

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7252327号

(P7252327)

(45)発行日 令和5年4月4日(2023.4.4)

(24)登録日 令和5年3月27日(2023.3.27)

(51)国際特許分類

G 0 6 F 16/90 (2019.01)

F I

G 0 6 F 16/90 1 0 0

請求項の数 13 (全41頁)

(21)出願番号 特願2021-519867(P2021-519867)
(86)(22)出願日 平成30年10月10日(2018.10.10)
(65)公表番号 特表2022-515005(P2022-515005 A)
(43)公表日 令和4年2月17日(2022.2.17)
(86)国際出願番号 PCT/CN2018/109704
(87)国際公開番号 WO2020/073248
(87)国際公開日 令和2年4月16日(2020.4.16)
審査請求日 令和3年5月21日(2021.5.21)

(73)特許権者 503433420
華為技術有限公司
HUAWEI TECHNOLOGIES
CO., LTD.
中華人民共和國 5 1 8 1 2 9 広東省深
▲チェン▼市龍崗区坂田 華為總部▲ベ
ン▼公樓
Huawei Administrat
ion Building, Banti
an, Longgang Distri
ct, Shenzhen, Guang
dong 5 1 8 1 2 9, P. R. C
hina
(74)代理人 100110364
弁理士 実広 信哉

最終頁に続く

(54)【発明の名称】 人間とコンピュータとの相互作用方法および電子デバイス

(57)【特許請求の範囲】

【請求項1】

人間とコンピュータとの相互作用方法であって、前記方法は、人間とコンピュータとの対話システムに適用され得、

サーバによって、第1の入力を受信するステップであって、前記第1の入力は、ユーザのサービス要求を含む、ステップ、

前記サーバによって、前記第1の入力に基づいて、前記第1の入力に対応する第1の分野を判定するステップであって、前記第1の分野は、前記ユーザの前記サービス要求に対応するタスクシナリオである、ステップ、

前記サーバによって、前記第1の入力を前記第1の分野に対応する意図認識モデルに割り当て、前記第1の入力に対応する第1の意図を認識するステップであって、前記第1の意図は、前記第1の分野のサブシナリオである、ステップ、

前記サーバによって、前記第1の入力から、前記第1の意図の第1のロットに対応する情報を抽出するステップであって、前記第1のロットは、前記第1の意図に事前構成される、ステップ、

前記第1のロットに対応する前記情報の抽出が失敗したと判定した場合に、前記サーバによって、前記第1のロットの属性を判定するステップであって、前記第1のロットの前記属性は、前記第1の意図に対応して事前設定される、ステップ、

前記第1のロットの前記属性が任意キースロットであると判定したことに応答して、前記サーバによって、前記第1のロットに対応する前記情報が必要かどうかを判定するた

10

20

めに前記ユーザに質問を行うステップ、

前記サーバによって、第2の入力を受信するステップであって、前記第2の入力は、前記第1のロットに対応する前記情報が必要かどうかを判定するために前記ユーザによって回答された情報を含む、ステップ、および

前記ユーザが、前記第1のロットに対応する前記情報が必要な情報であると判定した場合に、前記サーバによって、前記第2の入力から、前記第1のロットに対応する前記情報を抽出し、前記サーバによって、前記第1の意図と、前記第1のロットに対応する前記情報とに基づいて、前記第1の意図に対応する動作を実行するステップ、または

前記ユーザが、前記第1のロットに対応する前記情報が不要な情報であると判定した場合に、前記サーバによって、前記第1のロットに対応する前記情報を抽出することをスキップし、前記サーバによって、前記第1の意図に基づいて、前記第1の意図に対応する動作を実行するステップ

10

を含む、人間とコンピュータとの相互作用方法。

【請求項2】

前記サーバによって、前記第1の入力から、前記第1の意図の第1のロットに対応する情報を抽出する前記ステップは、

前記サーバによって、前記第1の入力から認識された各単語または各エンティティを前記第1のロットに対応するロット抽出モデルに入力し、前記第1の入力内の各単語または各エンティティに対応する信頼性レベルを計算するステップ、および

前記第1の入力内の第1の単語または第1のエンティティの信頼性レベルが第1の閾値以上である場合に、前記サーバによって、前記第1の単語または前記第1のエンティティが前記第1のロットに対応する前記情報であると判定するステップ、または

20

前記第1の入力内の各単語または各エンティティの前記信頼性レベルが前記第1の閾値未満である場合に、前記サーバによって、前記第1のロットに対応する前記情報の抽出が失敗したと判定するステップ

を含む、請求項1に記載の人間とコンピュータとの相互作用方法。

【請求項3】

前記人間とコンピュータとの相互作用方法は、

前記第1のロットがユーザ定義ロットタイプに対応する場合に、前記サーバによって、前記第1の入力から認識された各エンティティと前記ユーザ定義ロットタイプの各単語との類似度を計算するステップ、および

30

前記第1の入力から認識された各エンティティと前記ユーザ定義ロットタイプの各単語との前記類似度が第2の閾値未満である場合に、前記サーバによって、前記第1の入力が前記第1のロットに対応する前記情報を含まないと判定し、前記第1の入力内の第2のエンティティと前記ユーザ定義ロットタイプの第2の単語との類似度が第3の閾値以上である場合に、前記サーバによって、前記第2の単語が前記第1のロットに対応する前記情報であると判定し、または前記第1の入力内の任意のエンティティと前記ユーザ定義ロットタイプの任意の単語との類似度が前記第2の閾値以上であり、かつ前記第3の閾値未満である場合に、前記サーバによって、前記第1のロットに対応する前記情報が必要かどうかを判定するために前記ユーザに質問を行うことを決定するステップ

40

をさらに含む、請求項1または2に記載の人間とコンピュータとの相互作用方法。

【請求項4】

前記人間とコンピュータとの相互作用方法は、

前記第1の入力内の各単語または各エンティティの前記信頼性レベルが第4の閾値未満である場合に、前記サーバによって、前記第1の入力が前記第1のロットに対応する前記情報を含まないと判定するステップ、または

前記第1の入力内の任意の単語または任意のエンティティの信頼性レベルが前記第1の閾値未満であり、かつ前記第4の閾値以上である場合に、前記サーバによって、前記第1のロットに対応する前記情報が必要かどうかを判定するために前記ユーザに質問を行うことを決定するステップ

50

をさらに含む、請求項2に記載の人間とコンピュータとの相互作用方法。

【請求項5】

前記ユーザが、前記第1のロットに対応する前記情報が必要な情報であると判定した場合に、前記サーバによって、前記第2の入力から前記第1のロットに対応する前記情報を抽出する前記ステップは、

前記ユーザが、前記第1のロットに対応する前記情報が必要な情報であると判定した場合に、前記サーバによって、前記第1のロットに対応する前記情報抽出モデルを使用してまたは規則を使用して前記第2の入力から前記第1のロットに対応する前記情報を抽出するステップ

を含む、請求項1から4のいずれか一項に記載の人間とコンピュータとの相互作用方法。

【請求項6】

第2のロットが、前記第1の意図にさらに事前構成され、前記第2のロットの属性は、必須ロットであり、前記人間とコンピュータとの相互作用方法は、

前記第2のロットに対応する情報の抽出が失敗したと判定した場合に、前記サーバによって、前記第2のロットに対応する前記情報を抽出するために前記ユーザに質問を行うステップと、

前記サーバによって、第3の入力を受信し、前記第3の入力から前記第2のロットに対応する前記情報を抽出するステップであって、前記第3の入力は、前記ユーザの回答を含む、ステップと、

前記サーバによって、前記第1の意図と、前記第1のロットに対応する前記情報と、前記第2のロットに対応する前記情報とに基づいて、前記第1の意図に対応する前記動作を実行するか、または、前記サーバによって、前記第1の意図と、前記第2のロットに対応する前記情報とに基づいて、前記第1の意図に対応する前記動作を実行するステップ

をさらに含む、請求項1から5のいずれか一項に記載の人間とコンピュータとの相互作用方法。

【請求項7】

第3のロットが、前記第1の意図にさらに事前構成され、前記第3のロットの属性は、任意非キースロットであり、前記人間とコンピュータとの相互作用方法は、

前記第3のロットに対応する情報の抽出が失敗したと判定した場合に、前記サーバによって、前記第3のロットに対応する前記情報を抽出することをスキップするステップ

をさらに含む、請求項1から6のいずれか一項に記載の人間とコンピュータとの相互作用方法。

【請求項8】

前記第1のロットの前記属性は、開発者によって設定される、請求項1から7のいずれか一項に記載の人間とコンピュータとの相互作用方法。

【請求項9】

前記第1のロットに関して行われる前記質問は、デフォルトの質問であるか、または開発者によってカスタマイズされた質問である、請求項1から7のいずれか一項に記載の人間とコンピュータとの相互作用方法。

【請求項10】

サーバであって、前記サーバは、人間とコンピュータとの対話システムに適用され得、通信インターフェースと、メモリと、プロセッサとを備え、前記通信インターフェースおよび前記メモリは、前記プロセッサに結合されており、前記メモリは、コンピュータプログラムコードを記憶するように構成されており、前記コンピュータプログラムコードは、コンピュータ命令を含み、前記プロセッサが前記メモリから前記コンピュータ命令を読み出すとき、前記サーバは、請求項1から9のいずれか一項に記載の人間とコンピュータとの相互作用方法を実行することが可能である、サーバ。

【請求項11】

電子デバイスおよびサーバを備え、コンピュータ命令を含む人間とコンピュータとの対話システムであって、前記コンピュータ命令が前記電子デバイスおよび前記サーバ上で実

10

20

30

40

50

行されるとき、前記システムは、請求項1から9のいずれか一項に記載の人間とコンピュータとの相互作用方法を実行することが可能である、人間とコンピュータとの対話システム。

【請求項12】

コンピュータ命令を含むコンピュータ記憶媒体であって、前記コンピュータ命令が端末上で実行されるとき、前記端末は、請求項1から9のいずれか一項に記載の人間とコンピュータとの相互作用方法を実行することが可能である、コンピュータ記憶媒体。

【請求項13】

プログラムであって、前記プログラムがコンピュータ上で実行されるとき、前記コンピュータは、請求項1から9のいずれか一項に記載の人間とコンピュータとの相互作用方法を実行することが可能である、プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、通信技術の分野に関し、特に、人間とコンピュータとの相互作用方法および電子デバイスに関する。

【背景技術】

【0002】

人間とコンピュータとの対話システムまたは人間とコンピュータとの対話プラットフォームもしくはチャットボット(chatbot)などと呼ばれるものは、新世代の人間とコンピュータとの相互作用インターフェースである。チャットボットは、ユーザとの対話を行い、対話中にユーザの意図を認識し、食事の注文、チケットの予約、および配車などのサービスをユーザに提供し得る。

【0003】

図1Aは、チャットボットとユーザとの間の対話のプロセスの例である。この例は、チャットボットの動作プロセスを説明するために使用され、この動作プロセスは、オープンドメイン対話、エントリ基準、およびクローズドドメイン対話を含み得る。オープンドメイン対話は、チャットボットがユーザの意図を認識していない場合に行われる対話である。ユーザが「私に車を呼んで」と発話したら、チャットボットは、論理的な判断(すなわち、エントリ基準)によってユーザの意図(すなわち、配車サービス)を判定し、次にクローズドドメイン対話にジャンプする。クローズドドメイン対話は、ユーザの意図が認識された後にユーザの目的を明確にする(またはタスクの詳細を明確にする)ために行われる対話である。

【0004】

クローズドドメイン対話は、具体的には、スロット記入プロセスと、文言明確化プロセスと、結果応答プロセスとを含む。スロット記入プロセスは、ユーザの意図が特定のユーザ命令に変換されることを可能にする情報をスロットに記入するプロセスである。スロットは、意図を表現するためにユーザによって使用されるキー情報として理解され得る。例えば、図1Aに示されている対話では、配車サービスに対応するスロットは、出発地スロット、目的地スロット、および出発時間スロットを含む。チャットボットは、ユーザとの対話に基づいて、スロットに対応する情報(例えば、スロット値)を抽出する。スロットに対応する情報の中に何らかの必要な情報が欠落している場合、チャットボットは能動的に質問を行い、ユーザは質問に回答し、これにより、チャットボットは、ユーザの回答に基づいて、スロットに対応する必要な情報を記入する。このプロセスは、文言明確化プロセスと呼ばれる。スロットに対応するすべての情報を収集した後、チャットボットは、対応する動作を実行し得る、例えば、配車アプリケーションを使用してユーザのために注文を行い、注文を行った後にユーザに通知し得る、すなわち、結果応答プロセスを実行し得る。

【0005】

現在、2つのタイプのスロット、すなわち、必須スロットおよび任意スロットが存在する。チャットボットが、必須スロットに対応する情報の抽出に失敗した場合、チャットボットは、必須スロットに対応する情報が抽出されるまで、ユーザに情報の明確化を要求す

10

20

30

40

50

るために能動的に質問を行う。チャットボットが任意スロットに対応する情報の抽出に失敗した場合、チャットボットは、質問を行わず、任意スロットに対応する情報が存在しない状況に基づいて、対応する動作を直接実行する。

【0006】

実際のシナリオでは、チャットボットは通常、ユーザ入力エラー、音声認識エラー、または不正確なスロット抽出アルゴリズムなどの要因により、何らかの任意スロットに対応するキー情報の抽出に失敗する。その結果、チャットボットによって実行されるその後の動作が、ユーザの要求を満たさない可能性がある。例えば、図1Aに示されている対話において、ユーザによって発話された「カープール」は、任意スロットに対応するキー情報であり得る。チャットボットがキー情報を正確に抽出しなかった場合、チャットボットは、ユーザに対してカープールサービスを呼び出し得ない。これは、ユーザの意図に反しており、ユーザ体験に深刻な影響を及ぼす。

10

【発明の概要】

【0007】

本出願は、ユーザの意図を正確に認識し、ユーザの要求を満たし、ユーザ体験を改善するために、人間とコンピュータとの相互作用方法および電子デバイスを提供する。

【0008】

第1の態様によれば、方法が、本出願で提供され、人間とコンピュータとの対話システムに適用され得る。本方法は、サーバが第1の入力を受信し、第1の入力がユーザのサービス要求を含むことを含む。サーバは、第1の入力に基づいて、第1の入力に対応する第1の分野を判定し、第1の分野は、ユーザのサービス要求に対応するタスクシナリオである。サーバは、第1の入力を第1の分野に対応する意図認識モデルに割り当て、第1の入力に対応する第1の意図を認識し、第1の意図は、第1の分野のサブシナリオである。サーバは、第1の入力から、第1の意図の第1のスロットに対応する情報を抽出し、第1のスロットは、第1の意図に事前構成され、第1のスロットは、任意キースロットである。サーバが、第1のスロットに対応する情報の抽出が失敗したと判定した場合、サーバは、第1のスロットに対応する情報が必要かどうかを判定するためにユーザに質問を行う。サーバは、第2の入力を受信し、第2の入力は、第1のスロットに対応する情報が必要かどうかを判定するためにユーザによって回答された情報を含む。ユーザが、第1のスロットに対応する情報が必要な情報であると判定した場合、サーバは、第2の入力から第1のスロットに対応する情報を抽出し、サーバは、第1の意図と、第1のスロットに対応する情報とに基づいて、第1の意図に対応する動作を実行するか、またはユーザが、第1のスロットに対応する情報が不要な情報であると判定した場合、サーバは、第1のスロットに対応する情報を抽出せず、サーバは、第1の意図に基づいて、第1の意図に対応する動作を実行する。

20

30

【0009】

第1の入力は、ユーザとサーバ200との間の1回の対話における1つの発話であってもよいし、またはユーザとサーバ200との間の複数回の対話における複数の発話であってもよい。これは、本出願の本実施形態では限定されない。

【0010】

第2の入力は、ユーザとサーバ200との間の1回の対話における1つの発話であってもよいし、またはユーザとサーバ200との間の複数回の対話における複数の発話であってもよい。これは、本出願の本実施形態では限定されない。

40

【0011】

任意キースロットは、意図を表現するときにユーザがスロットに対応する情報を必ずしも表現しないことを意味することが理解され得る。スロットに対応する情報が表現されなかった場合、チャットボットは、スロットに対応する情報を無視してもよい。しかしながら、ユーザがスロットに対応する情報を表現した場合、チャットボットは、スロットに対応する情報を正確に抽出する必要がある。

【0012】

本出願の本実施形態では、サーバが事前設定されたすべてのスロットに対応する情報を

50

ユーザの発話に基づいて自動的に抽出するプロセスにおいて、対応する情報の抽出が失敗したスロットが存在し、そのスロットが任意キースロットである場合、チャットボットは、任意キースロットに対応する情報が必要かどうかに関してユーザからの確認を能動的に求めることが知られ得る。その情報が必要な場合、チャットボットは、ユーザの回答に基づいて、任意キースロットに対応する情報を引き続き抽出する。その情報が不要な場合、チャットボットは、任意キースロットに対応する情報をもはや抽出しない、言い換えれば、チャットボットは、ユーザからのさらなる確認を求めない。このようにして、任意キースロットに対応する情報の抽出が失敗した場合、チャットボットは、ユーザの意図の正確な認識を保証し、ユーザの要求を満たし、ユーザ体験を改善するために、ユーザからの確認をさらに求め得る。

