

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6952690号

(P6952690)

(45) 発行日 令和3年10月20日 (2021. 10. 20)

(24) 登録日 令和3年9月30日 (2021. 9. 30)

(51) Int. Cl.	F I
<b>G 0 6 F 3/16 (2006.01)</b>	G 0 6 F 3/16 6 5 0
<b>G 1 0 L 15/22 (2006.01)</b>	G 1 0 L 15/22 3 0 0 Z
	G 0 6 F 3/16 6 1 0

請求項の数 17 (全 20 頁)

(21) 出願番号	特願2018-525352 (P2018-525352)	(73) 特許権者	314015767
(86) (22) 出願日	平成28年11月22日 (2016. 11. 22)		マイクロソフト テクノロジー ライセン
(65) 公表番号	特表2019-502981 (P2019-502981A)		シング, エルエルシー
(43) 公表日	平成31年1月31日 (2019. 1. 31)		アメリカ合衆国 ワシントン州 9805
(86) 国際出願番号	PCT/US2016/063263		2 レッドモンド ワン マイクロソフト
(87) 国際公開番号	W02017/091550		ウェイ
(87) 国際公開日	平成29年6月1日 (2017. 6. 1)	(74) 代理人	100079108
審査請求日	令和1年10月25日 (2019. 10. 25)		弁理士 稲葉 良幸
(31) 優先権主張番号	62/259, 849	(74) 代理人	100109346
(32) 優先日	平成27年11月25日 (2015. 11. 25)		弁理士 大貫 敏史
(33) 優先権主張国・地域又は機関	米国 (US)	(74) 代理人	100117189
			弁理士 江口 昭彦
		(74) 代理人	100134120
			弁理士 内藤 和彦

最終頁に続く

(54) 【発明の名称】 自動的な音声対話スクリプトの発見

(57) 【特許請求の範囲】

【請求項 1】

自動化された対話システムを構成する方法であって、

グラフィカルユーザーインターフェイス (GUI) の操作中に、前記 GUI の実行時状態データを観察するステップであって、前記 GUI は、データネットワーク上において前記対話システムへアクセス可能であるウェブに基づいたインターフェイスを有する、ステップと、

前記実行時状態データの複数のトレースを記録するステップであって、前記トレースは、前記実行時状態データにおいて、前記 GUI に提示される入力と前記 GUI により生成された出力の値とを識別する、ステップと、

前記ウェブに基づいたインターフェイスにより提示されるウェブページの対話型要素を抽出し、各テンプレートを前記対話型要素に割り当てることにより、複数のテンプレートのうちの各テンプレートを前記トレースと関連付けるステップであって、各テンプレートは、1つ又は複数の定義された入力と1つの定義された出力とを有する、ステップと、

前記複数のテンプレート及び関連付けられたトレースのそれぞれについて、前記自動化された対話システムを構成するステップであって、

前記 GUI に提供するための1つ又は複数の入力データの値を決定するために、ユーザーによる言語的入力発話を前記テンプレートに従って解釈し、

前記 GUI に前記1つ又は複数の入力データの値を前記言語的入力発話の解釈に従って提供し、

10

20

前記 G U I から出力の値を取得し、  
前記出力の値を示す言語的出力発話を出力する、  
ように前記自動化された対話システムを構成するステップと、  
を含む方法。

【請求項 2】

前記言語的入力発話は、発話に基づいた入力を有する、請求項 1 に記載の方法。

【請求項 3】

前記言語的入力発話は、テキストに基づいた入力を有する、請求項 1 に記載の方法。

【請求項 4】

複数の前記トレースを記録するステップは、ユーザーと前記 G U I との間におけるやり取りを監視するステップを含む、請求項 1 に記載の方法。 10

【請求項 5】

前記 G U I は、G U I クライアントのプログラムである、請求項 4 に記載の方法。

【請求項 6】

自動化された対話システムのシミュレーションを記録することにより、追加の複数のトレースを記録するステップ、を更に含む、請求項 1 に記載の方法。

【請求項 7】

前記自動化された対話システムの前記シミュレーションは、テキストに基づいたユーザー入力を使用する、請求項 6 に記載の方法。

【請求項 8】 20

前記自動化された対話システムの前記シミュレーションは、発話に基づいたユーザー入力を使用する、請求項 6 に記載の方法。

【請求項 9】

前記 G U I への入力を自動的に生成するステップと、前記入力、出力を正常に導出するトレースを形成している際に、そのトレースを記録するステップと、を更に含む、請求項 1 に記載の方法。

