動するためのアクティビティであることを判定することに基づいて、実行エンジン250は、クライアントデバイス202、またはアプリケーション、オペレーティングシステム、クライアントデバイス202にホストされた他のソフトウェアに対して、クライアントデバイス202において、「Cab Caller」というアプリケーションを起動させる情報を送信し得る。

10

20

30

40

50

【0078】

いくつかの事例では、インテントをトリガするステップは、インテント、または、インテントによって指定されたアクティビティを判定することなく、インテントに関連付けられた情報を送信するステップを含み得る。たとえば、実行エンジン250は、インテントによって指定されたアクティビティを判定することなく、識別されたインテントに関連付けられたデータを、クライアントデバイス202へ送信し得る。そのような場合、クライアントデバイス202、またはアプリケーション、オペレーティングシステム、または、クライアントデバイスによってホストされた他のソフトウェアは、インテントに基づいて実行するアクティビティを判定し、アクティビティが、インテントにおいて指定されたアプリケーションにおいて、または、インテントにおいて指定されたアプリケーションによって、実行されるようにし得る。

10

【0079】

クライアントデバイス202は、実行エンジン250から情報を受信し得、この情報に基づいて、指定されたインテントによって指定されたアクティビティを実行するための動作を実行し得る。たとえば、クライアントデバイス202は、明確化エンジン230によって識別されたインテントを受信し得るか、または、クライアントデバイス202、または、クライアントデバイス202によってホストされたソフトウェアに対して、「Cab Caller」というアプリケーションを起動するためのアクティビティのように、識別されたインテントによって指定されたアクティビティを実行するように指示する情報を受信し得る。受信した情報に基づいて、クライアントデバイス202、または、クライアントデバイス202によってホストされたソフトウェアは、アクティビティを実行し得る。たとえば、クライアントデバイス202のオペレーティングシステム、「Cab Caller」というアプリケーション自体、または、クライアントデバイス202にホストされた別のアプリケーションは、受信した情報に応じて、「Cab Caller」というアプリケーションを起動し得る。このように、ユーザ201によって提供される「Call taxi」という音声入力は、「Cab Caller」というタクシーサービスアプリケーションの起動を引き起こし得、これによって、ユーザ201は、その後、「Cab Caller」というアプリケーションを使用して、タクシーを頼むことができるであろう。

20

【0080】

図3は、開発者音声アクションサービスのための例示的な処理300のフローチャートを描写する。いくつかの実施では、図3の処理300は、図1の開発者音声アクションサービスシステム100によって実行され得るか、または、ローカルに、または、アクセス可能であり、1つまたは複数のネットワークを介してクライアントデバイスから情報を受信するように構成されたシステムのようなクライアントデバイスから離れて位置する別のシステムによって実行され得る。

30

【0081】

処理300は、少なくともアプリケーションおよびトリガフレーズを識別するデータを受信することによって開始する(302)。たとえば、図1の取得エンジン110は、1つまたは複数のネットワーク105を介して、アプリケーション開発者101に関連付けられたコンピューティングデバイス102から、文法を受信し得、受信した文法は、少なくともアプリケーションおよびトリガフレーズを指定し得る。受信したデータは、開発者音声アクションサービスシステム100に登録されるべき新たなインテントのための候補として取り扱われ得る。いくつかの例では、説明されるように、指定されたアプリケーションにおいて、または、指定されたアプリケーションによって実行されるべきアクティビティまたは複合アクティビティを指定する情報、新たなインテントに関連付けられるべきコンテキスト、トリガフレーズに関連付けられた1つまたは複数の引数、または、他の情報のような他の情報は、アプリケーション開発者101から受信した文法において指定され得る。

40

【0082】

アプリケーションおよびトリガフレーズを識別する情報は、新たなインテントを生成するために情報が使用され得るか否かを判定するために検証され得る(304)。たとえば、開