10

【0013】

可能な実施態様では、サーバが第1の入力から、第1の意図の第1のスロットに対応する情報を抽出することは、以下を含む。

【0014】

サーバは、第1の入力から認識された各単語または各エンティティを、第1のスロットに対応するスロット抽出モデルに入力し、第1の入力内の各単語または各エンティティに対応する信頼性レベルを計算する。第1の入力内の第1の単語または第1のエンティティの信頼性レベルが第1の閾値以上である場合、サーバは、第1の単語または第1のエンティティが第1のスロットに対応する情報であると判定し、または第1の入力内の各単語または各エンティティの信頼性レベルが第1の閾値未満である場合、サーバは、第1のスロットに対応する情報の抽出が失敗したと判定する。

20

【0015】

可能な実施態様では、本方法は、第1のスロットがユーザ定義スロットタイプに対応する場合に、サーバが第1の入力から認識された各エンティティとユーザ定義スロットタイプの各単語との類似度を計算することをさらに含む。

【0016】

第1の入力から認識された各エンティティとユーザ定義スロットタイプの各単語との類似度が第2の閾値未満である場合、サーバは、第1の入力が第1のスロットに対応する情報を含まないと判定し、第1の入力内の第2のエンティティとユーザ定義スロットタイプの第2の単語との類似度が第3の閾値以上である場合、サーバは、第2の単語が第1のスロットに対応する情報であると判定し、または第1の入力内の任意のエンティティとユーザ定義スロットタイプの任意の単語との類似度が第2の閾値以上であり、かつ第3の閾値未満である場合、サーバは、第1のスロットに対応する情報が必要かどうかを判定するためにユーザに質問を行うことを決定する。

30

【0017】

エンティティとユーザ辞書内のキーワードとの類似度が判定されるとき、エンティティとユーザ辞書内のキーワードとの類似度を判定するために、例えば、ピンイン類似度に基づくアルゴリズムまたは文字列類似度に基づくアルゴリズムを使用して、第1の入力から認識されたエンティティとユーザ辞書内のキーワードとのレーベンシュタイン距離が計算され得る。あるいは、深層学習単語ベクトルまたは文ベクトルなどを使用して、単語または句間の類似度が計算されてもよい。本出願の本実施形態では、類似度計算方法は限定されない。

40

【0018】

ユーザ入力エラーまたは音声認識エラーが発生したとき、ユーザが第1のスロットに対応する情報を発話した可能性があるが、サーバは、第1のスロットに対応する情報の抽出に失敗する。サーバ200は、サーバが、誤り訂正方法を使用して、ユーザの発話（すなわち、第1の入力）から認識されたエンティティがユーザ辞書内のキーワードに比較的類似していると判定した場合にのみ、ユーザ確認メカニズムをトリガし得る。これは、ユーザによる確認の回数を低減し、ユーザへの過度の妨害を回避し、ユーザ体験を改善する。

【0019】

50

可能な実施態様では、本方法は、第1の入力内の各単語または各エンティティの信頼性レベルが第4の閾値未満である場合に、サーバが、第1の入力が第1のロットに対応する情報を含まないと判定し、または第1の入力内の任意の単語または任意のエンティティの信頼性レベルが第1の閾値未満であり、かつ第4の閾値以上である場合に、サーバが、第1のロットに対応する情報が必要かどうかを判定するためにユーザに質問を行うことを決定することをさらに含む。

【0020】

ユーザが第1のロットに対応する情報を正しく表現している場合、不正確なロット抽出モデルに起因して、第1のロットに対応する情報の抽出が失敗し得る。例えば、サーバは、スキル開発者がロット抽出モデルをトレーニングする前、比較的少ないユーザの発話しか入力されないため、または不正確なユーザの発話が入力されるため、トレーニングによって不正確なロット抽出モデルを生成する。この場合、ユーザは、確認閾値を設定し得る。ユーザの発話から認識されたエンティティに関してロット抽出モデルによって提供されるロットマーキング確率値が確認閾値よりも大きい場合、サーバは、ユーザ確認メカニズムをトリガする。これは、ユーザによる確認の回数を低減し、ユーザへの過度の妨害を回避し、ユーザ体験を改善する。

【0021】

可能な実施態様では、ユーザが、第1のロットに対応する情報が必要な情報であると判定した場合に、サーバが第2の入力から第1のロットに対応する情報を抽出することは、ユーザが、第1のロットに対応する情報が必要な情報であると判定した場合に、サーバが、第1のロットに対応するロット抽出モデルを使用してまたは規則を使用して第2の入力から第1のロットに対応する情報を抽出することを含む。

【0022】

同じエンティティおよび同じロット抽出モデルに関して、ロット抽出モデルは、最初はエンティティを正しく認識しない可能性があるが、2回目はエンティティを正しく認識し得る。これは、ユーザが最初にエンティティを発話したときにステートメントが別のエンティティを含み得るからである。すなわち、エンティティはコンテキストを有する。ロット抽出モデルが不正確である場合、エンティティの認識は、コンテキストの認識が失敗することから失敗する可能性がある。このため、サーバが最初にエンティティを認識し得なかった場合、サーバは、エンティティに関してユーザに質問を行う。この場合、ユーザの回答はエンティティに関するものである。ユーザの回答は、エンティティのみ、またはごく少数のコンテキストを含み得、ロット抽出モデルが、今度はエンティティを認識する可能性が高い。一部の他の実施形態では、エンティティはまた、ロット抽出モデルが使用されない方法でユーザの回答から認識され得る。例えば、エンティティを認識するための規則が可能であり得る。この規則は、ユーザの回答のコンテキスト論理、ユーザの意図との関連性、およびエンティティと第1のロットとの対応関係などの要因を参照してエンティティが認識され得ることである。これはまた、サーバが2回目またはそれ以降にユーザによって発話されたエンティティを認識する確率を効果的に高め得る。

【0023】

可能な実施態様では、第2のロットが、第1の意図にさらに事前構成され、第2のロットは、必須ロットであり、人間とコンピュータとの相互作用方法は、サーバが、第2のロットに対応する情報の抽出が失敗したと判定した場合に、サーバが、第2のロットに対応する情報を抽出するためにユーザに質問を行うことをさらに含む。サーバは、第3の入力を受信し、第3の入力から第2のロットに対応する情報を抽出し、第3の入力は、ユーザの回答を含む。サーバは、第1の意図と、第1のロットに対応する情報と、第2のロットに対応する情報とに基づいて、第1の意図に対応する動作を実行するか、またはサーバは、第1の意図と、第2のロットに対応する情報とに基づいて、第1の意図に対応する動作を実行する。

【0024】

可能な実施態様では、第3のロットが、第1の意図にさらに事前構成され、第3のロット

ットは、任意非キースロットであり、人間とコンピュータとの相互作用方法は、サーバが、第3のスロットに対応する情報の抽出が失敗したと判定した場合に、サーバが、第3のスロットに対応する情報を抽出しないことをさらに含む。

【0025】

第2の態様によれば、サーバが提供される。サーバは、人間とコンピュータとの対話システムに適用され得、通信インターフェースと、メモリと、プロセッサとを含み、通信インターフェースおよびメモリは、プロセッサに結合されており、メモリは、コンピュータプログラムコードを記憶するように構成されており、コンピュータプログラムコードは、コンピュータ命令を含み、プロセッサがメモリからコンピュータ命令を読み出すとき、サーバは、以下のステップ、すなわち、

通信インターフェースを介して第1の入力を受信するステップであって、第1の入力は、ユーザのサービス要求を含む、ステップ、第1の入力に基づいて、第1の入力に対応する第1の分野を判定するステップであって、第1の分野は、ユーザのサービス要求に対応するタスクシナリオである、ステップ、第1の入力を第1の分野に対応する意図認識モデルに割り当て、第1の入力に対応する第1の意図を認識するステップであって、第1の意図は、第1の分野のサブシナリオである、ステップ、第1の入力から、第1の意図の第1のスロットに対応する情報を抽出するステップであって、第1のスロットは、第1の意図に事前構成され、第1のスロットは、任意キースロットである、ステップ、サーバが、第1のスロットに対応する情報の抽出が失敗したと判定した場合に、第1のスロットに対応する情報が必要かどうかを判定するためにユーザに質問を行うステップ、通信インターフェースを介して第2の入力を受信するステップであって、第2の入力は、第1のスロットに対応する情報が必要かどうかを判定するためにユーザによって回答された情報を含む、ステップ、およびユーザが、第1のスロットに対応する情報が必要な情報であると判定した場合に、第2の入力から、第1のスロットに対応する情報を抽出し、第1の意図と、第1のスロットに対応する情報とに基づいて、第1の意図に対応する動作を実行するステップ、またはユーザが、第1のスロットに対応する情報が不要な情報であると判定した場合に、第1のスロットに対応する情報を抽出することをスキップし、第1の意図に基づいて、第1の意図に対応する動作を実行するステップを実行することが可能である。

【0026】

可能な実施態様では、プロセッサが第1の入力から、第1の意図の第1のスロットに対応する情報を抽出することは、具体的には、プロセッサが、第1の入力から認識された各単語または各エンティティを第1のスロットに対応するスロット抽出モデルに入力し、第1の入力内の各単語または各エンティティに対応する信頼性レベルを計算すること、第1の入力内の第1の単語または第1のエンティティの信頼性レベルが第1の閾値以上である場合に、第1の単語または第1のエンティティが第1のスロットに対応する情報であると判定すること、または第1の入力内の各単語または各エンティティの信頼性レベルが第1の閾値未満である場合に、第1のスロットに対応する情報の抽出が失敗したと判定することを含む。

【0027】

可能な実施態様では、プロセッサは、第1のスロットがユーザ定義スロットタイプに対応する場合、第1の入力から認識された各エンティティとユーザ定義スロットタイプの各単語との類似度を計算し、

第1の入力から認識された各エンティティとユーザ定義スロットタイプの各単語との類似度が第2の閾値未満である場合、第1の入力が第1のスロットに対応する情報を含まないと判定し、第1の入力内の第2のエンティティとユーザ定義スロットタイプの第2の単語との類似度が第3の閾値以上である場合、第2の単語が第1のスロットに対応する情報であると判定し、または第1の入力内の任意のエンティティとユーザ定義スロットタイプの任意の単語との類似度が第2の閾値以上であり、かつ第3の閾値未満である場合、第1のスロットに対応する情報が必要かどうかを判定するためにユーザに質問を行うことを決定するようにさらに構成される。

【0028】

可能な実施態様では、プロセッサは、第1の入力内の各単語または各エンティティの信頼性レベルが第4の閾値未満である場合、第1の入力が第1のロットに対応する情報を含まないと判定し、または第1の入力内の任意の単語または任意のエンティティの信頼性レベルが第1の閾値未満であり、かつ第4の閾値以上である場合、第1のロットに対応する情報が必要かどうかを判定するためにユーザに質問を行うことを決定するようにさらに構成される。

【0029】

可能な実施態様では、ユーザが、第1のロットに対応する情報が必要な情報であると判定した場合に、プロセッサが第2の入力から第1のロットに対応する情報を抽出することは、具体的には、ユーザが、第1のロットに対応する情報が必要な情報であると判定した場合に、プロセッサが、第1のロットに対応するロット抽出モデルを使用してまたは規則を使用して第2の入力から第1のロットに対応する情報を抽出することを含む。

【0030】

可能な実施態様では、第2のロットが第1の意図にさらに事前構成され、第2のロットが必須ロットである場合、プロセッサは、プロセッサが、第2のロットに対応する情報の抽出が失敗したと判定した場合、第2のロットに対応する情報を抽出するためにユーザに質問を行い、通信インターフェースを介して第3の入力を受信し、第3の入力から第2のロットに対応する情報を抽出し、第3の入力は、ユーザの回答を含み、第1の意図と、第1のロットに対応する情報と、第2のロットに対応する情報とに基づいて、第1の意図に対応する動作を実行するか、または第1の意図と、第2のロットに対応する情報とに基づいて、第1の意図に対応する動作を実行するようにさらに具体的に構成される。

【0031】

可能な実施態様では、第3のロットが第1の意図にさらに事前構成され、第3のロットが任意非キースロットである場合、プロセッサは、プロセッサが、第3のロットに対応する情報の抽出が失敗したと判定した場合、第3のロットに対応する情報を抽出することをスキップするようにさらに具体的に構成される。

【0032】

第3の態様によれば、コンピュータ記憶媒体が提供され、コンピュータ命令を含む。コンピュータ命令が端末上で実行されるとき、端末は、第1の態様または第1の態様の可能な実施態様のいずれか1つにおける方法を実行することが可能である。

【0033】

第4の態様によれば、コンピュータプログラム製品が提供される。コンピュータプログラム製品がコンピュータ上で実行されるとき、コンピュータは、第1の態様または第1の態様の可能な実施態様のいずれか1つにおける方法を実行することが可能である。

【図面の簡単な説明】

【0034】

【図1A】従来技術における人間とコンピュータとの対話に関する端末インターフェースの概略図である。

【図1B】本出願の一実施形態による人間とコンピュータとの対話に関する端末インターフェースの概略図である。

【図2】本出願の一実施形態による人間とコンピュータとの対話システムの構成要素の概略図1である。

【図3】本出願の一実施形態による人間とコンピュータとの対話システムの構成要素の概略図2である。

【図4】本出願の一実施形態による電子デバイスの概略構造図である。

【図5A】本出願の一実施形態による電子デバイスのインターフェースの概略図である。

【図5B】本出願の一実施形態による電子デバイスのインターフェースの概略図である。

【図5C】本出願の一実施形態による電子デバイスのインターフェースの概略図である。

【図6A】本出願の一実施形態による電子デバイスの他のインターフェースの概略図である。

【図 6 B】本出願の一実施形態による電子デバイスの他のインターフェースの概略図である。

【図 6 C】本出願の一実施形態による電子デバイスの他のインターフェースの概略図である。

【図 7】本出願の一実施形態による電子デバイスの他のインターフェースの概略図である。

【図 8 A】本出願の一実施形態による電子デバイスの他のインターフェースの概略図である。

【図 8 B】本出願の一実施形態による電子デバイスの他のインターフェースの概略図である。

【図 9 A】本出願の一実施形態による人間とコンピュータとの相互作用方法の概略フローチャート1である。

10

【図 9 B】本出願の一実施形態による人間とコンピュータとの相互作用方法の概略フローチャート1である。

【図 10 A】本出願の一実施形態による人間とコンピュータとの相互作用方法の概略フローチャート2である。

【図 10 B】本出願の一実施形態による人間とコンピュータとの相互作用方法の概略フローチャート2である。

【図 11】本出願の一実施形態によるサーバの概略構造図である。

【発明を実施するための形態】

【0035】

20

以下では、本出願の実施形態の添付の図面を参照して、本出願の実施形態の技術的解決策を説明する。本出願の実施形態の説明では、別段の指定がない限り、「／」は、「または」を意味する。例えば、A／Bは、AまたはBを表し得る。本明細書では、「および／または」は、関連付けられた対象を説明するための関連付けの関係のみを説明し、3つの関係が存在し得ることを表す。例えば、Aおよび／またはBは、以下の3つの状況、すなわち、Aのみが存在する状況、AとBとの両方が存在する状況、およびBのみが存在する状況を表し得る。

【0036】

以下の「第1の」および「第2の」という用語は、説明を目的としたものにすぎず、相対的な重要性の指示もしくは暗示または示された技術的特徴の数の暗黙の指示として理解されるべきではない。したがって、「第1の」または「第2の」によって限定される特徴は、1つ以上の特徴を明示的または暗黙的に含み得る。本出願の実施形態の説明では、別段の記載がない限り、「複数」は、2つまたは2つより多くを意味する。

30

【0037】

従来技術では、チャットボットは、ユーザ入力エラー、音声認識エラー、または不正確なスロット抽出アルゴリズムなどの要因により、何らかの任意スロットに対応するキー情報の抽出に失敗する。その結果、チャットボットによって実行されるその後の動作が、ユーザの要求を満たさない。本出願の実施形態は、任意スロットを任意キースロットと任意非キースロットとにさらに分類するために、人間とコンピュータとの相互作用方法を提供する。さらに、ユーザ確認メカニズムが、任意キースロットのために構成される。具体的には、チャットボットが事前設定されたすべてのスロットに対応する情報をユーザの発話に基づいて自動的に抽出するプロセスにおいて、対応する情報の抽出が失敗したスロットが存在し、そのスロットが任意キースロットである場合、チャットボットは、任意キースロットに対応する情報が必要かどうかに関してユーザからの確認を能動的に求める。その情報が必要な場合、チャットボットは、ユーザの回答に基づいて、任意キースロットに対応する情報を引き続き抽出する。その情報が不要な場合、チャットボットは、任意キースロットに対応する情報をもはや抽出しない、言い換えれば、チャットボットは、ユーザからのさらなる確認を求めない。このようにして、任意キースロットに対応する情報の抽出が失敗した場合、チャットボットは、ユーザの意図の正確な認識を保証し、ユーザの要求を満たし、ユーザ体験を改善するために、ユーザからの確認をさらに求め得る。

40

50

【0038】

任意キースロットは、意図を表現するときにユーザがスロットに対応する情報を必ずしも表現しないことを意味することが理解され得る。スロットに対応する情報が表現されなかった場合、チャットボットは、スロットに対応する情報を無視してもよい。しかしながら、ユーザがスロットに対応する情報を表現した場合、チャットボットは、スロットに対応する情報を正確に抽出する必要がある。