【請求項 10】

アプリケーションの複数のウェブに基づいたインターフェイスを自動的に見出すステップと、前記インターフェイスごとに、前記ウェブに基づいたインターフェイスを介して前記テンプレートへの入力の値を提供すると共に前記インターフェイスを介して前記テンプレートの出力を読み出すための対応する手順によって 1 つ又は複数のテンプレートを構成するステップと、を更に含む、請求項 1 に記載の方法。 30

【請求項 11】

ロジックデバイスと、

前記ロジックデバイスにより方法を実行するための実行可能な命令を保持している記録デバイスと、を備え、

前記方法は、

グラフィカルユーザーインターフェイス ( G U I ) の操作中に、前記 G U I の実行時状態データを観察するステップであって、前記 G U I は、データネットワーク上において対話システムへアクセス可能であるウェブに基づいたインターフェイスを有する、ステップと、 40

前記実行時状態データにおいて前記 G U I に提示される入力と前記 G U I により生成された出力の値とを識別する、前記実行時状態データの複数のトレースを記録するステップと、

前記ウェブに基づいたインターフェイスにより提示されるウェブページの対話型要素を抽出し、各テンプレートを前記対話型要素に割り当てることにより、複数のテンプレートのうちの各テンプレートを前記トレースと関連付けるステップであって、各テンプレートは、1 つ又は複数の定義された入力と 1 つの定義された出力とを有する、ステップと、

前記複数のテンプレート及び関連付けられたトレースのそれぞれについて、自動化された対話システムを構成するステップであって、 50

前記 G U I に提供するための 1 つ又は複数の入力データの値を決定するために、ユーザーによる言語的入力発話を前記テンプレートに従って解釈し、

前記 G U I に前記 1 つ又は複数の入力データの値を前記言語的入力発話の解釈に従って提供し、

前記 G U I から出力の値を取得し、

前記出力の値を示す言語的出力発話を出力する、

ように前記自動化された対話システムを構成するステップと、  
を含む、コンピュータシステム。

【請求項 1 2】

前記言語的入力発話は、発話に基づいた入力を有する、請求項 1 1 に記載のコンピュータシステム。

【請求項 1 3】

前記言語的入力発話は、テキストに基づいた入力を有する、請求項 1 1 に記載のコンピュータシステム。

【請求項 1 4】

複数のテンプレートのうちの各テンプレートを前記トレースと関連付けるステップは、前記複数のトレースのうちのトレースの中で、1 つ又は複数の規則的な構造を検出することと、

前記トレースの中で検出された規則的な構造のそれぞれについて、前記規則的な構造に含まれる 1 つ又は複数のデータタイプを識別し、前記規則的な構造に含まれるデータタイプのそれぞれについて、前記データタイプに対応する、利用可能な N L P テンプレートのフィールドを選択することと、

を含む、請求項 1 に記載の方法。

【請求項 1 5】

前記利用可能な N L P のフィールドを選択することは、予めトレーニングされた離散時間マルコフモデルを操作すること、を含む、請求項 1 4 に記載の方法。

【請求項 1 6】

前記 G U I は、1 つ又は複数のハイパーテキストマークアップ言語 ( H T M L ) の要素により定義され、

前記 1 つ又は複数の規則的な構造は、H T M L 入力の要素を含む、

請求項 1 4 に記載の方法。

【請求項 1 7】

自動化された対話システムと、

グラフィカルユーザーインターフェイス ( G U I ) の操作中に、前記 G U I の実行時状態データを観察し、前記実行時状態データにおいて前記 G U I に提示される入力と前記 G U I により生成された出力の値とを識別する、前記実行時状態データの複数のトレースを記録する、監視システムであって、前記 G U I は、データネットワーク上において前記対話システムへアクセス可能であるウェブに基づいたインターフェイスを有する、監視システムと、

テンプレートコンフィギュレータであって、

前記ウェブに基づいたインターフェイスにより提示されるウェブページの対話型要素を抽出し、各テンプレートを前記対話型要素に割り当てることにより、複数のテンプレートのうちの各テンプレートを前記トレースと関連付けるステップであって、各テンプレートは、1 つ又は複数の定義された入力と 1 つの定義された出力とを有する、ステップと、