50

発者音声アクションサービスシステム100の検証エンジン120は、データによって識別されたトリガフレーズが、受け入れ可能なトリガフレーズであること、たとえば、トリガフレーズがまだ別のインテントに関連付けられていないこと、または、トリガフレーズが、識別されたアプリケーションに十分に関連していないこと、を判定することによって受信された情報を検証し得る。受信した情報を検証するステップはさらに、受信した情報のフォーマットが、新たなインテントを生成するために変換され得るか否かを判定するステップを含み得、受信したデータにおいて識別されるアプリケーションまたはアクティビティが、有効なアプリケーションであるか、または、識別されたアプリケーションによって実行され得るアクティビティであるかを判定するステップを含み得るか、または、他の検証手順を含み得る。

10

【0083】

いくつかの事例では、受信した情報について実行された検証手順に関連する、または、受信した情報について実行された検証手順に基づいて判定された情報を含むフィードバックが、アプリケーション開発者101に関連付けられたコンピューティングデバイス102へ提供され得る。たとえば、フィードバックは、新たなインテントを生成するように文法が変換され得るように、アプリケーション開発者101によって発行された文法が、検証処理を通過したか否かを示し得るか、または、アプリケーション開発者101によって発行された文法が、なぜ検証処理を通過するのに失敗したかの理由を示し得る。

【0084】

受信した情報は、少なくとも、アプリケーション、受信したデータによって識別されたトリガフレーズ、および、受信したデータによって識別されたトリガフレーズに少なくとも基づいて判定される1つまたは複数の他のトリガフレーズを指定するインテントを生成するように誘導され得る(306)。受信した情報を誘導するステップは、1つまたは複数の関連するトリガフレーズを生成するように、識別されたトリガフレーズを拡張することによって、少なくとも、識別されたトリガフレーズに基づいて判定された1つまたは複数の追加のトリガフレーズを判定するステップを含む。たとえば、説明されるように、受信したデータにおいて識別されたトリガフレーズは、1つまたは複数の拡張規則、類義語規則、用語オプション化規則、または他の規則を使用することによって拡張され得、識別されたトリガフレーズおよび関連するトリガフレーズの他の言語への変換を含むように拡張され得る。

20

30

【0085】

受信した情報を誘導するステップはまた、少なくとも、アプリケーション、識別されたトリガフレーズ、識別されたトリガフレーズに少なくとも基づいて判定された1つまたは複数の追加のトリガフレーズを指定する新たなインテントを生成するために、受信した情報を、コンピューティングデバイス102から受信される文法フォーマットから、インテントのためのフォーマットへ変換するステップを含み得る。いくつかの事例では、受信した情報を変換するステップはまた、新たなインテントを生成するために、受信した情報の文法フォーマットのような第1のフォーマットから、識別されたトリガフレーズに少なくとも基づいて判定される1つまたは複数のトリガフレーズをインテントフォーマットへ変換するステップを含み得る。

40

【0086】

少なくとも、アプリケーション、受信した情報において識別されるトリガフレーズ、および1つまたは複数の追加のトリガフレーズを指定する新たなインテントが、文脈インテントデータベースにおいて記憶される(308)。たとえば、図1の誘導エンジン130が、アプリケーション開発者101から受信した文法を、新たなインテントを生成するように誘導することに基づいて、開発者音声アクションサービスシステム100は、文脈インテントデータベース140において新たなインテントを記憶し得る。記憶されたインテントは、少なくとも、識別されたアプリケーション、識別されたトリガフレーズ、および、少なくとも識別されたトリガフレーズに基づいて判定された1つまたは複数のトリガフレーズを指定し得る。文脈インテントデータベース140において記憶された新たなインテントは、他の情

50

報のみならず、インテントがトリガされた場合に、アプリケーションにおいて、または、アプリケーションによって実行されるべきアクティビティ、インテントに関連付けられた1つまたは複数のコンテキスト、トリガフレーズとアプリケーションまたはアクティビティとの間の関係の強度を推定する関係スコアの1つまたは複数の強度、または他の情報を指定し得る。新たなインテントは、ユーザによって提供される将来の音声入力が、記憶されたインテントのうちの1つによって指定されるトリガフレーズへ一致されることを有効にするために、1つまたは複数の他のインテントとともに文脈インテントデータベース140において記憶され得る。一致を検出することに応じて、一致したインテントによって指定されたアクティビティが、一致したインテントによって指定されたアプリケーションにおいて、または、一致したインテントによって指定されたアプリケーションによって実行されるようにするために、一致したインテントがトリガされ得る。