【0039】

例えば、図1Bは、本出願の一実施形態によるチャットボットとユーザとの間の対話のプロセスの例を示している。「カープール」は、配車アプリケーションの任意キースロットとして構成される。チャットボットとユーザとの間の対話において、ユーザは「カープール」と発話するが、チャットボットは、情報（すなわち、任意キースロットに対応する情報）の抽出に失敗する。この場合、チャットボットは、ユーザからのさらなる確認を求める、すなわち、ユーザに「カープールしますか？」と質問を行う。次に、チャットボットは、ユーザの意図に対応する動作が正確に実行されることを保証するために、ユーザの回答から任意キースロットに対応する情報を抽出する。

【0040】

ユーザの回答が「カープール」である場合、それは、ユーザがカープールを望んでおり、任意キースロットに対応する情報が重要であることを示す。任意キースロットに対応する情報を抽出した後、チャットボットは、ユーザのためにカープールの注文を行い得る。ユーザの回答が「カープールしない」である場合、それは、ユーザがカープールを望んでおらず、任意キースロットに対応する情報が重要であることを示す。任意キースロットに対応する情報を抽出した後、チャットボットは、ユーザのために非カープールの注文を行い得る。ユーザの回答が「どちらでも構わない」である場合、それは、任意キースロットに対応する情報が重要ではないことを示す。チャットボットは、カープールの要因を考慮せずにユーザのために注文を行い得る。任意キースロットに対応する情報に関してユーザからの確認を求めることは、任意キースロットに対応する重要な情報の抽出を遂行するだけでなく、ユーザの意思を確かめることにも役立ち得ることが理解され得る。これは、チャットボットによってユーザの意図に対応する動作を実行する精度を改善することにさらに役立ち、ユーザ体験を改善する。

【0041】

本出願の実施形態で提供される人間とコンピュータとの相互作用方法は、図2に示されている人間とコンピュータとの対話システムに適用され得る。人間とコンピュータとの対話システムは、電子デバイス100と、1つ以上のサーバ200（例えば、チャットボット）とを含む。電子デバイス100は、電気通信ネットワーク（3G／4G／5G通信ネットワークなどの通信ネットワーク）またはWIFIネットワークなどを介してサーバ200に接続される。これは、本出願の実施形態では限定されない。

【0042】

ユーザは、電子デバイス100を使用してサーバ200との間で人間とコンピュータとの対話を行い得る。電子デバイス100は、携帯電話、タブレットコンピュータ、パーソナルコンピュータ（Personal Computer、PC）、携帯情報端末（personal digital assistant、PDA）、スマートウォッチ、ネットブック、ウェアラブル電子デバイス、拡張現実（Augmented Reality、AR）デバイス、仮想現実（Virtual Reality、VR）デバイス、車載デバイス、スマート自動車、およびスマートスピーカなどであり得る。電子デバイス100の具体的な形態は、本出願では特に限定されない。

【0043】

サーバ200は、電子デバイス100に人間とコンピュータとの対話サービスを提供し得、ユーザの要求を把握し、対応するサービスをユーザに提供するために、電子デバイスによって入力されたユーザの発話に基づいてユーザの意図を認識し得る。サーバ200は、電子デバイス100のベンダのサーバであってもよく、例えば、電子デバイス100上の音声アシスタントのクラウドサーバであってもよい。あるいは、サーバ300は、別のアプリケーシ

ョンに対応するサーバであってもよい。これは、本出願の実施形態では限定されない。

【0044】

一部の実施形態では、サーバ200は、1つ以上の第三者アプリケーションに対応するサーバ300への通信接続をさらに確立し得、これにより、サーバ200は、ユーザの要求の把握後に、対応する第三者アプリケーションに対応するサーバ300に、対応するサービス要求を送信し、第三者アプリケーションに対応するサーバ300の応答情報を電子デバイス100に返す。一部の他の実施形態では、サーバ200は、第三者アプリケーションを用いて電子デバイス400への通信接続をさらに確立し得、これにより、第三者アプリケーションの開発者または管理者は、第三者アプリケーションによって提供されるサービスを構成および管理するために、電子デバイス400を使用してサーバ200にログインする。

10

【0045】

図3は、本出願の一実施形態による別の人間とコンピュータとの対話システムのフレームワーク図である。このフレームワーク図を参照して、以下では最初に、本出願の実施形態に適用可能な人間とコンピュータとの相互作用プロセスについて簡単に説明する。

【0046】

最初に、ユーザは、電子デバイス100を介してサーバ200にユーザステートメント（音声形式またはテキスト形式であり得る）を入力し得る。ユーザステートメントが音声形式である場合、電子デバイス100が、音声形式のユーザステートメントをテキスト形式のユーザステートメントに変換し、次に、テキスト形式のユーザステートメントをサーバ200に送信してもよいし、またはサーバ200が、音声形式のユーザステートメントをテキスト形式のユーザステートメントに変換してもよい。これは、本出願の本実施形態では限定されない。

20

【0047】

サーバ200が電子デバイス100によって送信されたユーザステートメントを受信した後、サーバ200内の自然言語理解（Natural Language Understanding、NLU）モジュールが、最初にユーザステートメントの意味理解を実行する。具体的には、ユーザステートメントが自然言語理解モジュールによって処理されるとき、ユーザステートメントは、分野分類モジュール、意図分類モジュール、およびスロット抽出モジュールの3つのサブモジュールによって処理される必要がある。一般に、サーバ200は、複数の特定のタスクシナリオ、例えば、食事注文シナリオ、配車シナリオ、および気象確認シナリオを一元化している。したがって、分野分類モジュールは、最初に、ユーザステートメントが属する特定のタスクシナリオを認識し、ユーザステートメントを特定のタスクシナリオに割り当て得る。意図認識モジュールは、ユーザの意図を認識し、特定のタスクシナリオ内のサブシナリオにユーザステートメントを割り当て得る。スロット抽出モジュールは、ユーザステートメントからエンティティ（entity）を認識し、スロット記入（Slot Filling）を実行し得る。例えば、名前付きエンティティ認識（Named Entity Recognition、NER）が、人名、地名、時間、日付、機関名、組織名、または金銭的価値など、ユーザステートメント内の特定の意味を有するエンティティを認識するために使用され得る。ユーザステートメントから対応するエンティティを認識するために、ユーザステートメントから抽出された各単語の特徴が各エンティティの所定の特徴と比較されることが簡単に理解され得る。

30

40

【0048】

例えば、図1Bに示されている対話が、説明のための例として引き続き使用される。分野分類モジュールは、ユーザの「私に車を呼んで」に基づいて、配車タスクがユーザのために実行される必要があると判定し得る（サブシナリオは、自家用車タスク、エクスプレスタスク、および相乗りタスクをさらに含み得る）。次に、意図分類モジュールは、ユーザの「DiDiエクスプレス」に基づいて、ユーザのためにエクスプレスタスクを実行することを決定し得る。次に、スロット抽出モジュールは、目的地スロットに対応する情報を「深セン湾公園」として抽出し、出発時間に対応する情報を「8:30」として抽出し得る。図1Bでは、ユーザは、出発地スロットに対応する情報を発話しておらず、スロット抽出モジュールは、出発地スロットに対応する情報として、ユーザによって設定されたデフォルト

50

の出発地を抽出し得るか、またはスロット抽出モジュールは、GPSを使用して測位を実行し、測位された位置を、出発地スロットに対応する情報として使用することに留意されたい。

【0049】

自然言語理解モジュールの出力は、対話管理（Dialog Management）モジュールの入力として使用される。対話管理モジュールは、状態追跡モジュールおよび対話ポリシーモジュールを含む。状態追跡モジュールは、継続的な対話の様々なタイプの情報を含み、以前の状態、ユーザ状態（自然言語理解モジュールによって出力された情報）、およびシステム状態（すなわち、データベースにおける照会に基づく）に基づいて現在の対話状態を更新する。対話ポリシーは、タスクシナリオに密接に関連しており、通常、対話管理モジュールの出力として使用され、例えば、欠落している必須スロットに関する質問メカニズムであり得る。

【0050】

本出願の実施形態では、対話ポリシーは、欠落している任意キースロットに関する確認メカニズムをさらに含む。具体的には、欠落している任意キースロットに関する確認メカニズムおよび欠落している必須スロットに関する質問メカニズムは、並行してまたは連続して実行され得る。言い換えれば、確認メカニズムおよび質問メカニズムの実行順序は、本出願の実施形態では限定されない。具体的な確認メカニズムは、以下の実施形態で詳細に説明される予定であり、ここでは詳細は説明されない。

【0051】

自然言語生成（Natural Language Generation、NLG）モジュールは、対話管理モジュールの出力に基づいてテキスト情報を生成し、テキスト情報をユーザにフィードバックする、すなわち、ユーザとの間で人間とコンピュータとの相互作用プロセスを遂行する。自然言語生成モジュールは、テンプレートベース、シンタックスベース、またはモデルベースの方法などで自然言語を生成し得る。テンプレートベース方法およびシンタックスベース方法では、規則ベースのポリシーが主に使用され、モデルベース方法では、例えば、長・短期記憶（Long Short-Term Memory、LSTM）が使用され得る。自然言語を生成する具体的な実施態様は、本出願の実施形態では限定されない。

【0052】

図4は、電子デバイス100の概略構造図である。

【0053】

電子デバイス100は、プロセッサ110、外部メモリインターフェース120、内部メモリ121、ユニバーサルシリアルバス（universal serial bus、USB）ポート130、充電管理モジュール140、電力管理モジュール141、バッテリー142、アンテナ1、アンテナ2、移動通信モジュール150、ワイヤレス通信モジュール160、音声モジュール170、スピーカ170A、受話器170B、マイクロフォン170C、ヘッドセットジャック170D、センサモジュール180、ボタン190、モータ191、インジケータ192、カメラ193、ディスプレイ194、および加入者識別モジュール（subscriber identification module、SIM）カードインターフェース195などを含み得る。センサモジュール180は、圧力センサ180A、ジャイロスコープセンサ180B、気圧センサ180C、磁気センサ180D、加速度センサ180E、距離センサ180F、光学式近接センサ180G、指紋センサ180H、温度センサ180J、タッチセンサ180K、周囲光センサ180L、および骨伝導センサ180Mなどを含み得る。

【0054】

本出願の実施形態に示されている構造は、電子デバイス100に対する特定の限定を構成しないことが理解され得る。本出願の一部の他の実施形態では、電子デバイス100は、図に示されているものよりも多いもしくは少ない構成要素を含み得るし、または一部の構成要素は組み合わせられ得るし、または一部の構成要素は分割され得るし、または異なる構成要素の配置が使用され得る。図に示されている構成要素は、ハードウェア、ソフトウェア、またはソフトウェアとハードウェアとの組み合わせによって実施され得る。

【0055】

プロセッサ110は、1つ以上の処理ユニットを含み得る。例えば、プロセッサ110は、アプリケーションプロセッサ (application processor、AP)、モデムプロセッサ、グラフィック処理ユニット (graphics processing unit、GPU)、画像信号プロセッサ (image signal processor、ISP)、コントローラ、メモリ、映像コーデック、デジタル信号プロセッサ (digital signal processor、DSP)、ベースバンドプロセッサ、および／またはニューラル処理ユニット (neural processing unit、NPU) を含み得る。異なる処理ユニットは、独立したデバイスであってもよいし、または1つ以上のプロセッサに統合されてもよい。

【0056】

コントローラは、電子デバイス100の中核およびコマンドセンタであり得る。コントローラは、命令を読み出して命令を実行するように制御するために、命令動作コードおよび時系列信号に基づいて動作制御信号を生成し得る。

【0057】

メモリは、プロセッサ110にさらに配置されてもよく、命令およびデータを記憶するように構成される。一部の実施形態では、プロセッサ110内のメモリは、キャッシュメモリである。このメモリは、プロセッサ110によって使用されたばかりの、または周期的に使用される命令またはデータを記憶し得る。プロセッサ110が命令またはデータを再び使用する必要がある場合、プロセッサ110は、繰り返しアクセスを回避するために、メモリから命令またはデータを直接呼び出し得る。これは、プロセッサ110の待ち時間を短縮し、システム効率を改善する。

【0058】

一部の実施形態では、プロセッサ110は、1つ以上のインターフェースを含み得る。インターフェースは、集積回路間 (inter-integrated circuit、I2C) インターフェース、集積回路間音響 (inter-integrated circuit sound、I2S) インターフェース、パルス符号変調 (pulse code modulation、PCM) インターフェース、汎用非同期送受信機 (universal asynchronous receiver/transmitter、UART) インターフェース、モバイルインダストリプロセッサインターフェース (mobile industry processor interface、MIPI)、汎用入出力 (general-purpose input/output、GPIO) インターフェース、加入者識別モジュール (subscriber identity module、SIM) インターフェース、および／またはユニバーサルシリアルバス (universal serial bus、USB) ポートなどを含み得る。

【0059】

I2Cインターフェースは、双方向同期シリアルバスであり、シリアルデータライン (serial data line、SDA) およびシリアルクロックライン (serial clock line、SCL) を含む。一部の実施形態では、プロセッサ110は、I2Cバスの複数のグループを含み得る。プロセッサ110は、異なるI2Cバスインターフェースを介してタッチセンサ180K、充電器、フラッシュライト、およびカメラ193などに別個に結合され得る。例えば、プロセッサ110は、I2Cインターフェースを介してタッチセンサ180Kに結合され得、これにより、プロセッサ110は、電子デバイス100のタッチ機能を実施するために、I2Cバスインターフェースを介してタッチセンサ180Kと通信する。

【0060】

I2Sインターフェースは、音声通信を実行するように構成され得る。一部の実施形態では、プロセッサ110は、I2Sバスの複数のグループを含み得る。プロセッサ110は、プロセッサ110と音声モジュール170との間の通信を実施するために、I2Sバスを介して音声モジュール170に結合され得る。一部の実施形態では、音声モジュール170は、ブルートゥースヘッドセットを介して通話に応答する機能を実施するために、I2Sインターフェースを介してワイヤレス通信モジュール160に音声信号を送信し得る。

【0061】

PCMインターフェースもまた、音声通信を実行し、アナログ信号をサンプリング、量子化、および符号化するように構成され得る。一部の実施形態では、音声モジュール170は

10

20

30

40

50

、PCMバスインターフェースを介してワイヤレス通信モジュール160に結合され得る。一部の実施形態では、音声モジュール170は、あるいは、ブルートゥースヘッドセットを介して通話に応答する機能を実施するためにPCMインターフェースを介してワイヤレス通信モジュール160に音声信号を送信してもよい。I2SインターフェースとPCMインターフェースのどちらも、音声通信を実行するように構成され得る。

【0062】

UARTインターフェースは、ユニバーサルシリアルデータバスであり、非同期通信を実行するように構成される。このバスは、双方向通信バスであり得、シリアル通信とパラレル通信との間で送信データを変換する。一部の実施形態では、UARTインターフェースは通常、プロセッサ110とワイヤレス通信モジュール160とを接続するように構成される。例えば、プロセッサ110は、ブルートゥース機能を実施するために、UARTインターフェースを介してワイヤレス通信モジュール160内のブルートゥースモジュールと通信する。一部の実施形態では、音声モジュール170は、ブルートゥースヘッドセットを介して音楽を再生する機能を実施するために、UARTインターフェースを介してワイヤレス通信モジュール160に音声信号を送信し得る。

【0063】

MIPIインターフェースは、プロセッサ110と、ディスプレイ194またはカメラ193などの周辺デバイスとを接続するように構成され得る。MIPIインターフェースは、カメラシリアルインターフェース(camera serial interface、CSI)およびディスプレイシリアルインターフェース(display serial interface、DSI)などを含む。一部の実施形態では、プロセッサ110は、電子デバイス100の撮影機能を実施するために、CSIインターフェースを介してカメラ193と通信する。プロセッサ110は、電子デバイス100の表示機能を実施するために、DSIインターフェースを介してディスプレイ194と通信する。

【0064】

GPIOインターフェースは、ソフトウェアを使用して構成され得る。GPIOインターフェースは、制御信号として構成され得るし、またはデータ信号として構成され得る。一部の実施形態では、GPIOインターフェースは、プロセッサ110と、カメラ193、ディスプレイ194、ワイヤレス通信モジュール160、音声モジュール170、およびセンサモジュール180などとを接続するように構成され得る。GPIOインターフェースは、あるいは、I2Cインターフェース、I2Sインターフェース、UARTインターフェース、またはMIPIインターフェースなどとして構成されてもよい。

【0065】

USBポート130は、USB標準仕様に準拠したポートであり、具体的には、Mini USBポート、Micro USBポート、またはUSB Type-Cポートなどであり得る。USBインターフェース130は、電子デバイス100を充電するために充電器に接続されるように構成され得るし、または電子デバイス100と周辺デバイスとの間でデータを送信するように構成され得るし、またはヘッドセットを介して音声を再生するためにヘッドセットに接続されるように構成され得るし、またはARデバイスなどの別の電子デバイスに接続されるように構成され得る。