前記複数のテンプレート及び関連付けられたトレースのそれぞれについて、前記自動化された対話システムを構成するステップであって、

前記 G U I に提供するための 1 つ又は複数の入力データの値を決定するために、ユーザーによる言語的入力発話を前記テンプレートに従って解釈し、

前記 G U I に前記 1 つ又は複数の入力データの値を前記言語的入力発話の解釈に従って提供し、

10

20

30

40

50

前記 G U I から出力の値を取得し、  
前記出力の値を示す言語的出力発話を出力する、  
ように前記自動化された対話システムを構成するステップと、  
を行うように構成されているテンプレートコンフィギュレータと、  
を備える、コンピュータシステム。

【発明の詳細な説明】

【技術分野】

【 0 0 0 1 】

関連出願の相互参照

本出願は、2015年11月25日付けで出願された米国仮特許出願第62/259, 849号明細書の利益を主張するものであり、この特許文献の内容は、引用により、本明細書に包含される。

10

【 0 0 0 2 】

本発明は、自動的な音声対話スクリプトの発見に関し、且つ、更に詳しくは、アプリケーションに対するスクリプトに基づいたアクセスのための音声対話システムの自動的な構成に関する。

【背景技術】

【 0 0 0 3 】

ワールドワイドウェブ(「ウェブ」)は、人間のユーザーによってアクセスされるように意図された様々なタイプのリソースを含み、これには、コンテンツを直接的に内蔵する文書のみならず、インターフェイスを通じて提供される情報に応答してコンテンツを提供するコンピュータ実装システムに対するインターフェイスも含まれる。

20

【 0 0 0 4 】

又、コンピュータ実装システムに対するコンピュータ指向のインターフェイスには、例えば、XML及びJSONなどのフォーマットを使用したデータ交換の方法を提供しうる「ウェブサービス」インターフェイスを使用することにより、ウェブ上においてアクセス可能である。ウェブサービスを使用するために、「クライアント」システムの作成者は、一般に、そのウェブサービスを介して情報にアクセスするプログラムを作成するべく、そのウェブサービス用のドキュメンテーションを使用している。ウェブサービスに対する要求及び応答の構造を(例えば、XMLスキーマとして)規定することにより、クライアントの作成のある程度の自動化を許容することができる。

30

【 0 0 0 5 】

実際に、ウェブ上においてアクセス可能である大部分のコンピュータ実装システムは、人間指向のインターフェイスを提供しており、且つ、コンピュータ指向(例えば、ウェブサービス)のインターフェイスを提供するものは、非常に少ない。従って、それぞれのインターフェイスごとに大規模なプログラミングを必要とせずに人間指向のインターフェイスを自動的に使用することができるようにするというニーズが存在している。

【 0 0 0 6 】

益々、ユーザーは、従来のGUIに基づいたインターフェイスを使用せずにコンピュータ実装システムに対してアクセスすることを所望するようになっている。例えば、現在の音声に基づいたパーソナルアシスタント(例えば、AppleのSiri)は、GUIを使用せずに、音声に基づいた対話を使用して情報を提供するように試みている。

40

【 0 0 0 7 】

1990年代の初頭から、アクセス可能なコンテンツを「クロール」する自動化された「ウェブロボット」を使用することにより、ウェブのコンテンツの自動的なインデックス付けが行われている。一般に、このようなウェブクロラーは、1つのウェブサイトにおいて開始し、且つ、次いで、それぞれのサイト上のその他のサイトへのハイパーリンクを辿り、且つ、このサーチを適切な方式によって継続することにより、そのインデックス付けサーチを拡大している。これらのウェブサーチブラウザに対する大きな改善は、Googleによって導入されたものであり、Googleは、それぞれのページをポイン

50

トしているその他のページの数の関数により、それぞれのページをランク付けしたのである。

#### 【 0 0 0 8 】

サーチエンジンによるウェブの「クローリング」は、例えば、キーワードに基づいたクエリに応答して、ウェブページ上において内蔵されているコンテンツに自動的にアクセスする方法を提供している。但し、このような方式は、一般に、ウェブ上において提示されているインターフェイスを介してアクセス可能であるコンテンツに、ではなく、ウェブページ上の明示的なコンテンツに対して着目している。