10

【0087】

多くの実施が説明された。しかしながら、様々な修正が、本開示の精神および範囲から逸脱することなくなされ得ることが理解されるであろう。たとえば、上記に図示された様々な形式のフローが、順序換えされ、追加され、または削除されたステップとともに使用され得る。したがって、他の実施は、以下の特許請求の範囲内である。

【0088】

本明細書において議論されたシステムおよび/または方法が、ユーザに関する個人情報を収集し得るか、または、個人情報を活用し得る事例のために、ユーザは、プログラムまたは特徴が、たとえば、ユーザのソーシャルネットワーク、ソーシャルアクションまたはアクティビティ、職業、嗜好、または現在位置に関する情報のような個人情報を収集するか否かを制御するための、または、システムおよび/または方法が、よりユーザに関連する動作を実行するか否か、および/または、どのように実行するかを制御するための機会を提供され得る。それに加えて、あるデータは、記憶または使用される前に、1つまたは複数の方式で匿名化され得、これによって、個人的に識別可能な情報が削除されるようになる。たとえば、ユーザのための個人的に識別可能な情報が判定されないように、ユーザの識別情報が匿名化され得るか、または、ユーザの地理的位置が一般化され得、都市、郵便番号、または州レベルのような位置情報が取得され、これによって、ユーザの特定の位置が決定されなくなる。したがって、ユーザは、彼または彼女に関する情報がどのように収集され、使用されるのかについての制御を有し得る。

20

30

【0089】

この明細書において説明された実施形態、および機能的な動作のすべては、デジタル電子回路において、または、この明細書において開示された構成およびその構成的等価物を含むコンピュータソフトウェア、ファームウェア、またはハードウェアにおいて、または、これらの1つまたは複数の組合せにおいて実施され得る。実施形態は、1つまたは複数のコンピュータプログラム製品すなわち、データ処理装置による実行のために、または、データ処理装置の動作を制御するために、コンピュータ読取可能な媒体においてエンコードされたコンピュータプログラム命令の1つまたは複数のモジュールとして実施され得る。コンピュータ読取可能な媒体は、マシン読取可能な記憶デバイス、マシン読取可能な記憶基板、メモリデバイス、マシン読取可能な伝搬信号に影響を与える組成物、またはこれらのうちの1つまたは複数の組合せであり得る。「データ処理装置」という用語は、例としてプログラマブルプロセッサ、コンピュータ、または、マルチプロセッサまたはコンピュータを含む、データを処理するためのすべての装置、デバイス、およびマシンを包含する。装置は、ハードウェアに加えて、問題になっているコンピュータプログラムのための実行環境を生成するコード、たとえば、プロセッサファームウェア、プロトコルスタック、データベース管理システム、オペレーティングシステム、または、これらの1つまたは複数の組合せを構築するコードを含み得る。伝搬信号は、たとえば、適切な受信装置へ送信するための情報をエンコードするために生成された、マシン生成電気、光、または電磁信号のような、人工的に生成された信号である。

40

【0090】

50

(プログラム、ソフトウェア、ソフトウェアアプリケーション、スクリプト、またはコードとしても知られる)コンピュータプログラムは、コンパイルされたまたは解釈された言語を含む任意の形式のプログラミング言語で記述され得、スタンドアロンプログラムとして、または、モジュール、構成要素、サブルーチン、または、コンピュータ環境における使用のために適切な他のユニットとして含む任意の形態で展開され得る。コンピュータプログラムは、必ずしも、ファイルシステム内のファイルに対応する必要はない。プログラムは、他のプログラムまたはデータを保持するファイルの一部(たとえば、マークアップ言語ドキュメントに記憶された1つまたは複数のスクリプト)に、問題となっているプログラムへ特化された単一のファイルに、または、多数の調整されたファイル(たとえば、1つまたは複数のモジュール、サブプログラム、またはコードの一部を記憶するファイル)に記憶され得る。コンピュータプログラムは、1つのサイトにおいて、または、多数のサイトにわたって分散して配置され、通信ネットワークによって相互接続された1つのコンピュータまたは多数のコンピュータにおいて実行されるように展開され得る。