【0066】

本発明の実施形態に示されているモジュール間のインターフェース接続関係は、説明のための例として使用されているにすぎず、電子デバイス100の構造に対する限定を構成しないことが理解され得る。本出願の一部の他の実施形態では、電子デバイス100は、あるいは、前述の実施形態とは異なるインターフェース接続方式、または複数のインターフェース接続方式の組み合わせを使用してもよい。

【0067】

充電管理モジュール140は、充電器から充電入力を受け取るように構成される。充電器は、ワイヤレス充電器または有線充電器であり得る。有線充電が使用される一部の実施形態では、充電管理モジュール140は、USBポート130を介して有線充電器から充電入力を受け取り得る。ワイヤレス充電が使用される一部の実施形態では、充電管理モジュール14

10

20

30

40

50

0は、電子デバイス100のワイヤレス充電コイルを介してワイヤレス充電入力を受け取り得る。充電管理モジュール140は、バッテリー142が充電されているときに、電力管理モジュール141を介して電子デバイスに電力をさらに供給し得る。

【0068】

電力管理モジュール141は、バッテリー142および充電管理モジュール140とプロセッサ110とを接続するように構成される。電力管理モジュール141は、バッテリー142および／または充電管理モジュール140から入力を受け取り、プロセッサ110、内部メモリ121、外部メモリ、ディスプレイ194、カメラ193、およびワイヤレス通信モジュール160などに電力を供給する。電力管理モジュール141は、バッテリー電力、バッテリーサイクル数、およびバッテリーの健全状態（漏電またはインピーダンス）などのパラメータを監視するようにさらに構成され得る。一部の他の実施形態では、電力管理モジュール141は、あるいは、プロセッサ110に配置されてもよい。一部の他の実施形態では、電力管理モジュール141および充電管理モジュール140は、あるいは、同じデバイスに配置されてもよい。

10

【0069】

電子デバイス100のワイヤレス通信機能は、アンテナ1、アンテナ2、移動通信モジュール150、ワイヤレス通信モジュール160、モデムプロセッサ、およびベースバンドプロセッサなどによって実施され得る。

【0070】

アンテナ1およびアンテナ2はそれぞれ、電磁波信号を送受信するように構成される。電子デバイス100の各アンテナは、1つ以上の通信帯域をカバーするように構成され得る。アンテナ利用を改善するために、異なるアンテナがさらに多重化されてもよい。例えば、アンテナ1は、ワイヤレスローカルエリアネットワークにおいてダイバーシティアンテナとして多重化されてもよい。一部の他の実施形態では、アンテナは、チューニングスイッチと組み合わせて使用され得る。

20

【0071】

移動通信モジュール150は、電子デバイス100に適用される、例えば2G／3G／4G／5Gワイヤレス通信を含むワイヤレス通信の解決策を提供し得る。移動通信モジュール150は、少なくとも1つのフィルタ、スイッチ、電力増幅器、および低ノイズ増幅器（low noise amplifier、LNA）などを含み得る。移動通信モジュール150は、アンテナ1を介して電磁波を受信し、受信された電磁波に対してフィルタリングまたは増幅などの処理を実行し、この電磁波を復調のためにモデムプロセッサに送信し得る。移動通信モジュール150はさらに、モデムプロセッサによって変調された信号を増幅し、増幅された信号を、放射のためにアンテナ1を介して電磁波に変換し得る。一部の実施形態では、移動通信モジュール150内の少なくとも一部の機能モジュールは、プロセッサ110に配置され得る。一部の実施形態では、移動通信モジュール150内の少なくとも一部の機能モジュールは、プロセッサ110内の少なくとも一部のモジュールと同じデバイスに配置され得る。

30

【0072】

モデムプロセッサは、変調器および復調器を含み得る。変調器は、送信される低周波ベースバンド信号を中高周波信号に変調するように構成される。復調器は、受信された電磁波信号を低周波ベースバンド信号に復調するように構成される。次に、復調器は、復調によって取得された低周波ベースバンド信号を処理のためにベースバンドプロセッサに送信する。低周波ベースバンド信号は、ベースバンドプロセッサによって処理され、次にアプリケーションプロセッサに転送される。アプリケーションプロセッサは、音声デバイス（スピーカ170Aまたは受話器170Bなどに限定されない）を介して音響信号を出力したり、またはディスプレイ194を介して画像もしくは映像を表示したりする。一部の実施形態では、モデムプロセッサは、独立したデバイスであり得る。一部の他の実施形態では、モデムプロセッサは、プロセッサ110から独立していてもよく、移動通信モジュール150または別の機能モジュールと同じデバイスに配置される。

40

【0073】

ワイヤレス通信モジュール160は、電子デバイス100に適用される例えばワイヤレスロ

50

ーカルエリアネットワーク (wireless local area network、WLAN) (例えば、ワイヤレスフィデリティ (wireless fidelity、Wi-Fi) ネットワーク)、ブルートゥース (bluetooth、BT)、全地球航法衛星システム (global navigational satellite system、GNSS)、周波数変調 (frequency modulation、FM)、近距離無線通信 (near field communication、NFC) 技術、または赤外線 (infrared、IR) 技術を含むワイヤレス通信の解決策を提供し得る。ワイヤレス通信モジュール160は、少なくとも1つの通信処理モジュールを統合した1つ以上のデバイスであり得る。ワイヤレス通信モジュール160は、アンテナ2を介して電磁波信号を受信し、電磁波信号に対して周波数変調およびフィルタリング処理を実行し、処理された信号をプロセッサ110に送信する。ワイヤレス通信モジュール160はさらに、プロセッサ110から送信信号を受信し、この信号に対して周波数変調および増幅を実行し、処理された信号を、放射のためにアンテナ2を介して電磁波に変換し得る。

10

【0074】

一部の実施形態では、電子デバイス100において、アンテナ1は、移動通信モジュール150に結合され、アンテナ2は、ワイヤレス通信モジュール160に結合され、これにより、電子デバイス100は、ワイヤレス通信技術を使用してネットワークおよび別のデバイスと通信し得る。ワイヤレス通信技術は、グローバル・システム・フォー・モバイル・コミュニケーションズ (global system for mobile communications、GSM)、汎用パケット無線サービス (general packet radio service、GPRS)、符号分割多元接続 (code division multiple access、CDMA)、広帯域符号分割多元接続 (wideband code division multiple access、WCDMA)、時分割同期符号分割多元接続 (time-division synchronous code division multiple access、TD-SCDMA)、ロングタームエボリューション (long term evolution、LTE)、BT、GNSS、WLAN、NFC、FM、および／またはIR技術などを含み得る。GNSSは、全地球測位システム (global positioning system、GPS)、全地球航法衛星システム (global navigation satellite system、GLONASS)、北斗衛星導航システム (beidou navigation satellite system、BDS)、準天頂衛星システム (quasi-zenith satellite system、QZSS)、および／または静止衛星型衛星航法補強システム (satellite based augmentation systems、SBAS) を含み得る。

20

【0075】

電子デバイス100は、GPU、ディスプレイ194、およびアプリケーションプロセッサなどによって表示機能を実施する。GPUは、画像処理のためのマイクロプロセッサであり、ディスプレイ194およびアプリケーションプロセッサに接続される。GPUは、数学的および幾何学的計算を実行し、グラフィックレンダリングを実行するように構成される。プロセッサ110は、表示情報を生成または変更するためにプログラム命令を実行する1つ以上のGPUを含み得る。

30

【0076】

ディスプレイ194は、画像および映像などを表示するように構成される。ディスプレイ194は、表示パネルを含む。表示パネルは、液晶ディスプレイ (liquid crystal display、LCD)、有機発光ダイオード (organic light-emitting diode、OLED)、アクティブマトリックス有機発光ダイオード (active-matrix organic light emitting diode、AMOLED)、フレキシブル発光ダイオード (flexible light-emitting diode、FLED)、Miniled、MicroLed、Micro-oLed、または量子ドット発光ダイオード (quantum dot light emitting diode、QLED) などであり得る。一部の実施形態では、電子デバイス100は、1つまたはN個のディスプレイ194を含んでもよく、Nは、1よりも大きい正の整数である。

40

【0077】

電子デバイス100は、ISP、カメラ193、映像コーデック、GPU、ディスプレイ194、およびアプリケーションプロセッサなどによって撮影機能を実施し得る。

【0078】

ISPは、カメラ193によってフィードバックされたデータを処理するように構成される

50

。例えば、撮影時、シャッタが押され、光がレンズを介してカメラの受光素子に伝達される。カメラの受光素子は、光信号を電気信号に変換し、この電気信号を処理のためにISPに送信する。ISPは、電気信号を眼に知覚可能な画像に変換する。ISPは、画像のノイズ、明るさ、および色合いに対してアルゴリズム最適化をさらに実行し得る。ISPは、撮影シナリオの露出および色温度などのパラメータをさらに最適化し得る。一部の実施形態では、ISPは、カメラ193に配置され得る。

【0079】

カメラ193は、静止画像または映像を取り込むように構成される。物体の光学画像が、レンズを介して生成され、受光素子に投影される。受光素子は、電荷結合素子 (charge coupled device、CCD) または相補型金属酸化膜半導体金 (complementary metal-oxide-semiconductor、CMOS) 光電トランジスタであり得る。受光素子は、光信号を電気信号に変換し、次に電気信号をISPに送信する。ISPは、電気信号をデジタル画像信号に変換し、デジタル画像信号を処理のためにDSPに出力する。DSPは、デジタル画像信号を、RGBまたはYUVなどのフォーマットの標準画像信号に変換する。一部の実施形態では、電子デバイス100は、1つまたはN個のカメラ193を含み得、Nは、1よりも大きい正の整数である。

【0080】

デジタル信号プロセッサは、デジタル信号を処理するように構成される。デジタル画像信号を処理することに加えて、デジタル信号プロセッサは、別のデジタル信号をさらに処理し得る。例えば、電子デバイス100が周波数を選択すると、デジタル信号プロセッサは、周波数エネルギーに対してフーリエ変換を実行するように構成される。

【0081】

映像コーデックは、デジタル映像を圧縮または解凍するように構成される。電子デバイス100は、1つ以上の映像コーデックをサポートし得る。したがって、電子デバイス100は、複数の符号化フォーマット、例えば、MPEG (moving picture experts group、MPEG) -1、MPEG-2、MPEG-3、およびMPEG-4で映像を再生または記録し得る。

【0082】

NPUは、ニューラルネットワーク (neural-network、NN) コンピューティングプロセッサである。NPUは、生物学的ニューラルネットワーク構造を参照することによって、例えば、人間の脳のニューロン間の伝達のモードを参照することによって、入力情報を迅速に処理し、さらに自己学習を継続的に実行し得る。電子デバイス100は、NPUによって画像認識、顔認識、音声認識、およびテキスト理解などのインテリジェントな認知を実施し得る。

【0083】

外部メモリインターフェース120は、電子デバイス100の記憶容量を拡張するために、Micro SDカードなどの外部メモリカードに接続されるように構成され得る。外部メモリカードは、データ記憶機能を実施するために、例えば、音楽および映像などのファイルを外部メモリカードに記憶するために外部メモリインターフェース120を介してプロセッサ110と通信する。

【0084】

内部メモリ121は、コンピュータ実行可能プログラムコードを記憶するように構成され得、実行可能プログラムコードは命令を含む。プロセッサ110は、電子デバイス100の様々な機能アプリケーションおよびデータ処理を実施するために、内部メモリ121に記憶された命令を実行する。内部メモリ121は、プログラム記憶領域およびデータ記憶領域を含み得る。プログラム記憶領域は、オペレーティングシステム、および少なくとも1つの機能 (例えば、音響再生機能または画像再生機能) によって必要とされるアプリケーションなどを記憶し得る。データ記憶領域は、電子デバイス100の使用時に作成されたデータ (音声データおよびアドレス帳など)などを記憶し得る。さらに、内部メモリ121は、高速ランダムアクセスメモリを含み得るし、または不揮発性メモリ、例えば、少なくとも1つの磁気ディスク記憶デバイス、フラッシュメモリデバイス、もしくはユニバーサルフラッ

10

20

30

40

50

シュストレージ (universal flash storage、UFS) をさらに含み得る。

【0085】

電子デバイス100は、音声モジュール170、スピーカ170A、受話器170B、マイクロフォン170C、ヘッドセットジャック170D、およびアプリケーションプロセッサなどによって音楽再生および録音機能などの音声機能を実施し得る。

【0086】

音声モジュール170は、出力のためにデジタル音声情報をアナログ音声信号に変換するように構成され、アナログ音声入力をデジタル音声信号に変換するようにも構成される。音声モジュール170は、音声信号の符号化および復号化を実行するようにさらに構成され得る。一部の実施形態では、音声モジュール170は、プロセッサ110に配置され得るし、
10

【0087】

「ホーン」とも呼ばれるスピーカ170Aは、音声電気信号を音響信号に変換するように構成される。電子デバイス100は、スピーカ170Aを介して音楽を聴いたり、またはハンズフリー通話に応答したりし得る。

【0088】

「イヤピース」とも呼ばれる受話器170Bは、音声電気信号を音響信号に変換するように構成される。電子デバイス100が通話に応答したり、または音声情報を受信したりするとき、受話器170Bは、音声を聴くために人間の耳の近くに配置され得る。「マイク」または「マイクロフォン」とも呼ばれるマイクロフォン170Cは、音響信号を電気信号に変換するように構成される。通話を行うときまたは音声メッセージを送信するとき、ユーザは、マイクロフォン170Cに音響信号を入力するために、マイクロフォン170Cの近くで声を出し得る。少なくとも1つのマイクロフォン170Cが、電子デバイス100に配置され得る。一部の他の実施形態では、2つのマイクロフォン170Cが、音響信号を収集し、ノイズ低減機能を実施するために電子デバイス100に配置され得る。一部の他の実施形態では、3つ、4つ、またはこれより多くのマイクロフォン170Cが、あるいは、音響信号を収集し、ノイズを低減し、音源を識別し、指向性録音機能を実施するなどのために、電子デバイス100に配置されてもよい。ヘッドセットジャック170Dは、有線ヘッドセットに接続されるように構成される。ヘッドセットジャック170Dは、USBポート130であってもよいし、
20
30
または3.5mmオープンモバイル電子デバイスプラットフォーム (open mobile terminal platform、OMTP) 標準インターフェースもしくは米国セルラー通信工業会 (cellular telecommunications industry association of the USA、CTIA) 標準インターフェースであってもよい。

【0089】

圧力センサ180Aは、圧力信号を検知するように構成され、圧力信号を電気信号に変換し得る。一部の実施形態では、圧力センサ180Aは、ディスプレイ194上に配置され得る。抵抗式圧力センサ、誘導式圧力センサ、および静電容量式圧力センサなど、多くのタイプの圧力センサ180Aが存在する。静電容量式圧力センサは、導電性材料で作られた少なくとも2つの平行なプレートを含み得る。圧力センサ180Aに力が加わると、電極間の静電容量が変化する。電子デバイス100は、静電容量の変化に基づいて圧力強度を判定する。
40
ディスプレイ194上でタッチ操作が実行されると、電子デバイス100は、圧力センサ180Aによってタッチ操作の強度を検出する。電子デバイス100は、圧力センサ180Aの検出信号に基づいてタッチ位置も計算し得る。一部の実施形態では、同じタッチ位置で実行されるが、異なるタッチ操作強度を有するタッチ操作は、異なる操作命令に対応し得る。例えば、タッチ操作強度が第1の圧力閾値未満であるタッチ操作が、メッセージのアイコンに対して実行されると、SMSを見るための命令が実行される。タッチ操作強度が第1の圧力閾値以上であるタッチ操作が、メッセージのアイコンに対して実行されると、新しいSMSメッセージを作成するための命令が実行される。

【0090】

ジャイロ스코プセンサ180Bは、電子デバイス100の動作姿勢を判定するように構成
50

され得る。一部の実施形態では、3つの軸（すなわち、軸x、y、およびz）の周りの電子デバイス100の角速度が、ジャイロスコープセンサ180Bによって判定され得る。ジャイロスコープセンサ180Bは、写真撮影時の画像安定化のために使用され得る。例えば、シャッターが押されたとき、ジャイロスコープセンサ180Bは、画像安定化を実施するために、電子デバイス100がぶれた角度を検出し、この角度に基づいて、レンズモジュールが補償する必要がある距離を計算し、レンズが逆方向の動作によって電子デバイス100のぶれを除去することを可能にする。ジャイロスコープセンサ180Bは、ナビゲーションシナリオおよび身体ゲームシナリオでさらに使用され得る。

【0091】

気圧センサ180Cは、気圧を測定するように構成される。一部の実施形態では、電子デバイス100は、測位およびナビゲーションを支援するために、測定により気圧センサ180Cによって取得された気圧値に基づいて高度を計算する。

【0092】

磁気センサ180Dは、ホールセンサを含む。電子デバイス100は、磁気センサ180Dによってフリップカバーの開閉を検出し得る。一部の実施形態では、電子デバイス100が折り畳み式携帯電話である場合、電子デバイス100は、フリップカバーの検出された開閉状態に基づく、フリッピングによる自動ロック解除などの特徴を設定するために、磁気センサ180Dによってフリップカバーの開閉を検出し得る。

【0093】

加速度センサ180Eは、電子デバイス100の様々な方向（通常は3軸上の）の加速度の大きさを検出し得、電子デバイス100が静止しているときの重力の大きさおよび方向を検出し得る。加速度センサ180Eは、電子デバイスの姿勢を識別するようにさらに構成され得、歩数計および横向きモードと縦向きモードとの切り替えなどのアプリケーションに適用される。