#### 【 0 0 0 9 】

その後の努力においては、HTML又はDOM ( Document Object Module ) 情報を解析することにより、ウェブページから情報を抽出することが試みられている。DOM情報は、相対的に静的であることから、これらの技法は、ニュース、天気、及びその他の情報ページのリサンプリングを許容している。但し、これらの進歩した技法の場合にも、情報がサイトに供給されると共に、データ、マップ、写真、又はオーディオが返されるウェブサイトをユーザーが活用することは、許容されていない。ウェブサイトの自動的な使用を許容する制限が注釈付けされると共にカタログされている「セマンティックウェブ」の生成を狙った早期の試みは、その大部分が失敗している。例えば、W3Cは、リンクされたデータのウェブの1つのビジョンとしての「セマンティックウェブ」を参照している。セマンティックウェブ技術によれば、人々は、ウェブに関するデータストアを生成することが可能であり、ボキャブラリを構築することが可能であり、且つ、データの取扱いのための規則を記述することができる。但し、この技術は、広く採用されてはいない。

#### 【 0 0 1 0 】

ウェブに基づいたインターフェイスを有するロボット的なやり取り（例えば、これらは、人間指向のもの又はウェブサービスに基づいたものである）は、人間によるプログラミングに基づいてスクリプト化することができる（しばしば、「スクリーンスクレーピング」と呼称されている）。例えば、プログラマは、コンピュータ実装システム内において情報を取得するべく、ユーザーの動作を真似ている。

#### 【 0 0 1 1 】

汎用の対話システムの構築における課題の1つが、新しいサービス又は新しいやり取りをカバーする更なる機能の追加である。例えば、人々は、ユーザーによる新しい航空会社におけるチケットの予約、近所のレストランにおける料理の注文、映画チケットの購入、又は新しいソーシャルネットワーキングサービスの使用を可能にすることを所望しうであろう。従来においては、プログラマが、それぞれの更なるサービスとやり取りするために必要とされるAPIを生成又は使用することになるが、これには、多くの工数が所要されることになる。更には、人々が、（例えば、連絡先情報又はフライト旅程について同一の表現を使用することによって）これらの新しいコンポーネントを対話システムの残りの部分に統合することを所望した場合には、コンポーネントのすべてが1つにフィットすることを保証するべく、更なる注意を払わなければならない。

#### 【 0 0 1 2 】

但し、多くの方法においては、これは、重複した作業であり、HTML搭載型のディスプレイインターネットは、周知のように、広範なオンラインサービス及び情報とやり取りするべく必要とされるコンポーネントの大部分を含んでいる。実際に、大部分でない場合にも、多くのオンラインサービスは、明示的にディスプレイインターネットに留意しつつ構築されている。但し、表示重視型のサービスへ着目することは、インターネット上において利用可能である情報の大部分が、機械にとっては、或いは、オーディオのみのインターフェイスにとっては、アクセス可能性が相対的に低く、或いは、アクセス不能であることを意味している。これは、ウェブを（人々及び機械の両方にとって）相対的にアクセス可能な状態とするための努力が払われていないことを意味するものではない。ARIAのような規格によれば、例えば、スクリーンリーダーを使用することにより、視覚障害を含

む 障害を有するユーザーは、ウェブサイトを相対的に容易にナビゲートすることができる。同様に、セマンティックウェブ規格も、ウェブサイトの機械による解釈を相対的に容易にすることを狙った試みである。

【 0 0 1 3 】

これらの規格は、均等に、或いは、（特に、セマンティックウェブのケースにおいては）広範に、利用されてはいない。アクセス可能性の規格が利用されている際にも、これらの規格は、既定の視覚的表示インターフェイスを有するものと同程度に使用が容易である経験を促進してはいない。換言すれば、スクリーンリーダーは、その名称が意味しているだけのものに過ぎず、それらは、スクリーンを読み上げて、情報の統合をユーザーに任せている。それは、どんな理由にしろ（例えば、視覚的な障害を有しているために、運転中であるために、或いは、単にコンピュータのそばにいないために）、対象のウェブサイトを見ることができない別の人物とやり取りしなければならない場合に人物が得ることができるコヒーレントな対話型の経験ではない。

10

【 0 0 1 4 】

コンピュータ実装システムに対する音声に基づいたインターフェイスは、一般に、人間 - コンピュータ対話コンポーネントとコンピュータ実装システムに対するインターフェイスとの間における「コネクタ」のプログラミングを必要としている。例えば、実験的旅行予約システムは、音声に基づいた対話を実装しており、且つ、旅行予約システムとの間におけるやり取りをプログラミングしている（例えば、S a b r e）。