10

【0091】

この明細書において説明された処理および論理フローは、入力されたデータに対する演算を行い、出力を生成することによって機能を実行するために、1つまたは複数のコンピュータプログラムを実行する1つまたは複数のプログラマブルプロセッサによって実行され得る。処理および論理フローはまた、たとえば、FPGA(フィールドプログラマブルゲートアレイ)またはASIC(特定用途向け集積回路)のような、専用論理回路によって実行され得、装置もまた、専用論理回路として実施され得る。

20

【0092】

コンピュータプログラムの実行のために適切なプロセッサは、例として、汎用マイクロプロセッサと専用マイクロプロセッサとの両方、および、任意の種類のデジタルコンピュータの任意の1つまたは複数のプロセッサを含む。一般に、プロセッサは、読取専用メモリまたはランダムアクセスメモリまたはその両方から、命令およびデータを受信するであろう。

【0093】

コンピュータの必須の要素は、命令を実行するためのプロセッサと、命令およびデータを記憶するための1つまたは複数のメモリデバイスである。一般に、コンピュータはまた、たとえば磁気ディスク、磁気光ディスク、または光ディスクのように、データを記憶するための1つまたは複数の大容量記憶デバイスを含んでいるか、あるいは、データを記憶するための1つまたは複数の大容量記憶デバイスからデータを受け取るため、または、データを記憶するための1つまたは複数の大容量記憶デバイスへデータを転送するため、またはその両方のために、動作可能に結合され得る。しかしながら、コンピュータは、そのようなデバイスを有する必要はない。さらに、コンピュータは、別のデバイス、たとえば、いくつかの名前を挙げると、タブレットコンピュータ、モバイル電話、携帯情報端末(PDA)、モバイルオーディオプレーヤ、全地球測位システム(GPS)受信機のような別のデバイスに埋め込まれ得る。コンピュータプログラム命令およびデータを記憶するために適切なコンピュータ読取可能な媒体は、例によれば、たとえば、EPROM、EEPROM、およびフラッシュメモリデバイスのような半導体メモリデバイスと、たとえば内部ハードディスクまたはリムーバブルディスクのような磁気ディスクと、磁気光ディスクと、CD-ROMディスクおよびDVD-ROMディスクとを含む、すべての形態の不揮発性メモリ、媒体、およびメモリデバイスを含む。プロセッサおよびメモリは、専用論理回路によって補足され得るか、または、専用論理回路に組み込まれ得る。

30

40

【0094】

ユーザとのインタラクションを提供するために、実施形態は、情報をユーザへ表示するための、たとえばCRT(陰極放電管)またはLCD(液晶ディスプレイ)モニタのようなディスプレイデバイスと、ユーザがコンピュータへ入力を提供し得る、キーボードと、たとえばマウスまたはトラックボールのようなポインティングデバイスを有するコンピュータにおいて実施され得る。他の種類のデバイスも同様に、ユーザとのインタラクションを提供する

50

ために使用され得、たとえば、ユーザへ提供されるフィードバックは、たとえば、視覚的フィードバック、聴覚的フィードバック、または触覚的フィードバックのような任意の形態の知覚的フィードバックであり得、ユーザからの入力、音響、音声、または触覚入力を含む任意の形態で受信され得る。

【0095】

実施形態は、たとえばデータサーバとしてバックエンド構成要素を含む、または、たとえばアプリケーションサーバのようなミドルウェア構成要素を含む、または、たとえばユーザが実施とインタラクトし得るグラフィックユーザインターフェースまたはWebブラウザを有するクライアントコンピュータのようなフロントエンド構成要素を含むコンピューティングシステム、または、1つまたは複数のそのようなバックエンド構成要素、ミドルウェア構成要素、またはフロントエンド構成要素の任意の組合せにおいて実施され得る。システムの構成要素は、たとえば通信ネットワークのような、デジタルデータ通信の任意の形式または媒体によって相互接続され得る。通信ネットワークの例は、ローカルエリアネットワーク(「LAN」)と、たとえばインターネットのような広域ネットワーク(「WAN」)とを含む。