【0094】

距離センサ180Fは、距離を測定するように構成される。電子デバイス100は、赤外線方式またはレーザ方式で距離を測定し得る。一部の実施形態では、撮影シナリオにおいて、電子デバイス100は、迅速な焦点合わせを実施するために距離センサ180Fによって距離を測定し得る。

【0095】

光学式近接センサ180Gは、例えば、発光ダイオード（LED）、およびフォトダイオードなどの光検出器を含み得る。発光ダイオードは、赤外線発光ダイオードであり得る。電子デバイス100は、発光ダイオードによって赤外光を発する。電子デバイス100は、フォトダイオードによって近傍の物体からの赤外反射光を検出する。十分な反射光を検出したとき、電子デバイス100は、電子デバイス100の近くに物体が存在すると判定し得る。不十分な反射光を検出したとき、電子デバイス100は、電子デバイス100の近くに物体が存在しないと判定し得る。電子デバイス100は、光学式近接センサ180Gによって、ユーザが通話を行うために電子デバイス100を耳の近くに保持していることを検出し得る。この場合、電子デバイス100は、省電力のために画面を自動的にオフにする。光学式近接センサ180Gはまた、画面のロック解除またはロックを自動的に実行するために、スマートカバーモードまたはポケットモードで使用され得る。

【0096】

周囲光センサ180Lは、周囲光の明るさを検知するように構成される。電子デバイス100は、検知された周囲光の明るさに基づいてディスプレイ194の明るさを適応的に調整し得る。周囲光センサ180Lは、撮影時にホワイトバランスを自動的に調整するようにさらに構成され得る。周囲光センサ180Lは、偶発的なタッチを回避する目的で電子デバイス100がポケット内にあるかどうかを検出するために光学式近接センサ180Gと協働し得る。

【0097】

指紋センサ180Hは、指紋を収集するように構成される。電子デバイス100は、収集された指紋の特徴を使用して、指紋ベースのロック解除、アプリケーションロックアクセス

10

20

30

40

50

、指紋ベースの撮影、および指紋ベースの通話応答などを実施し得る。

【0098】

温度センサ180Jは、温度を検出するように構成される。一部の実施形態では、電子デバイス100は、温度センサ180Jによって検出された温度に基づいて温度処理ポリシーを実行する。例えば、温度センサ180Jによって報告された温度が閾値を超えたとき、電子デバイス100は、熱保護の目的で消費電力を低減するために、温度センサ180Jの近傍のプロセッサのパフォーマンスを低下させる。一部の他の実施形態では、温度が別の閾値未満であるとき、電子デバイス100は、低温に起因して電子デバイス100が異常に電源オフされるのを防止するために、バッテリー142を加熱する。一部の他の実施形態では、温度がさらに別の閾値未満であるとき、電子デバイス100は、低温によって発生する異常な電源オフを防止するためにバッテリー142の出力電圧を上げる。

10

【0099】

タッチセンサ180Kは、「タッチパネル」とも呼ばれる。タッチセンサ180Kは、ディスプレイ194上に配置され得る。タッチセンサ180Kおよびディスプレイ194は、「タッチスクリーン」とも呼ばれるタッチスクリーンを形成する。タッチセンサ180Kは、タッチセンサ180K上またはその近くで実行されたタッチ操作を検出するように構成される。タッチセンサは、タッチイベントのタイプを判定するために、検出されたタッチ操作をアプリケーションプロセッサに転送し得る。ディスプレイ194は、タッチ操作に関連する視覚的出力を提供する。一部の他の実施形態では、タッチセンサ180Kは、あるいは、ディスプレイ194の位置とは異なる位置で電子デバイス100の表面上に配置されてもよい。

20

【0100】

骨伝導センサ180Mは、振動信号を取得し得る。一部の実施形態では、骨伝導センサ180Mは、人間の声帯部分の振動骨の振動信号を取得し得る。骨伝導センサ180Mはまた、血圧拍動信号を受信するために身体の脈に接触し得る。一部の実施形態では、骨伝導センサ180Mは、あるいは、骨伝導ヘッドセットを構成するためにヘッドセットに配置されてもよい。音声モジュール170は、音声機能を実施するために、骨伝導センサ180Mによって取得された、声帯部分の振動骨の振動信号に基づく解析によって音声信号を取得し得る。アプリケーションプロセッサは、心拍検出機能を実施するために、骨伝導センサ180Mによって取得された血圧拍動信号に基づいて心拍情報を解析し得る。

【0101】

ボタン190は、電源ボタンおよび音量ボタンなどを含む。ボタン190は、機械式ボタンであってもよいし、またはタッチボタンであってもよい。電子デバイス100は、ボタン入力を受け取り、電子デバイス100のユーザ設定および機能制御に関連するボタン信号入力を生成し得る。

30

【0102】

モータ191は、振動プロンプトを生成し得る。モータ191は、着信振動プロンプトに使用され得るし、またはタッチ振動フィードバックに使用され得る。例えば、異なるアプリケーション（例えば、撮影および音声再生）に対して実行されるタッチ操作は、異なる振動フィードバック効果に対応し得る。ディスプレイ194上の異なる領域に対して実行されるタッチ操作は、モータ191の異なる振動フィードバック効果に対応し得る。異なる利用シナリオ（例えば、時間リマインダシナリオ、情報受信シナリオ、アラームクロックシナリオ、およびゲームシナリオ）も、異なる振動フィードバック効果に対応し得る。タッチ振動フィードバック効果は、あるいは、カスタマイズされてもよい。

40

【0103】

インジケータ192は、充電状態および電源変化を示すために使用され得る、またはメッセージ、不在着信、および通知などを示すために使用され得るインジケータライトであり得る。

【0104】

SIMカードインターフェース195は、SIMカードに接続されるように構成される。SIMカードは、電子デバイス100との接触または電子デバイス100からの分離を実施するため

50

に、SIMカードインターフェース195に挿入され得る、またはSIMカードインターフェース195から取り外され得る。電子デバイス100は、1つまたはN個のSIMカードインターフェースをサポートし得、Nは、1よりも大きい正の整数である。SIMカードインターフェース195は、ナノSIMカード、マイクロSIMカード、およびSIMカードなどをサポートし得る。複数のカードが、1つのSIMカードインターフェース195に同時に挿入され得る。複数のカードは、同じタイプまたは異なるタイプのものであり得る。SIMカードインターフェース195はまた、様々なタイプのSIMカードと互換性があり得る。SIMカードインターフェース195はまた、外部メモリカードと互換性があり得る。電子デバイス100は、通話およびデータ通信などの機能を実施するためにSIMカードを介してネットワークと相互作用する。一部の実施形態では、電子デバイス100は、eSIM、すなわち組み込みSIMカードを使用する。eSIMカードは、電子デバイス100に組み込まれ得、電子デバイス100から分離され得ない。

10

【0105】

電子デバイス400の構造については、図4に示されている電子デバイス100の構造を参照されたい。詳細は再び説明されない。

【0106】

以下の実施形態におけるすべての技術的解決策は、前述のハードウェアアーキテクチャを有する電子デバイス100、サーバ200、または電子デバイス400で実施され得る。

【0107】

本出願の実施形態で提供される技術的解決策は、添付の図面を参照して以下で詳細に説明される。

20

【0108】

最初に、スキル開発者（第三者アプリケーションの開発者またはサービスプロバイダなどであり得る）は、新しいサービスを構成するために、電子デバイス400を使用してサーバ200にログインし得る。サービスは、あるいは、人間とコンピュータとの対話プラットフォーム上のスキルと呼ばれ得る。

【0109】

図5から図8は、スキル開発者によって新しいスキルを構成するプロセスで使用されるいくつかのインターフェースを示している。新しいスキルを構成するために、スキル開発者は、以下のステップを実行する必要がある。

30

【0110】

1. 新しいスキルの基本情報を設定する。

【0111】

スキル開発者は、電子デバイス400を介して人間とコンピュータとの対話プラットフォームのスキル管理ページにログインし、新しいスキルの構成を開始し得る。例えば、スキル開発者は、スキル管理ページにログインするために、電子デバイス400のブラウザにおいて、人間とコンピュータとの対話プラットフォームに関連するURLを入力し得る。あるいは、スキル開発者は、対応するアプリを介してスキル管理ページにログインしてもよい。

【0112】

図5Aは、スキル管理ページのホームページ400を示している。ホームページ400は、コントロール401およびコントロール402を含み得る。スキルテンプレートが、人間とコンピュータとの対話プラットフォーム上で提供される。これらのテンプレートは、いくつかの利用シナリオをカバーする。スキル開発者は、スキル開発者の個人化要求を満たすようにこれらのテンプレートを修正し得る。一部の実施形態では、スキル開発者は、人間とコンピュータとの対話プラットフォーム上で提供されているスキルテンプレートを使用して新しいスキルを構成するためにコントロール402を選択し得る。一部の他の実施形態では、スキル開発者は、音声相互作用および対応するサービスをエンドユーザに提供するために、コントロール401を選択し、スキル開発者によって提供されるサービスに基づいてユーザ定義スキルを追加し得る。以下では、スキル開発者がコントロール401を選択し、ユーザ定義スキルを追加する例を使用して説明を提供する。

40

50

【0113】

スキル開発者がコントロール401を選択したことを検出した後、電子デバイス400は、ユーザ定義スキルを追加するためのインターフェースに入る。図5Bは、新しいスキルの基本情報を設定するためのページ500を示している。新しいスキルのスキル識別子、スキル名、スキルカテゴリ、およびウェイクアップワードなどの基本情報が、ページ500上で設定され得る。

【0114】

スキル識別子は、スキルのグローバルに一意の識別子であり、各スキルのスキル識別子は繰り返し使用され得ない。スキル名は、スキル開発者が作成されたスキルを管理するのを容易にするための、スキルの記述名であり、スキル名が繰り返し使用されるかどうかについての要求はない。スキル開発者は、ユーザの発話のマッピング時に検索およびスクリーニングするために、各スキルに関してカテゴリ（上記の特定のシナリオとしても理解され得る）を選択する必要がある。各スキルは、1つのカテゴリのみに属し得る。正確なスキルカテゴリ選択は、ユーザの発話をスキルの意図に迅速かつ正確にマッピングすることに役立つ。ウェイクアップワードは、スキルの別の名前として理解され得る。ユーザが別の名前を発話した後、人間とコンピュータとの対話プラットフォームは、スキルに対応するサービスを迅速に取得し得る。

【0115】

新しいスキルの基本情報の設定は、1つずつ列挙されなかった他の内容をさらに含み得ることが理解され得る。

【0116】

2. 新しいスキルの意図を作成する。

【0117】

新しいスキルの基本情報を設定した後、スキル開発者は、新しいスキルの意図の作成を開始し得る。図5Cは、電子デバイス400に表示された意図作成ページ600を示している。ページ600は、意図名、先行コンテキスト、および後続コンテキストなどを含み得る。

【0118】

意図名は、スキルにおいて繰り返し使用され得ない。コンテキストは、複数回の対話のシナリオで主に使用される。先行コンテキスト、現在の意図をトリガするために使用され、後続コンテキストは、次の意図を関連付けるために使用される。

【0119】

ユーザの各文は、ユーザの意図に対応し、ユーザの意図は、ユーザがその文を発話する目的であることに留意されたい。各スキルは、いくつかの意図を含む。ユーザの各文は、対応するサービスが提供されるように、ユーザの要求の把握のためにスキルの意図にマッピングされる。このスキルを使用するとき、ユーザは、様々な方法で意図を表現する。したがって、スキル開発者は、意図構成時に、日常生活で意図を表現するためにユーザによって使用される表現（すなわち、ユーザの発話）をできるだけ多く入力する必要がある。このようにして、意図はより正確に認識され得る。

【0120】

3. ユーザの発話を入力し、ユーザの発話内のスロットをマークする（スロット属性の設定およびスロットタイプの関連付けを含む）。

【0121】

図6Aは、電子デバイス400に表示された意図作成ページ上のユーザ発話設定ページ601を示している。ページ601は、1つ以上のユーザ発話追加コントロール602を含み得る。ページ601は、既存のユーザの発話の情報項目603をさらに表示し得る。

【0122】

スキル開発者は、ユーザの発話を追加するために、コントロール602のテキストボックスに新しいユーザの発話を入力し、機能ボタン「追加」をクリックし得る。一部の実施形態では、人間とコンピュータとの対話プラットフォームは、追加されたユーザの発話からエンティティを自動的に認識し、認識されたエンティティとスロットおよびスロットタイ

10

20

30

40

50

ブとを関連付け得る。一部の他の実施形態では、人間とコンピュータとの対話プラットフォームがスロットを自動的にマークしないか、またはスロットを間違えてマークする場合、スキル開発者は、手動でスロットをマークし、マークされたスロットとスロットタイプとを関連付けることを選択し得る。

【0123】

スロットは、ユーザの発話の意図を表現するために使用されるキー情報であり、ユーザの発話のキーワードとして理解され得る。1つのスロットは、1つのスロットタイプに対応し、スロットの値には、このスロットタイプの別の単語が記入され得る。スロットタイプは、一分野の単語のセットとして理解され得る。ユーザの発話内の、スロットに対応する情報は、様々なスロットタイプのものであり、同じスロットタイプの単語が、対応するスロット情報から置換、認識、および抽出され得る。

10

【0124】

例えば、スキル開発者がコントロール602のテキストボックスに新しいユーザの発話を入力し、機能ボタン「追加」をクリックする操作に回答して、電子デバイス400は、図6Bに示されている対話ボックス604をポップアップ表示する。スキル開発者は、対話ボックス604で、ユーザの発話にマークされたスロットおよび関連するスロットタイプを見て修正し得る。対話ボックス604には、スロットタイプ追加コントロール605がさらに表示され得、これにより、スロットタイプ関連付け時に適切なスロットタイプが選択され得ない場合に、対応するスロットタイプが追加され得る。対話ボックス604には、スロットリスト確認コントロール606がさらに表示され得る。スキル開発者によるコントロール606のクリックに回答して、電子デバイス400は、図6Cに示されているページ608を表示する。ページ608は、ユーザの発話に含まれるスロット、スロットに関連付けられたスロットタイプ、スロットの属性（必須スロット、任意キースロット、および任意非キースロット）、ならびにスロットに対応する情報が欠落しているときに尋ねられる質問などの情報を示している。ページ608は、新スロット追加コントロール607をさらに含み得、コントロール607は、ユーザの発話内のスロットを追加するために使用され得る。

20

【0125】

一般に、ページ608において、スキル開発者は、必須スロットに関する質問メカニズムを構成し、任意キースロットに関する確認メカニズムを構成し、任意非キースロットに関しては質問を設定しない。すなわち、スロットが必須スロットまたは任意キースロットに設定される場合、そのスロットに関して質問が設定される。質問は、デフォルトの質問であってもよいし、またはスキル開発者によってカスタマイズされた質問であってもよい。スロットが任意非キースロットに設定される場合、スロットの質問は、デフォルトで「なし」に設定され、変更され得ない。別の例では、新しいユーザの発話は、「今度の金曜日に首都では雨が降るかな？」である。マークされたスロットは、時間スロットおよび都市スロットを含む。時間スロットに対応するスロットタイプは、sys. timeであり、時間スロットの属性は、任意キースロットである。具体的には、人間とコンピュータとの対話プラットフォームが、時間スロットに対応する情報の抽出に失敗した場合、人間とコンピュータとの対話プラットフォームは、能動的にユーザに「質問1」を発し、ユーザは、時間スロットに対応する情報が必要かどうかを判定する。時間スロットに対応する情報が必要な場合、人間とコンピュータとの対話プラットフォームは、ユーザの回答から時間スロットに対応する情報を抽出し、次に、その後の動作を実行する。時間スロットに対応する情報が不要な場合、人間とコンピュータとの対話プラットフォームは、時間スロットに対応する情報が存在しないと考え、その後の動作を直接実行する。

30

40

【0126】

都市スロットに対応するスロットタイプは、sys. local. cityであり、都市スロットの属性は、必須スロットである。具体的には、人間とコンピュータとの対話プラットフォームが、都市スロットに対応する情報の抽出に失敗した場合、人間とコンピュータとの対話プラットフォームは、能動的にユーザに「質問2」を発する。次に、人間とコンピュータとの対話プラットフォームは、ユーザの回答から都市スロットに対応する情報を抽出し、

50

次に、その後の動作を実行する。

【0127】

追加されたユーザの発話において、別の任意非キースロットもマークされ得る。具体的には、人間とコンピュータとの対話プラットフォームが、任意非キースロットに対応する情報の抽出に失敗した場合、人間とコンピュータとの対話プラットフォームは、任意非キースロットに対応する情報が存在しないと考え、その後の動作を直接実行する。

【0128】

本出願の実施形態では、スロットタイプは、システムスロットタイプおよびユーザ定義スロットタイプ（ユーザ辞書とも呼ばれ得る）を主に含む。システムスロットタイプは、人間とコンピュータとの対話プラットフォームによって事前設定されたスロットタイプであり、システムスロットタイプの単語は、無数にあり、例えば、sys. time、sys. location、city、sys. name、およびsys. phoneNumである。ユーザ定義スロットタイプは、スキル開発者によって定義され、ユーザ定義スロットタイプの単語の数は限られている。