【 発明の概要 】

20

【 0 0 1 5 】

一態様である、コンテンツに対するインターフェイスの自動化された発見は、一般に、インターフェイスを介してアクセス可能なコンテンツの自動化された取得を許容している。

【 0 0 1 6 】

いくつかの例においては、この発見は、インターフェイス上におけるコンテンツ（例えば、人間のユーザーへの提示が意図されたラベル及びプロンプト）の分析のみならず、入力のどの組合せが、コンピュータ実装システムからコンテンツを取得する十分に形成された入力を提供するのかを判定するためのインターフェイスの自動化された「プロービング」を伴うことができる。この発見されたインターフェイス規則は、人物による、或いは、人物の代理人による、将来における使用のために、データ構造内においてキャッシングされることになる。

30

【 0 0 1 7 】

別の態様は、コンピュータアクセス可能なインターフェイスの組（例えば、インターネットアクセス可能なウェブ上のURL）の「クローリング」と、ウェブサーチ可能なコンテンツのうちのコンテンツへの自動的なアクセス用のシステムを提供するためのインターフェイスの自動化された発見の使用と、を含む。このクローリング及び分析は、人間 / ウェブのやり取りを観察するトレーニングプログラムを通じて学習されてもよく、或いは、これは、サイトによって要求されたそれぞれの情報要素の制約と、更には、それぞれの要素が必要とされているのか又は任意選択であるのかと、を発見するべくウェブサイトを探る自動化されたシステムによって発見されてもよい。

40

【 0 0 1 8 】

別の態様は、例えば、その特性が自動的に発見されているインターフェイスなどの、インターフェイスを介してコンテンツにアクセスするべく必要とされる情報を導出するべく人間 - コンピュータのやり取りをシミュレートするステップを含む。次いで、この情報は、自動化されたシステム内のインターフェイスを介して提供され、且つ、インターフェイスを介して取得されたコンテンツは、ユーザーに提供される。例えば、情報の導出及び取得された情報の提示は、（例えば、音声及び / 又はテキストによって）単一の人間 - コンピュータの対話において制御されている。

【 0 0 1 9 】

50

別の態様は、テキストに基づいた情報交換を使用することによってウェブサイトとやり取りするステップと、いずれのタイプの情報がウェブサイトによって必要とされており、且つ、いずれが任意選択であるのか、についてそれらのクエリから学習するステップと、を含む。ウェブサイトのテキストに基づいた使用のための要件の記述は、将来における使用のために保存されている。

【0020】

別の態様は、人間のユーザーが情報を求め、且つ、対話に基づいて、複数のコンピュータ実装システムのうちの1つが、ユーザーのニーズを充足するべく、コンテンツの取得のために選択される、人間-コンピュータの対話を実施するステップ（或いは、実施するべく機械可読形態において命令を規定するステップ）を含む。

10

【0021】

人間指向のコンピュータ実装されたインターフェイスは、一般に、インターフェイスを介して利用可能であるコンテンツに対するコンピュータ実装されたアクセスを許容する形態において、関連する「命令」を有してはいない。一般に明示的ではない態様は、限定を伴うことなしに、

インターフェイスを介してアクセス可能であるシステムの一般的な主題（例えば、天気予想やエアラインスケジュールなど）

インターフェイス内の要素と関連するセマンティック情報（例えば、フォームフィールドをユーザー名と関連付けることなど）

有効なクエリを定義するためのインターフェイス内の要素の組に対する全体的な制約（例えば、天気予想に必要とされる都市及び州）を含む。

20

【0022】

1つの技術的な問題は、インターフェイスを介したコンテンツに対する自動化された（例えば、ロボット的な）アクセスに必要とされる又は有用である態様を判定するためのコンピュータ実装されたインターフェイスの自動化された分析である。

【0023】

解決されるべき別の技術的問題は、（人間指向のインターフェイスから自動的に発見された）インターフェイスを特徴付ける情報の使用（或いは、さもなければ、データスキーマの使用など）と、そのインターフェイス用の特定のプログラミングを必要とせずにインターフェイスを介して入力用の情報を導出するための人間-コンピュータの対話の制御と、である。