10

【0096】

コンピューティングシステムは、クライアントおよびサーバを含み得る。クライアントおよびサーバは、一般に、互いに離れており、典型的には、通信ネットワークを通じてインタラクトする。クライアントとサーバとの関係は、おのこのコンピュータにおいて実行し、互いにクライアント-サーバ関係を有している、コンピュータプログラムによって

20

【0097】

この明細書は、多くの詳細を含んでいるが、これらは、本開示または特許請求される範囲に関する限定としてではなく、特定の実施形態に特有の特徴の説明として解釈されるべきである。個別の実施形態のコンテキストにおいてこの明細書において説明されるいくつかの特徴はまた、単一の実施形態における組合せにおいても実施され得る。逆に、単一の実施形態のコンテキストにおいて説明される様々な特徴はまた、多数の実施形態において個別に、または、任意の適切な部分組合せにおいて実施され得る。さらに、特徴は、いくつかの実施形態において動作するものとして上記で説明され、また、そのように最初に特許請求もされているが、特許請求された組合せからの1つまたは複数の特徴は、いくつか

30

【0098】

同様に、動作は、特定の順序で図面に描写されているが、これは、所望される結果を達成するために、このような動作が、図示された特定の順序または連続順で実行されることも、または、例示されたすべての動作が実行されることも、必要であるとは理解されるべきではない。いくつかの状況では、マルチタスクおよび並列処理が有利であり得る。さらに、上記で説明された実施形態における様々なシステム構成要素の分離は、すべての実施形態において、そのような分離を必要とするものとして理解されるべきではなく、説明されたプログラム構成要素およびシステムは一般に、単一のソフトウェア製品へともに統合

40

【0099】

HTMLファイルが言及される各事例では、他のファイルタイプまたはフォーマットが、代用され得る。たとえば、HTMLファイルは、XML、JSON、プレーンテキスト、または他のタイプのファイルと交換され得る。さらに、テーブルまたはハッシュテーブルが言及される場合、(スプレッドシート、リレーショナルデータベース、または構造化されたファイルのような)他のデータ構造も使用され得る。

【0100】

したがって、特定の実施形態が説明された。他の実施形態は、以下の特許請求の範囲内

50

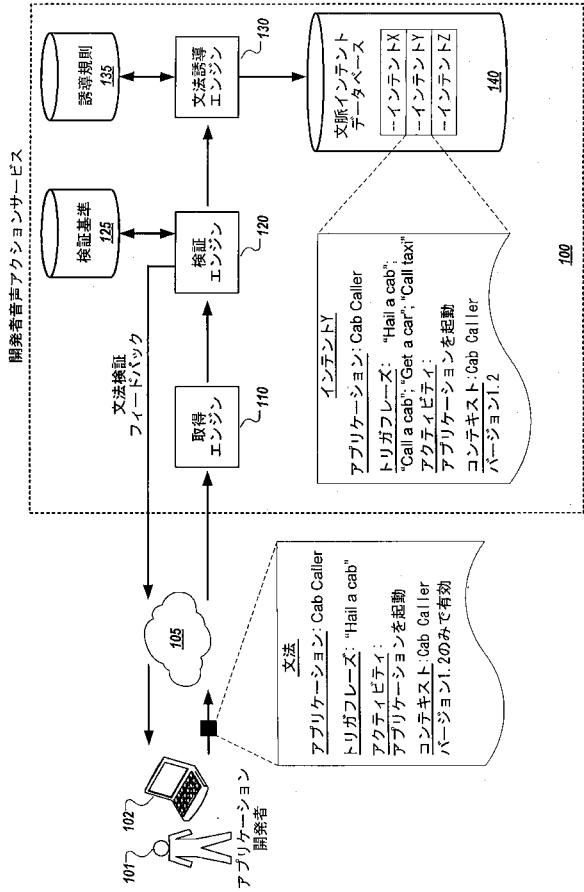
である。たとえば、特許請求の範囲に詳述されたアクションは、異なる順序で実行され得、さらに、所望される結果を達成し得る。