【0129】

図7は、電子デバイス400に表示されたスロットタイプ編集ページ700を示している。スキル開発者は、追加されるユーザ定義スロットタイプのテキストをテキストボックス701に入力し、エンターを押し得る。スキル開発者は、値の項目702の下に、追加されるユーザ定義スロットタイプの値を入力し得、同義語の項目703の下に、追加されるユーザ定義スロットタイプに対応する同義語を入力し、次に「保存」ボタンをクリックし得る。このようにして、ユーザ定義スロットタイプが追加される。ユーザ定義スロットタイプは、スロットタイプ編集ページ700の領域704に示されている複数のコントロールを使用してさらに修正および削除などされ得る。一部の実施形態では、スロットタイプ編集ページ700は、スロットタイプのバッチ追加をさらにサポートし得る。例えば、スキル開発者は、スロットタイプをバッチで追加するために、バッチ追加ボタン705をクリックし、指定されたファイルタイプのファイルまたは指定されたファイルフォーマットのファイルをアップロードすることを選択し得る。特定のファイルタイプのファイルまたは特定のファイルフォーマットのファイルは、追加されるスロットタイプに対応する1つ以上の情報を含む。これは、本出願の本実施形態では限定されない。

【0130】

4. 新しいスキルが設定された後、新しいスキルに対応する人間とコンピュータとの対話モデルをトレーニングおよびリリースする。

【0131】

スキル開発者が、新しいスキルによって必要とされるユーザの発話を入力し、ユーザの発話内のスロットに対してマーキング、属性設定、およびスロットタイプ関連付けを実行した後、電子デバイス400は、ページ800を表示し得る。スキル開発者は、新しいスキルに対応する人間とコンピュータとの対話モデルのトレーニングを開始するように人間とコンピュータとの対話プラットフォームに通知するためにコントロール801「トレーニング開始」をクリックし得る。人間とコンピュータとの対話プラットフォームによってトレーニングされる、新しいスキルに対応する人間とコンピュータとの対話モデルは、分野分類モデル、意図分類モデル、およびスロット抽出モデルなどを含み得る。分野分類モデルは、ユーザの発話に対して分野分類を実行するために使用され得る。意図分類モデルは、ユーザの発話に対応する新しいスキルの意図を認識する目的で、ユーザの発話に対してさらなる分野分類を実行するために使用され得る。スロット抽出モデルは、ユーザの発話内のスロットに対応する情報を抽出するために使用され得る。このようにして、ユーザの意図と、スロット抽出モデルによって出力された、スロットに対応する情報とに基づいて、意図分類モデルによって出力されたユーザの意図に対応するその後の動作が実行され得る。

【0132】

人間とコンピュータとの対話プラットフォームが、新しいスキルに対応する人間とコンピュータとの対話モデルを生成した後、電子デバイス400は、ページ900を表示し得る。スキル開発者は、新しいスキルをリリースするように人間とコンピュータとの対話プラッ

10

20

30

40

50

トフォームに通知し、新しいスキルに対応する人間とコンピュータとの対話モデルを人間とコンピュータとの対話プラットフォームにプッシュするためにコントロール902「スキルをリリースする」をクリックし得る。次に、別の端末は、人間とコンピュータとの対話プラットフォームによって提供される新しいスキルを取得するために、人間とコンピュータとの対話プラットフォームとの対話を行い得る。ページ900は、コントロール901「再トレーニング」をさらに含み得、スキル開発者は、コントロール901を使用して、新しいスキルに対応する人間とコンピュータとの対話モデルを再トレーニングし得る。

【0133】

図9Aおよび図9Bは、本出願の一実施形態による人間とコンピュータとの相互作用方法を示している。本方法は、電子デバイス100とサーバ200との間の相互作用において使用され得る。本方法は、具体的には以下のステップを含む。

【0134】

S101. サーバ200は、第1の入力を受信する。

【0135】

サーバ200との対話相互作用を実行するために電子デバイス100を使用するとき、ユーザは、音声形式またはテキスト形式で、対応するサービス要求をサーバ200に転送し得る。ユーザが音声を入力する場合、サーバ200は、自動音声認識モジュールを使用して音声を認識し、音声をテキスト、すなわち第1の入力に変換し、第1の入力を自然言語理解モジュールに入力し得る。ユーザがテキストを入力する場合、サーバ200は、ユーザによって入力されたテキストを第1の入力として使用し、第1の入力を自然言語理解モジュールに入力する。

【0136】

第1の入力は、ユーザとサーバ200との間の1回の対話における1つの発話であってもよいし、またはユーザとサーバ200との間の複数回の対話における複数の発話であってもよい。これは、本出願の本実施形態では限定されない。

【0137】

S102. サーバ200は、第1の入力に対応する第1の分野を判定するために第1の入力に基づいて分野分類を実行する。

【0138】

第1の入力は、ユーザの意図に対応する。すなわち、ユーザは、サーバ200にサービスの提供またはいくつかの動作の実行を望む。自然言語理解モジュール内の分野分類モジュールは、第1の入力のユーザの意図が属する特定のタスクシナリオ（すなわち、第1の分野）を判定し、第1の入力を特定のタスクシナリオ（すなわち、第1の分野）に割り当てるために、第1の入力に基づいて検索およびスクリーニングを実行し得る。

【0139】

S103. サーバ200は、第1の入力を第1の分野に割り当て、第1の入力に対応する第1の意図を認識する。

【0140】

自然言語理解モジュール内の意図認識モジュールは、第1の入力のユーザの意図を特定のタスクシナリオのサブシナリオにさらに割り当て得る、すなわち、第1の入力に対応するユーザの意図（すなわち、第1の意図）を認識し得る。

【0141】

S104. サーバ200は、第1の意図に対応するスロット構成に基づいて、第1の意図の各スロットに対応する情報を第1の入力から抽出する。

【0142】

第1の意図は、サーバ200にあるスキルの意図である。スキルを構成するとき、スキル開発者は、スキルの第1の意図に対応するスロット、すなわち、第1の意図において抽出される必要があるスロットと、スロットの属性とを構成する。したがって、第1の入力に対応する第1の意図が判定された後、サーバ200内のスロット抽出モジュールは、第1の意図に対応するスロット構成を見つけ得る。

10

20

30

40

50

【0143】

サーバ200内のスロット抽出モジュールは、第1の入力に含まれるエンティティを認識し、スロット抽出モジュールに記憶されたスロット抽出モデルを呼び出し、これらのエンティティに対応する、第1の意図のスロットを判定する動作をこれらのエンティティに対して実行し、対応するスロットのラベルでこれらのエンティティをマークし得る。これらのエンティティが、対応するスロットの値として判定されることも考えられ得る。すなわち、これらのスロットに対応する情報が抽出される。例えば、スロット抽出モジュールは、第1の入力からエンティティAを認識し、エンティティAを、スロット抽出モデル内の各スロットに対応するアルゴリズムに入力し、エンティティAに対応する各信頼性レベルを計算する。エンティティAがスロットAに対応するアルゴリズムに入力された後に計算された信頼性レベルが事前設定された条件を満たさない場合、例えば、信頼性レベルが事前設定された閾値、例えば閾値C未満である場合、エンティティAはスロットAに対応する情報ではないと考えられる。エンティティAがスロットBに対応するアルゴリズムに入力された後に計算された信頼性レベルが事前設定された条件を満たす場合、例えば、信頼性レベルが閾値C以上である場合、エンティティAはスロットBに対応する情報であると考えられる。

10

【0144】

一部のスロットに対応する情報は、ユーザによってデフォルトで設定されてもよいし、または別の方法で取得された情報であってもよく、必ずしも第1の入力から抽出されないことに留意されたい。

20

【0145】

例えば、第1の意図は、「航空券の予約」であり、「航空券の予約」の事前設定されたスロット構成は、時間スロット、出発地スロット、および目的地スロットを含み得る。ユーザが「明日の上海への航空券を予約して」（すなわち、第1の入力）と発話した場合、サーバ200は、第1の入力から複数のエンティティ、例えば「明日」および「上海」を認識し得る。サーバ200は、計算のためにスロット抽出モデル内の時間スロットに対応するアルゴリズムに「明日」を入力して、時間スロット「明日」の信頼性レベルが事前設定された条件を満たすことを把握し得、「明日」は「航空券の予約」の時間スロットの値であると考えられ得る。言い換えれば、サーバ200は、第1の意図の時間スロットに対応する情報を抽出する。同様に、サーバ200は、計算のためにスロット抽出モデル内の目的地スロットに対応するアルゴリズムに「上海」を入力して、目的地スロット「上海」の信頼性レベルが事前設定された条件を満たすことを把握し得、「上海」は「航空券の予約」の目的地スロットの値であると考えられ得る。言い換えれば、サーバ200は、第1の意図の目的地スロットに対応する情報を抽出する。第1の入力に出発地スロットに対応するエンティティは存在しない。ユーザは、GPSを使用して電子デバイス100の現在位置を取得し得、現在位置を出発地スロットの値として使用する。あるいは、ユーザは、出発地スロットの値などとして、ユーザによってデフォルトで設定された住所を使用してもよい。言い換えれば、サーバ200は、第1の意図の出発地スロットに対応する情報を抽出する。

30

【0146】

S105. サーバ200は、第1の意図の第1のスロットに対応する情報の抽出が失敗したと判定する。

40

【0147】

ステップS104が実行されるとき、第1の入力は、第1の意図の一部のスロットに対応する情報を含まないことがある（例えば、ユーザがその情報を発話しないか、またはユーザはその情報を発話するものの、自動音声認識エラーもしくはユーザ入力エラーが発生する）。あるいは、サーバ200のスロット抽出モデルが不正確である場合、サーバ200は、第1の意図の一部のスロットに対応する情報を第1の入力から抽出することに失敗することがある。この場合、ステップS106およびその後のステップが実行される必要がある。

【0148】

S106. サーバ200は、第1のスロットの属性を判定する。第1のスロットは、必須スロ

50

ット、任意キースロット、または任意非キースロットであり得る。第1のロットが必須ロットである場合、サーバ200は、ステップS107を実行する。第1のロットが任意非キースロットである場合、サーバ200は、ステップS110を実行する。第1のロットが任意キースロットである場合、サーバ200は、ステップS111を実行する。

【0149】

具体的には、サーバ200内のロット抽出モジュールは、第1のロットに対応する情報の抽出が失敗したという結果を対話管理モジュールに送信する。対話管理モジュールは、第1のロットの属性に基づいてその後の動作を判定するために第1のロットの属性を判定する。

【0150】

S107. サーバ200は、第1のロットに対応する情報に関してユーザに質問を行う。

【0151】

対話管理モジュールは、第1のロットの属性および事前設定された対話ポリシーに基づいて、第1のロットに対応する情報に関してユーザに質問を行う。例えば、サーバ200は、ユーザに繰り返しを要求してもよいし、または以前にユーザと交わされた質問を再びユーザに発してもよいし、または欠落している第1のロットに関して質問を行ってもよい。質問内容および質問方法は、本出願の本実施形態では限定されない。

【0152】

S108. サーバ200は、第2の入力を受信する。

【0153】

第2の入力は、サーバ200の質問に対してユーザによって与えられた回答である。ユーザが音声形式で回答を与えた場合、サーバ200内の自動音声認識モジュールは、第2の入力を取得するために音声をテキストに変換し得る。ユーザがテキスト形式で回答を与えた場合、サーバ200は、ユーザによって入力されたテキストを第2の入力として使用する。サーバは、判定された第2の入力を自然言語理解モジュールに送信する。

【0154】

第2の入力は、ユーザとサーバ200との間の1回の対話における1つの発話であってもよいし、またはユーザとサーバ200との間の複数回の対話における複数の発話であってもよい。これは、本出願の本実施形態では限定されない。

【0155】

S109. サーバ200は、第2の入力に基づいて第1の意図の第1のロットに記入する。

【0156】

自然言語理解モジュール内のロット抽出モジュールは、第2の入力からエンティティを認識し、第1のロットに対応するエンティティを認識するために、計算のためにロット抽出モジュールに記憶されたロット抽出モデル内の、第1のロットに対応するアルゴリズムを呼び出す。言い換えれば、判定されたエンティティが第1のロットの値として使用される場合、それは、第1のロットに対応する情報が抽出されていることを示している。次に、ステップS116が実行される。

【0157】

S110. サーバ200は、第1のロットに記入しない。

【0158】

自然言語理解モジュール内のロット抽出モジュールは、第1のロットに対応する情報を第1のロットに記入しないことを決定する、すなわち、第1のロットの値を決定する必要がある。次に、ステップS116が実行される。

【0159】

S111. サーバ200は、第1のロットに対応する情報が必要かどうかを判定するためにユーザに質問を行う。

【0160】

本出願の一部の実施形態では、第1のロットが任意キースロットであると判定した（すなわち、ステップS106）後、ユーザが第1のロットに対応する情報を発話したかどうか

10

20

30

40

50

かにかかわらず、対話管理モジュールは、第1のロットに対応する情報に関してユーザに直接質問を行い得、ユーザは、第1のロットに対応する情報を第1のロットに記入する必要があるかどうかを判定する。第1のロットに対応する情報の質問方法および質問内容は、本出願の本実施形態では限定されない。

【0161】

本出願の一部の他の実施形態では、第1のロットが任意キースロットであると判定した（すなわち、ステップS106）後、対話管理モジュールは、ユーザが第1のロットに対応する情報を発話した可能性があるかどうかをさらに判定し得る。ユーザが第1のロットに対応する情報を発話した可能性がある場合、サーバ200は、第1のロットに対応する情報に関してユーザに質問を行い、ユーザは、第1のロットに対応する情報を第1のロットに記入する必要があるかどうかを判定する。このようにして、ユーザへの妨害を低減することに役立つように、ターゲットを絞った、ユーザからの確認が求められる。サーバ200が、ユーザが第1のロットに対応する情報を発話した可能性があるかどうかを判定するプロセスについては、以下の説明を参照されたい。ここでは詳細は説明されない。

【0162】

S112. サーバ200は、電子デバイスの第3の入力を受信する。

【0163】

第3の入力は、サーバ200の質問に対してユーザによって与えられた回答である。ユーザが音声形式で回答を与えた場合、サーバ200内の自動音声認識モジュールは、第3の入力を取得するために音声をテキストに変換し得る。ユーザがテキスト形式で回答を与えた場合、サーバ200は、ユーザによって入力されたテキストを第3の入力として使用する。サーバは、判定された第3の入力を自然言語理解モジュールに送信する。

【0164】

第3の入力は、ユーザとサーバ200との間の1回の対話における1つの発話であってもよいし、またはユーザとサーバ200との間の複数回の対話における複数の発話であってもよい。これは、本出願の本実施形態では限定されない。

【0165】

S113. サーバ200は、電子デバイス100の第3の入力に基づいて、第1のロットに対応する情報が必要かどうかを判定する。第1のロットに対応する情報が必要な場合、サーバ200は、ステップS114を実行する。第1のロットに対応する情報が不要な場合、サーバ200は、ステップS115を実行する。

【0166】

S114. サーバ200は、第3の入力に基づいて第1のロットに記入する。

【0167】

詳細については、ステップS109を参照されたい。次に、ステップS116が実行される。

【0168】

S115. サーバ200は、第1のロットに記入しない。

【0169】

次に、ステップS116が実行される。

【0170】

S116. サーバ200は、第1の意図と、第1の意図のロットに対応する抽出情報とに基づいて、第1の意図に対応する動作を実行する。

【0171】

図10Aおよび図10Bは、本出願の一実施形態による人間とコンピュータとの別の相互作用方法の概略フローチャートである。本方法は、ユーザが第1のロットに対応する情報を発話した可能性があるかどうかを判定するプロセスをさらに含む。すなわち、ステップS106の後かつステップS111の前に、ステップS201がさらに含まれ、具体的には以下の通りである。

【0172】

S201. サーバ200は、第1の入力が第1のロットに対応する情報を含む可能性がある

10

20

30

40

50

かどうかを判定する。サーバ200が、第1の入力が第1のロットに対応する情報を含む可能性があると判定した場合、サーバ200は、ステップS111を実行する。サーバ200が、第1の入力が第1のロットに対応する情報を含む可能性がないと判定した場合、サーバ200は、ステップS115を実行する。

【0173】

例えば、ユーザが第1のロットに対応する情報を発話するが、サーバ200が第1のロットに対応する情報の抽出に失敗する以下の2つの理由が存在し得る。

【0174】

理由1：ユーザ入力エラーまたは音声認識エラーにより、第1のロットに対応する情報の抽出は失敗する。例えば、サーバ200上に2つの意図が存在することが想定される。意図1は「設定を開いて」であり、対応する動作は、システム設定を開くことである。意図2は「WeChatの設定を開いて」であり、対応する動作は、WeChatアプリケーションの設定を開くことである。ユーザが「WeChatの設定を開いて」と発話すると、サーバ200は、ユーザのアクセントまたは句切りからこの音声を「WeChaの設定を開いて」と認識し得る。この場合は、サーバ200は、「WeCha」を抽出しない。サーバ200がユーザからの確認を求めない場合、サーバ200は、ユーザの意図を「設定を開いて」として直接認識し得る。すなわち、サーバ200は、システム設定を開く。これは、ユーザがWeChatアプリケーションの設定を開くことを期待した意図とは異なる。