30

【0024】

別の態様においては、一般に、自動化された対話システムを構成する方法は、アプリケーションのグラフィカルユーザーインターフェイス（GUI：Graphical User Interface）サーバーを介してやり取りの複数のトレースを収集することにより、開始している。それぞれのトレースは、GUIの複数の提示のコンテキストにおけるやり取りを含んでいる。GUIの1つ又は複数の提示の要素が識別されている。テンプレートは、トレースの一部分と関連付けられている。それぞれのテンプレートは、1つ又は複数の定義された入力と、1つの定義された出力と、を有する。複数のテンプレートのそれぞれのテンプレートごとに、トレースの一部分は、GUIを介してテンプレートへの入力の値を提供すると共に出力の値を取得するための手順を規定することにより、テンプレートを構成するべく、処理されている。自動化された対話システムは、構成されたテンプレートによって構成されている。

40

【0025】

以下の特徴のうちの1つ又は複数が含まれていてもよい。

【0026】

方法は、構成された自動化された対話システムを使用することにより、ユーザー入力を処理するステップを更に含む。この処理は、ユーザーから言語的入力を受け取るステップと、複数のテンプレートに従って言語的入力を解釈するステップと、構成されたテンプレ

50

ートに従ってアプリケーションとやり取りするステップと、を含む。言語的入力、発話に基づいた入力及び／又はテキストに基づいた入力を含むことができる。

【0027】

グラフィカルユーザーインターフェイスは、データネットワーク上においてシステムからアクセス可能であるウェブに基づいたインターフェイスを有し、且つ、GUIの提示は、ウェブに基づいたインターフェイスのウェブページを有する。

【0028】

テンプレートをトレースの一部分と関連付けるステップは、ウェブページの対話型要素を抽出するステップと、1つ又は複数のテンプレートを対話型要素のそれぞれに対して割り当てるステップと、を含む。

10

【0029】

トレースを収集するステップは、ユーザーとアプリケーションとの間のやり取りを監視するステップを含む。例えば、ユーザーとアプリケーションとの間のやり取りは、GUIクライアントを介したものであり、或いは、自動化された対話システムのシミュレーションを伴うものである。自動化された対話システムのシミュレーションは、テキストに基づいたユーザー入力及び／又は発話に基づいたユーザー入力を使用することができる。

【0030】

トレースを収集するステップは、アプリケーションに対する入力を自動的に生成するステップと、入力が、出力を正常に導出するトレースを形成している際に、そのトレースを収集するステップと、を更に含む。

20

【0031】

トレースを収集するステップは、アプリケーションに対する入力を自動的に生成するステップと、入力が、出力を正常に導出するトレースを形成している際に、そのトレースを収集するステップと、を含む。

【0032】

アプリケーションは、データネットワーク上においてシステムにアクセス可能であるウェブに基づいたインターフェイスを有する。方法は、アプリケーションの複数のウェブに基づいたインターフェイスを自動的に見出すステップと、それぞれのインターフェイスごとに、ウェブに基づいたインターフェイスを介してテンプレートへの入力の値を提供すると共に前記インターフェイスを介してテンプレートの出力を取得するための対応する手順によって1つ又は複数のテンプレートを構成するステップと、を更に含む。

30

【0033】

本発明のその他の特徴及び利点については、以下の説明から、且つ、添付の請求項から、明らかとなる。

【図面の簡単な説明】

【0034】

【図1】スクリプト化されたウェブインターフェイスを使用する対話システムのブロックダイアグラムである。

【図2】ウェブに基づいたグラフィカルユーザーインターフェイスにおけるフォームの提示の概略図である。

40

【図3】図2のフォームのサブミッションの結果の提示の概略図である。

【図4】自動化されたテンプレート構成方式を示す図である。

【図5】自動化されたテンプレート構成方式を示す図である。

【発明を実施するための形態】

【0035】

図1を参照すれば、システム100は、例えば、情報の取得及び／又はサブミッションを目的として、システムとアプリケーションとの間において情報を伝達するべく、ユーザー110が、システムとの間において言語的対話を実施すると共に代表的アプリケーション162A、162Bとやり取りするための方法を提供している。図1においては、アプリケーション162Aは、データベース164A内の情報に対するアクセスを有するか又