【符号の説明】

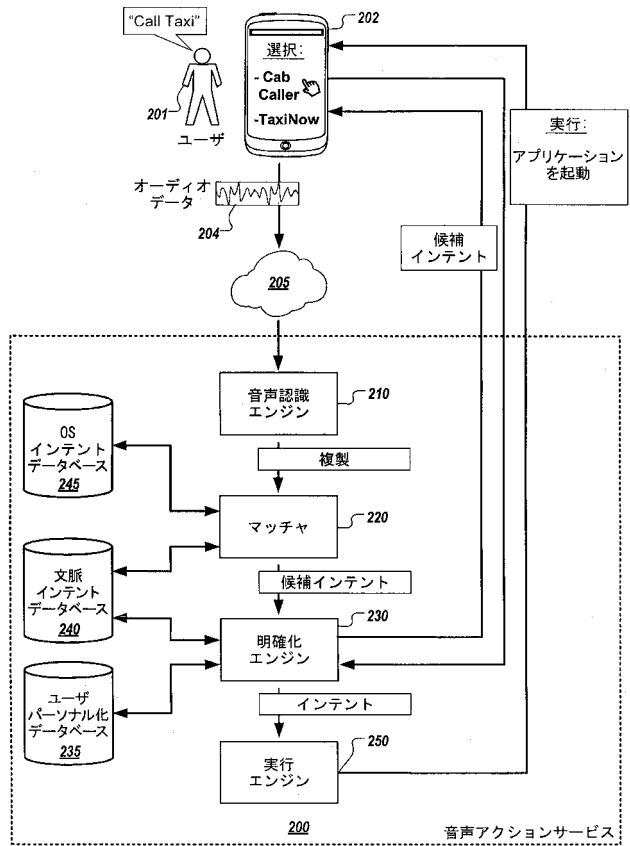
【0101】

100	開発者音声アクションサービスシステム	
101	アプリケーション開発者	
102	コンピューティングデバイス	
105	ネットワーク	
110	取得エンジン	
120	検証エンジン	10
125	検証基準	
130	文法誘導エンジン	
135	誘導規則	
140	文脈インテントデータベース	
200	音声アクションサービスシステム	
201	ユーザ	
202	クライアントデバイス	
204	オーディオデータ	
205	ネットワーク	
210	音声認識エンジン	20
220	マッチャ	
230	明確化エンジン	
235	ユーザパーソナリゼーションデータベース	
240	文脈インテントデータベース	
245	OSインテントデータベース	
250	実行エンジン	

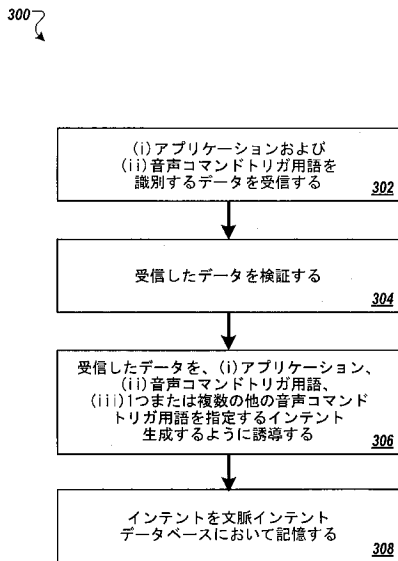
【 図 1 】



【 図 2 】



【 図 3 】



フロントページの続き

- (72)発明者 スニル・ヴェムリ
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク
ウェイ・1600
- (72)発明者 ニティン・マンゲシュ・シェティ
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク
ウェイ・1600
- (72)発明者 プラヴィール・クマール・グプタ
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク
ウェイ・1600
- (72)発明者 スコット・ビー・ハフマン
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク
ウェイ・1600
- (72)発明者 ハビエル・アレハンドロ・レイ
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク
ウェイ・1600
- (72)発明者 ジェフリー・エー・ボーツ
アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク
ウェイ・1600

【外国語明細書】

2019144598000001.pdf