【0175】

理由2：ユーザが第1のロットに対応する情報を正しく表現している場合、不正確なロット抽出モデルに起因して、第1のロットに対応する情報の抽出が失敗し得る。例えば、サーバ200は、スキル開発者がロット抽出モデルをトレーニングする前、比較的少ないユーザの発話しか入力されないため、または不正確なユーザの発話が入力されるため、トレーニングによって不正確なロット抽出モデルを生成する。

【0176】

2つの異なる理由から、本出願の本実施形態は、ユーザが第1のロットに対応する情報を発話した可能性があるかどうかを判定するために、以下の2つの方法を提供する。詳細は以下の通りである。

【0177】

方法1は、ユーザ入力エラーまたは音声認識エラーが発生したときに使用される。

【0178】

サーバ200は、サーバが、誤り訂正方法を使用して、ユーザの発話（すなわち、第1の入力）から認識されたエンティティがユーザ辞書内のキーワードに比較的類似していると判定した場合にのみ、ユーザ確認メカニズムをトリガし得る。エンティティとユーザ辞書内のキーワードとの類似度が判定されるとき、エンティティとユーザ辞書内のキーワードとの類似度を判定するために、例えば、ピンイン類似度に基づくアルゴリズムまたは文字列類似度に基づくアルゴリズムを使用して、第1の入力から認識されたエンティティとユーザ辞書内のキーワードとのレーベンシュタイン距離が計算され得る。あるいは、深層学習単語ベクトルまたは文ベクトルなどを使用して、単語または句間の類似度が計算されてもよい。本出願の本実施形態では、類似度計算方法は限定されない。

【0179】

以下では、レーベンシュタイン距離を計算する例を使用して、本出願の本実施形態で提供される第1の入力に含まれ得る、第1のロットに対応する情報を判定するための方法を説明する。レーベンシュタイン距離（Levenshtein Distance）は、1つの文字列を別の文字列に変換するために必要な編集操作の最小量である。編集操作は、1つの文字を別の文字に置き換えること、1つの文字を挿入すること、または1つの文字を削除することを含み得る。

【0180】

最初に、第1のロットがシステムロットタイプとユーザ辞書（すなわち、ユーザ定義ロットタイプ）とのどちらに対応するかが判定される。

【0181】

一部の実施形態では、サーバ200は、第1のスポットがユーザ定義スポットタイプに対応すると判定する。ユーザ定義スポットタイプの単語はユーザ定義であるため、ユーザ定義スポットタイプの単語の数は限られている。したがって、サーバ200は、第1のスポットに対応するユーザ定義スポットタイプのすべての単語をトラバースし、ユーザの発話内に判定されたエンティティとユーザ定義スポットタイプの単語とのレーベンシュタイン距離を計算し、これらのレーベンシュタイン距離から最小レーベンシュタイン距離を判定し得る。最小レーベンシュタイン距離に対応する、第1の入力内のエンティティは、第1のスポット内の潜在的なエンティティとして判定され得る、すなわち、第1のスポットに対応する情報であり得る。

10

【0182】

任意選択で、サーバ200は、最小値が閾値A未満であるかどうかを判定し得る。閾値Aは、開発者またはユーザによって指定された閾値であり得る。最小値が閾値A未満である場合、ユーザは第1のスポットに対応する情報を発話していないと考えられ得る。言い換えれば、第1の入力は、第1のスポットに対応する情報を含まない。この場合、サーバ200は、ユーザからの確認を求めなくてもよい。最小値が閾値A以上である場合、ユーザが第1のスポットに対応する情報を発話した可能性があると考えられ得る。この場合、サーバ200は、ユーザからの確認を求めてもよい。

【0183】

例えば、ユーザの意図は、「明日のShanghaiまでの航空券を予約して」であるが、「明日のshanghaまでの航空券を予約して」と間違っ

て入力される。第1の意図は、「航空券を予約して」であり、第1の意図は、時間スポット、出発地スポット、および目的地スポットを含む。ここでは、目的地スポットがユーザ辞書1に対応することが想定される。サーバ200は、目的地スポットに対応する情報の認識に失敗し、サーバ200は、第1の入力から複数のエンティティ、例えば「明日」および「shangha」を認識し得る。サーバ200は、「明日」とユーザ辞書1内の各単語との距離を計算し、「shangha」とユーザ辞書1内の各単語との距離を計算し、最短距離に対応する単語を選択する。例えば、「shangha」とユーザ辞書1内の「Shanghai」とのレーベンシュタイン距離が最も短い。この場合、「shangha」は第1のスポットの潜在的なエンティティであると判定され得る。さらに、「shangha」とユーザ辞書1内の最も類似する単語とのレーベンシュタイン距離が閾値Aよりも大きい場合、「shangha」は、ユーザによって発話された、第1のスポットに対応する情報である可能性があると考えられる。この場合、サーバ200は、ユーザからの確認を求めてもよい。

20

30

【0184】

任意選択で、開発者またはユーザは、閾値Bをさらに設定し得、閾値Bは閾値Aよりも大きい。最小値が閾値B以上である場合、それは、潜在的なエンティティがユーザ定義スポットタイプの単語に非常に類似していることを示し、基本的に、ユーザが第1のスポットに対応する情報を発話したと考えられ得る。すなわち、サーバは、ユーザからの確認を求めず、潜在的なエンティティを第1のスポットに対応する情報として直接判定し得る。最小値が閾値B以下であり、かつ閾値Aよりも大きい場合、ユーザが第1のスポットに対応する情報を発話した可能性があると考えられ得る。すなわち、潜在的なエンティティは、第1のスポットに対応する情報であり得る。この場合、サーバは、ユーザからの確認をさらに求めてもよい。これは、本出願の本実施形態では限定されない。

40

【0185】

一部の他の実施形態では、サーバ200は、第1のスポットがシステムスポットタイプに対応すると判定する。システムスポットタイプの無数の単語が存在する。したがって、第1の入力内のエンティティ単語とシステムスポットタイプのすべての単語とのレーベンシュタイン距離は、システムスポットタイプのすべての単語をトラバースすることによって計算され得ない。この場合、サーバ200は、ユーザが第1のスポットに対応する情報を発話したかどうかを判定し得ない。ユーザへの過度の妨害を回避するために、サーバ200は

50

、第1のロットに対応する情報に関してユーザからの確認を求めなくてもよい。

【0186】

方法2は、不正確なロット抽出モデルに起因してサーバ200が第1のロットに対応する情報の抽出に失敗した場合に使用される。

【0187】

ロット抽出モデルは、例えば名前付きエンティティ認識 (Named Entity Recognition, NER) 方法を使用して第1の入力からエンティティを認識し、各エンティティの信頼性レベルを計算するために、認識されたエンティティを、ロット抽出モデル内の第1のロットに対応するアルゴリズムに入力し得る。任意選択で、ロット抽出モデルは、エンティティを認識しなくてもよく、各単語の信頼性レベルを計算するために、第1の入力に含まれる各単語を、ロット抽出モデル内の第1のロットに対応するアルゴリズムに直接入力する。各エンティティまたは各単語の計算された信頼性レベルが特定の条件を満たした場合、ユーザが第1のロットに対応する情報を発話した可能性があると考えられ、次に、ユーザからの確認が求められる。第1の入力内の各エンティティまたは各単語の信頼性レベルの計算については、従来技術における分類ベースの計算方法およびシーケンスマーキングベースの計算方法などを参照されたい。ここでは詳細は説明されない。

【0188】

例えば、サーバ200は、各エンティティの信頼性レベルを計算するために、第1の入力内の各エンティティをロット抽出モデルに入力し得る。サーバ200は、エンティティの信頼性レベルが特定の条件を満たした場合にのみ、ユーザからの確認を求める。

【0189】

ロット抽出モデルが第1のロットに対応する情報の抽出に失敗することは、ユーザの発話から認識されたエンティティに関してサーバ200内のロット抽出モデルによって提供されるロットマーキング確率が認識閾値未満であることとして理解され得る。この場合、ユーザは、確認閾値を設定し得る。ユーザの発話から認識されたエンティティに関してロット抽出モデルによって提供されるロットマーキング確率が確認閾値よりも大きい場合、サーバ200は、ユーザ確認メカニズムをトリガする。すなわち、サーバ200は、ロット抽出モデルから取得された、第1の入力内の1つ以上のエンティティに対応する信頼性レベルが確認閾値よりも大きく、かつ認識閾値よりも小さい場合にのみ、第1のロットに対応する情報に関してユーザからの確認を求める。

【0190】

同じエンティティおよび同じロット抽出モデルに関して、ロット抽出モデルは、最初はエンティティを正しく認識しない可能性があるが、2回目はエンティティを正しく認識し得ることに留意されたい。これは、ユーザが最初にエンティティを発話したときにステートメントが別のエンティティを含み得るからである。すなわち、エンティティはコンテキストを有する。ロット抽出モデルが不正確である場合、エンティティの認識は、コンテキストの認識が失敗することから失敗する可能性がある。このため、サーバ200が最初にエンティティを認識し得なかった場合、サーバは、エンティティに関してユーザに質問を行う。この場合、ユーザの回答はエンティティに関するものである。ユーザの回答は、エンティティのみ、またはごく少数のコンテキストを含み得、ロット抽出モデルが、今度はエンティティを認識する可能性が高い。一部の他の実施形態では、エンティティはまた、ロット抽出モデルが使用されない方法でユーザの回答から認識され得る。例えば、エンティティを認識するための規則が可能であり得る。この規則は、ユーザの回答のコンテキスト論理、ユーザの意図との関連性、およびエンティティと第1のロットとの対応関係などの要因を参照してエンティティが認識され得ることである。これはまた、サーバ200が2回目またはそれ以降にユーザによって発話されたエンティティを認識する確率を効果的に高め得る。

【0191】

前述の機能を実施するために、前述の端末などは、機能を実行するための対応するハードウェア構造および／またはソフトウェアモジュールを含むことが理解され得る。当業者

は、本明細書に開示されている実施形態で説明された例と組み合わせて、ユニット、アルゴリズム、およびステップが、ハードウェアまたはハードウェアとコンピュータソフトウェアとの組み合わせによって実施され得ることに容易に気付くはずである。機能がハードウェアとコンピュータソフトウェアによって動かされるハードウェアとのどちらによって実行されるかは、技術的解決策の特定の用途および設計上の制約に依存する。当業者は、特定の用途ごとに、説明された機能を実施するために異なる方法を使用し得るが、その実施が本発明の実施形態の範囲を超えると考えられるべきではない。

【0192】

本出願の実施形態では、端末などは、前述の方法の例に基づいて機能モジュールに分割され得る。例えば、各機能モジュールは、対応する各機能に基づく分割によって得られ得るし、または2つ以上の機能が、1つの処理モジュールに統合され得る。統合モジュールは、ハードウェアの形態で実施され得るし、またはソフトウェア機能モジュールの形態で実施され得る。本発明の本実施形態では、モジュールへの分割は、例であり、論理的機能の分割にすぎないことに留意されたい。実際の実施態様では、別の分割方法が使用されてもよい。

10

【0193】

図11は、本出願の一実施形態によるサーバ200のハードウェア構造の概略図である。サーバ200は、少なくとも1つのプロセッサ201、少なくとも1つのメモリ202、および少なくとも1つの通信インターフェース203を含む。任意選択で、サーバ200は、図11に示されていない出力デバイスおよび入力デバイスをさらに含んでもよい。

20

【0194】

プロセッサ201、メモリ202、および通信インターフェース203は、バスを介して互いに接続される。プロセッサ201は、汎用中央処理ユニット（Central Processing Unit、CPU）、マイクロプロセッサ、特定用途向け集積回路（Application-Specific Integrated Circuit、ASIC）、または本出願の解決策のプログラム実行を制御するように構成された1つ以上の集積回路であり得る。プロセッサ201は、あるいは、複数のCPUを含んでもよく、プロセッサ201は、シングルコア（single-CPU）プロセッサまたはマルチコア（multi-CPU）プロセッサであってもよい。ここでのプロセッサは、データ（例えば、コンピュータプログラム命令）を処理するように構成された1つ以上のデバイス、回路、および／または処理コアであり得る。

30

【0195】

メモリ202は、読み出し専用メモリ（read-only memory、ROM）もしくは静的情報および命令を記憶し得る別のタイプの静的記憶デバイス、またはランダムアクセスメモリ（random access memory、RAM）もしくは情報および命令を記憶し得る別のタイプの動的記憶デバイスであってもよいし、または電氣的消去可能プログラマブル読み出し専用メモリ（Electrically Erasable Programmable Read-Only Memory、EEPROM）、コンパクトディスク読み出し専用メモリ（Compact Disc Read-Only Memory、CD-ROM）もしくは別のコンパクトディスク記憶媒体、光ディスク記憶媒体（コンパクトディスク、レーザディスク、光ディスク、デジタル多用途ディスク、もしくはブルーレイディスクなどを含む）、磁気ディスク記憶媒体、別の磁気記憶デバイス、または命令もしくはデータ構造の形式のしかるべきプログラムコードを保持もしくは記憶するために使用され得る、コンピュータによってアクセスされ得る任意の他の媒体であってもよい。これは限定されない。メモリ202は、独立して存在してもよく、バスを介してプロセッサ201に接続される。メモリ202は、あるいは、プロセッサ201と統合されてもよい。メモリ202は、本出願の解決策を実行するためのアプリケーションプログラムコードを記憶するように構成され、プロセッサ201は、実行を制御する。プロセッサ201は、本出願の実施形態における人間とコンピュータとの相互作用方法を実施するために、メモリ202に記憶されたコンピュータプログラムコードを実行するように構成される。

40

【0196】

通信インターフェース203は、別のデバイス、またはイーサネットもしくはワイヤレス

50

ローカルエリアネットワーク（wireless local area network、WLAN）などの通信ネットワークと通信するように構成され得る。

【0197】

プロセッサと通信するとき、出力デバイスは、複数の方法で情報を表示し得る。例えば、出力デバイスは、液晶ディスプレイ（Liquid Crystal Display、LCD）、発光ダイオード（Light-Emitting Diode、LED）ディスプレイデバイス、陰極線管（Cathode Ray Tube、CRT）ディスプレイデバイス、またはプロジェクタ（projector）などであり得る。プロセッサと通信するとき、入力デバイスは、複数の方法でユーザの入力を受信し得る。例えば、入力デバイスは、マウス、キーボード、タッチスクリーンデバイス、または感知デバイスであり得る。

10

【0198】

実施態様に関する前述の説明は、当業者が、簡便な説明のために前述の機能モジュールへの分割が例示のための例として用いられていることを理解することを可能にする。実際の用途では、前述の機能は、要求に応じて異なるモジュールに割り当てられて実施され得る、すなわち、装置の内部構造は、上記の機能の全部または一部を実施するために異なる機能モジュールに分割される。前述のシステム、装置、およびユニットの詳細な動作プロセスについては、前述の方法の実施形態における対応するプロセスを参照されたく、ここでは詳細は再び説明されない。

【0199】

本出願の実施形態における機能ユニットは、1つの処理ユニットに統合されてもよいし、または各ユニットは物理的に単独で存在してもよいし、または2つ以上のユニットが、1つのユニットに統合される。統合ユニットは、ハードウェアの形態で実施されてもよいし、またはソフトウェア機能ユニットの形態で実施されてもよい。

20

【0200】

統合ユニットが、ソフトウェア機能ユニットの形態で実施され、独立した製品として販売または使用される場合、統合ユニットは、コンピュータ可読記憶媒体に記憶されてもよい。このような理解に基づいて、本質的に本出願の実施形態の技術的解決策は、または従来技術に寄与する部分、または技術的解決策の全部もしくは一部は、ソフトウェア製品の形態で実施され得る。コンピュータソフトウェア製品は、記憶媒体に記憶され、本出願の実施形態で説明された方法のステップの全部または一部を実行するようにコンピュータデバイス（パーソナルコンピュータ、サーバ、またはネットワークデバイスであり得る）に命令するためのいくつかの命令を含む。前述の記憶媒体は、フラッシュメモリ、リムーバブルハードディスク、読み出し専用メモリ、ランダムアクセスメモリ、磁気ディスク、または光ディスクなどの、プログラムコードを記憶し得る任意の媒体を含む。

30

【0201】

前述の説明は、本出願の特定の実施態様にすぎず、本出願の保護範囲を限定することを意図されていない。本出願に開示されている技術的範囲内のいかなる変形または置換も、本出願の保護範囲内にあるものとする。したがって、本出願の保護範囲は、特許請求の範囲の保護範囲に従うものとする。

【符号の説明】

40

【0202】

- 100 電子デバイス
- 110 プロセッサ
- 120 外部メモリインターフェース
- 121 内部メモリ
- 130 ユニバーサルシリアルバスポート、USBインターフェース
- 140 充電管理モジュール
- 141 電力管理モジュール
- 142 バッテリ
- 150 移動通信モジュール