50



はこれを提供しており、且つ、ウェブサーバー 160Aを介してインターフェイスを提示している。例えば、ウェブサーバー 160Aは、インターネット 150上におけるグラフィカルウェブクライアント 195（例えば、Microsoft Internet Explorer）からの従来の要求に対して応答している。ユーザー 190は、例えば、マウス及び画面表示を使用することにより、グラフィカルウェブクライアントとやり取りしている。

【0036】

システム 100を使用することにより、別のユーザー 110は、例えば、発話又はタイプ入力された文章の形態において言語的入力を提供する言語的対話システム 120との間においてやり取りしており、且つ、システムは、例えば、言語的な形態によって（例えば、合成された発話又はテキストの表示など）、或いは、その他の出力の形態によって（例えば、データや画像などの画面提示）、応答している。言語的対話システム 120は、しばしば、対話の一部分として、自身が、リモートアプリケーション（例えば、アプリケーション 162A又はアプリケーション 162B）とやり取りする必要がある、従って、そのやり取りを実行するべくスクリプト化されたウェブインターフェイスモジュール 130を使用するものと判定している。ウェブインターフェイスモジュールは、情報を提供するべく、且つ/又は、アプリケーションから情報を取得するべく、ウェブサーバー 160A又はウェブサーバー 160Bとの間のやり取りの方式を制御するべく、インターフェイステンプレート 135を使用している。

【0037】

少なくともいくつかの実施形態においては、言語的対話システム 120は、2016年11月10付で出願された「INTERACTION ASSISTANT」という名称の同時係属中の国際特許出願第PCT/US2016/061303号パンフレットにおいて詳述されている構造を有しており、且つ、この特許文献の内容は、引用により、本明細書において包含される。基礎をなす対話システムは、以下のものに類似した表現を使用することにより、タスク指向の対話をモデル化するものと仮定されている。システムは、自身が対話タスクの一部分を完成させるために使用する小さな「ミニスクリプト」の集合体を有する。これらのミニスクリプトは、カードテンプレート（或いは、省略して、テンプレート）と呼称される。カードは、小さな「構築ブロック」であり、これらから、対話の全体が構築される。例えば、その名称がカードであってもよい場合に、データベース内の特定の人物の連絡先情報を見出すこと、特定の電子メールを読み取ること、クエリによって電子メールを見出すこと、電子メールドラフトを送信すること、電子メールドラフトを作成すること、が、電子メールを送信するべく構成された対話システム内のカードであってもよい。カードテンプレートは、プログラマによって構築されてもよく、或いは、自動的に学習されてもよく、或いは、これら2つの混合であってもよい。後述する方式は、リモートウェブに基づいたアプリケーションにアクセスするという目的のためにシステムによって使用されるカードテンプレートの自動的な生成に関係している。

【0038】

それぞれのカード（即ち、特定のユーザーダイアログ内のカードテンプレートのインスタンス）は、なんらかの予め規定されたタイプの結果を生成し、これは、電子メールメッセージ、連絡先、或いは、電子メールの読取りのような、なんらかの結果をも生まないタスクの完了の証明書を表しうる（タイプは、対話システムにおける、データベースのスキーマの均等物である。好適な実施形態において使用されているタイプシステムは、公式には、MLプログラミング言語のタイプシステムと同一である）。これに加えて、それぞれのカードは、いくつか（恐らくは、ゼロ）の実引数（入力とも呼称される）を必要とする。この場合に、実引数は、我々が名称と呼ぶ実因数は、自身をその他の実引数から区別するためのなんらかの識別子とタイプと、によって規定されている。次いで、カードは、システム外のAPIの呼び出し及びユーザーとの間のやり取りのなんからの組合せにより、結果タイプのオブジェクトを生成するための手順を有する。

【0039】

これらのカードは、1つのカードの結果が別のものに対する実引数の1つとして使用されるように、対話システムの計画モジュールにより、チェーン又はネストされた構造において1つに構成されている。この計画モジュールによって生成される、カードの構造は、「計画」又は「対話スクリプト」と呼称される。これらの計画は、ユーザーの要求又はその他のイベントに応答して構築することができる。計画は、対話システムがどのカードを最初に行うのかに関する部分的な順序又は全体的な順序を定義することができる。