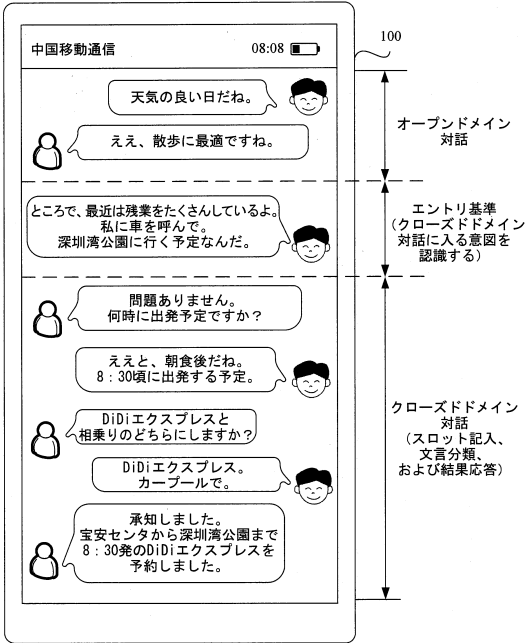
50

160	ワイヤレス通信モジュール	
170	音声モジュール	
170A	スピーカ	
170B	受話器	
170C	マイクロフォン	
170D	ヘッドセットジャック	
180	センサモジュール	
180A	圧力センサ	
180B	ジャイロ스코ープセンサ	
180C	気圧センサ	10
180D	磁気センサ	
180E	加速度センサ	
180F	距離センサ	
180G	光学式近接センサ	
180H	指紋センサ	
180J	温度センサ	
180K	タッチセンサ	
180L	周囲光センサ	
180M	骨伝導センサ	
190	ボタン	20
191	モータ	
192	インジケータ	
193	カメラ	
194	ディスプレイ	
195	加入者識別モジュールカードインターフェース	
200	サーバ	
201	プロセッサ	
202	メモリ	
203	通信インターフェース	
300	サーバ	30
400	電子デバイス、ホームページ	
401	コントロール	
402	コントロール	
500	ページ	
600	意図作成ページ	
601	ユーザ発話設定ページ	
602	ユーザ発話追加コントロール	
603	情報項目	
604	対話ボックス	
605	スロットタイプ追加コントロール	40
606	スロットリスト確認コントロール	
607	新スロット追加コントロール	
608	ページ	
700	スロットタイプ編集ページ	
701	テキストボックス	
702	値の項目	
703	同義語の項目	
704	スロットタイプ編集ページの領域	
705	バッチ追加ボタン	
800	ページ	50

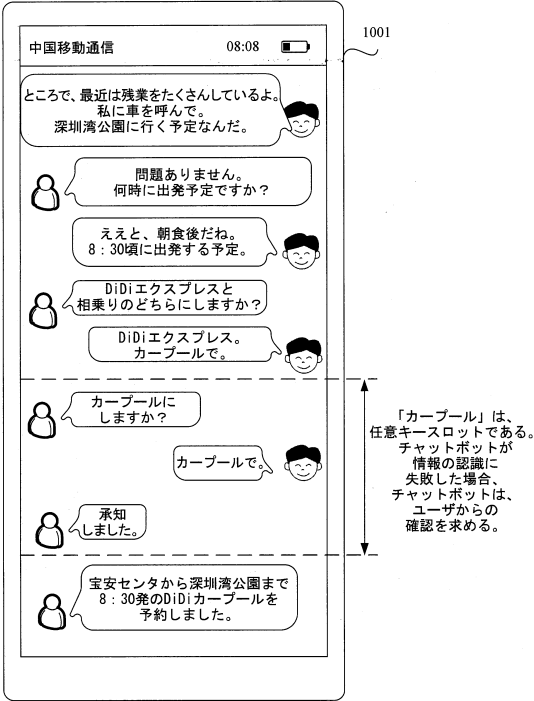
- 801 コントロール
- 900 ページ
- 901 コントロール
- 902 コントロール

【図面】

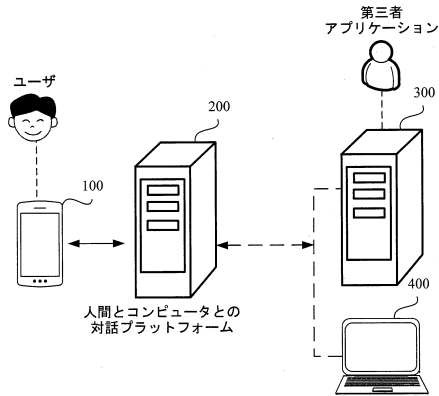
【図 1 A】



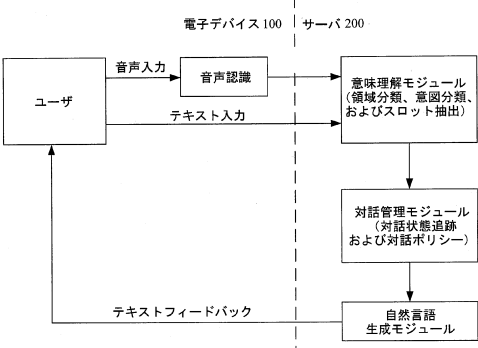
【図 1 B】



【図 2】



【図 3】



10

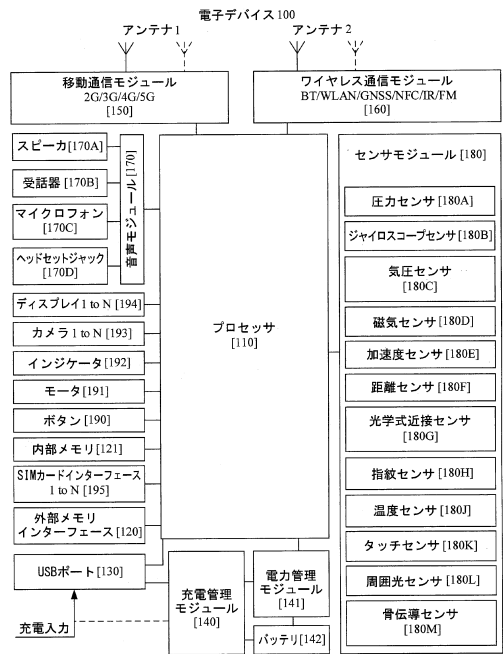
20

30

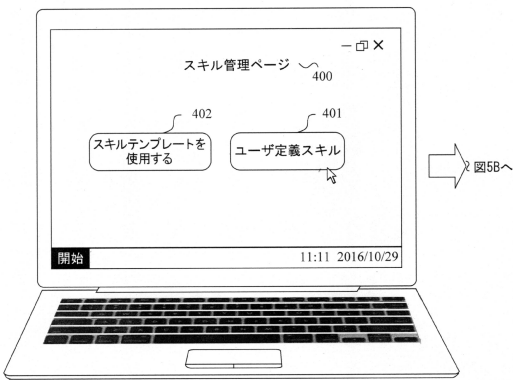
40

50

【図 4】



【図 5 A】



10

20

【図 5 B】

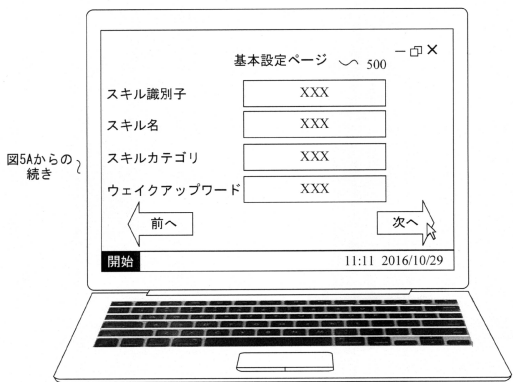


図5Cへ

【図 5 C】

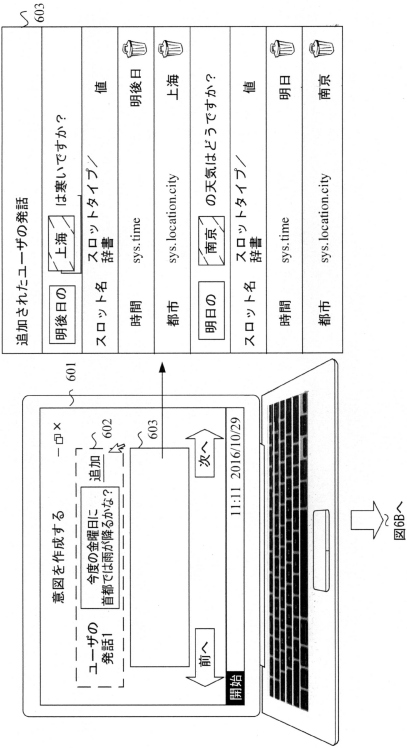


30

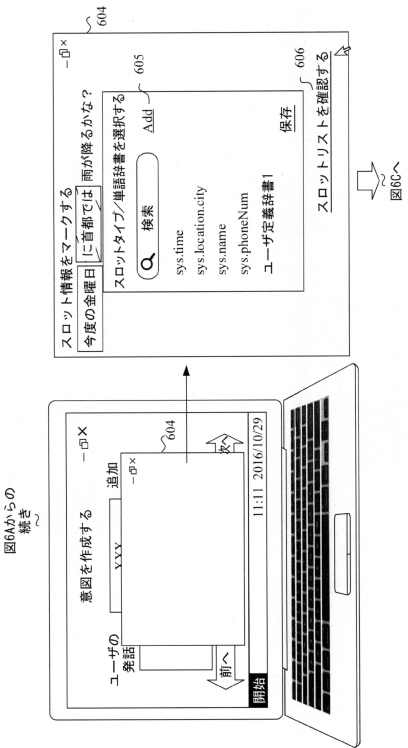
40

50

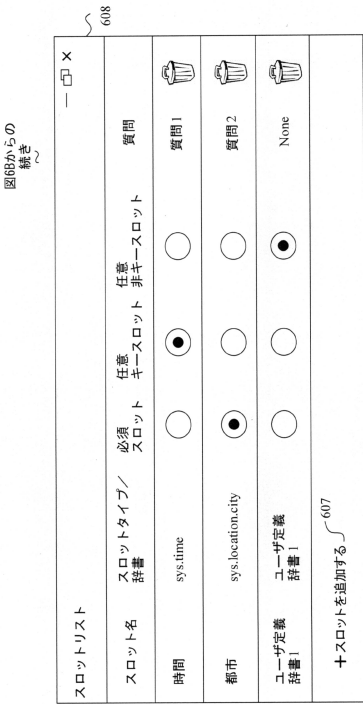
【図 6 A】



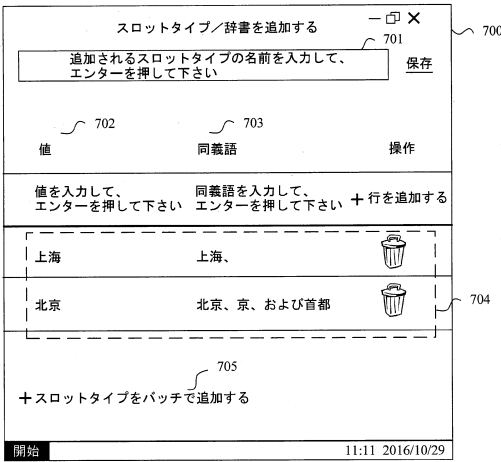
【図 6 B】



【図 6 C】



【図 7】



10

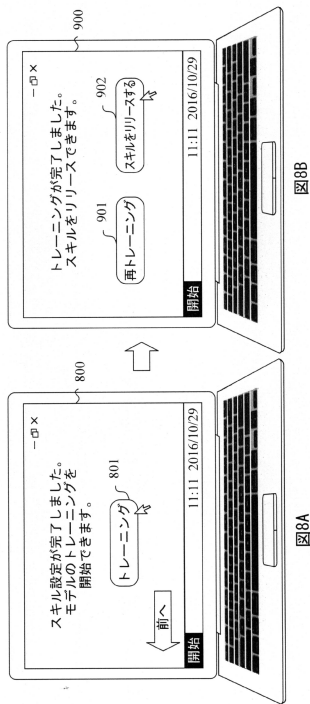
20

30

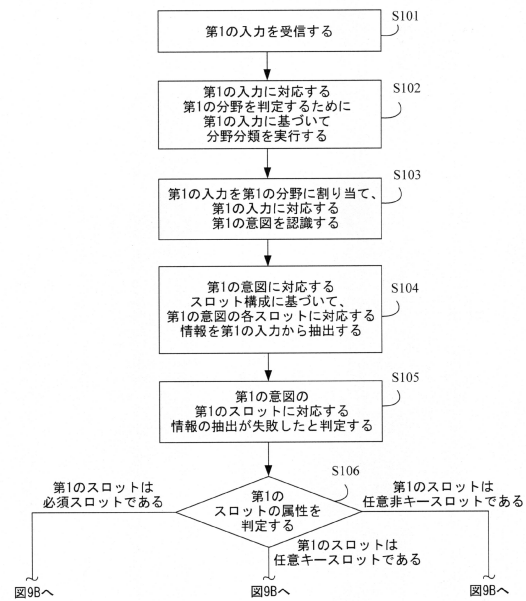
40

50

【図 8】



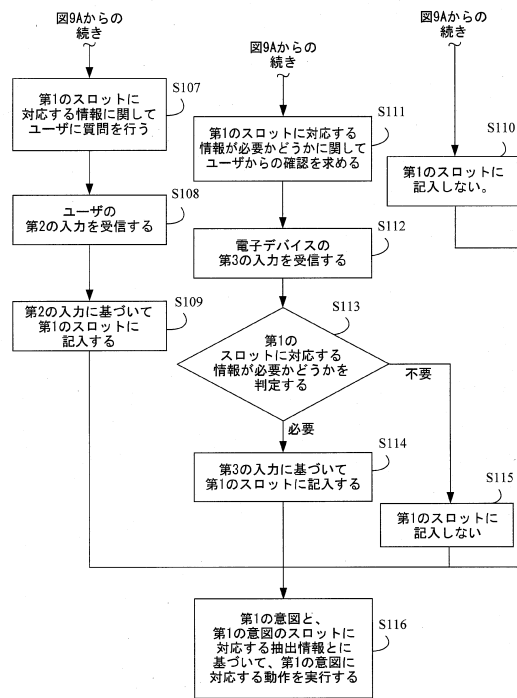
【図 9 A】



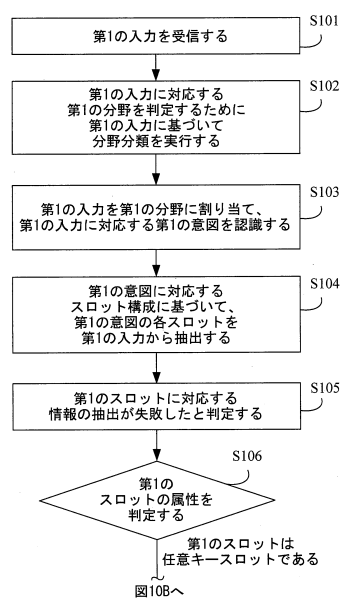
10

20

【図 9 B】



【図 10 A】

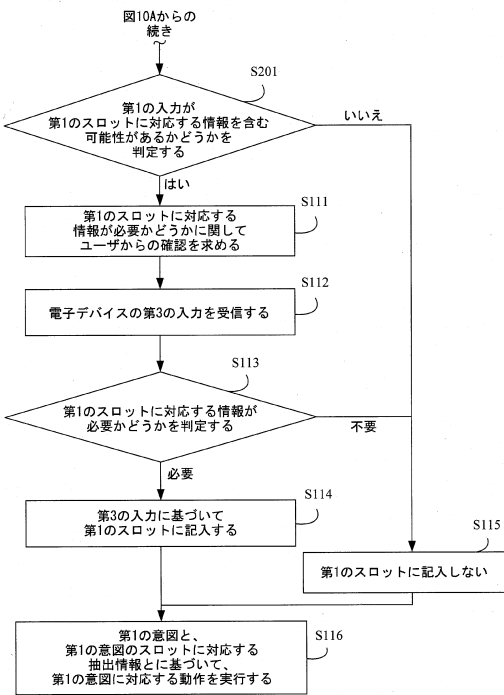


30

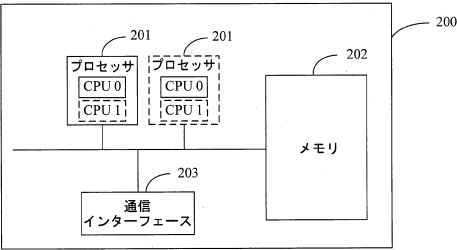
40

50

【図 10B】



【図 11】



10

20

30

40

50

フロントページの続き

- (74)代理人 100133569
弁理士 野村 進
- (72)発明者 ▲張▼ 晴
中華人民共和国 5 1 8 1 2 9 広東省深▲チェン▼市龍岗区坂田 華為総部▲ベン▼公楼
- (72)発明者 ▲張▼ ▲錦▼▲輝▼
中華人民共和国 5 1 8 1 2 9 広東省深▲チェン▼市龍岗区坂田 華為総部▲ベン▼公楼
- (72)発明者 ▲張▼ ▲軼▼博
中華人民共和国 5 1 8 1 2 9 広東省深▲チェン▼市龍岗区坂田 華為総部▲ベン▼公楼
- 審査官 松尾 真人
- (56)参考文献 特許第6370962 (JP, B1)
特表2008-506156 (JP, A)
米国特許出願公開第2017/0039181 (US, A1)
特開2016-212405 (JP, A)
- (58)調査した分野 (Int.Cl., DB名)
G 0 6 F 1 6 / 0 0 - 1 6 / 9 5 8