#### 【0040】

システムは、一般に、取得されるデータのそれぞれのタイプごとに、少なくとも1つのインターフェイステンプレートを有する。1つの方式においては、アプリケーション用のテンプレートは、アプリケーションのためにウェブサーバーによって提供されるウェブインターフェイスの生成において使用されるものと推定されるスキーマを含む。スキーマは、そのタイプ内の可能なフィールドを含むタイプの定義を含む。テンプレートの1つの共通的なタイプは、フォームへの入力に関連するウェブページ内の構造と、返される情報の特性と、を規定するフォームであって、グラフィカルインターフェイスにおいてユーザーによって入力されるように意図されたフォームに関係している。テンプレートは、フォームのサブミッションの結果として得られる情報と同様に、フォーム内の異なるスロットに提供されるべき情報がテンプレートにおいて規定されているという意味において、フォームの機能に係る情報を含んでいる。スクリプト化されたウェブインターフェイス130は、情報を取得するべく、テンプレートを使用する。いくつかの例においては、ウェブインターフェイスは、Selenium WebDriverなどのスクリプティングツールを使用している。

#### 【0041】

一例として、航空会社用のウェブフォームの例である、図2を参照すれば、ユーザーは、「ホテル」などではなく、「フライト」を選択し、且つ、「出発」及び「到着」場所と、望ましいフライトのその他の特性と、を入力し、且つ、次いで、「サーチ」を選択することにより、航空会社アプリケーションが、そのデータベースにアクセスし、且つ、どのフライトが要求された基準にマッチングしているのかを判定するようにすることができる。図3を参照すれば、結果的に得られた画面の一例は、フライト情報を表の形態において提示している。図2～図3の例においては、テンプレートは、入力すべきフィールドと、フィールド値が入力されるべき図2のフォームの構造と、を規定しており、且つ、取得された情報が見出される図3の結果的に得られる画面の構造を規定している。

#### 【0042】

この方式によってデータを取得するのに適したテンプレートのライブラリを開発するための1つの方法は、それぞれの手動で作成されたスクリプトがテンプレートと関連付けられている（例えば、これとの関連において又はその内部において保存されている）状態において、取得を自動化するべく必要とされるスクリプトを手動で作成するというものである。このような手動による作成は、有効であることが可能であり、且つ、恐らくは、複雑なウェブサイトの場合には、必要でありうるが、このようなスクリプトが、自動化された又は半自動化された方式によって判定されることの利点が存在している。主な利点は、このようなスクリプトを形成するべく必要とされる人間の作業の低減であり、この結果、対話システムから格段に大きな数のアプリケーションへのアクセスを得ることが可能となる。

#### 【0043】

後述するいくつかの方式は、情報の取得及び/又はサブミッションを自動化するべく使用されるスクリプトの作成に対する純粋に手動的な方式の使用ではなく、自動化された技法を使用している。図4を参照すれば、第1技法は、グラフィカルインターフェイス195のユーザー190又はシミュレートされた対話システム480のユーザー490を伴うウェブに基づいたアプリケーションのウェブサーバー160との間における記録された交換と、記録された交換に基づいたスクリプトの自動化された又は半自動化された推定と、を利用している。第1技法との関連において使用される、第2の技法は、例えば、既知

のアプリケーションと類似した情報を提供する、ウェブに基づいたアプリケーションの特性を発見するべく、アクセス可能なウェブに基づいたアプリケーションの「クローリング」と、自動的なプロービングシステム 470 と、を利用している。

【0044】

図4を更に参照すれば、第1技法においては、システムは、1つ又は複数のウェブクライアント（例えば、グラフィカルウェブクライアント195、或いは、シミュレートされた対話システム480）からのウェブサーバー160（例えば、図1のウェブサーバー160A）との間におけるやり取りのシーケンス（本明細書においては、「実行トレース」と呼称される）412、414を監視及び記録しており、且つ、トレースの分析に基づいてテンプレートライブラリ内においてインストールされうる可能なテンプレートを推定している。それぞれのトレースは、一人のユーザーとウェブサーバーとの間の単一のセッションを表している。更に詳しくは、システムは、実行トレースを収集しており、この場合に、それぞれのトレースは、ウェブページと、テキストボックスへの特定の値の入力、或いは、リンク又はボタンのクリックなどの、システムによってウェブページに対して行われた動作と、のシーケンスを含む。

【0045】

図5を参照すれば、収集された実行トレース410のうちのそれぞれの実行トレースごとに、且つ、1つの定義された出力と、入力の組と、をそれぞれが有する予め定義されたテンプレートの組について、システムのテンプレートコンフィギュレータ420は、実行トレースをテンプレートのインスタンスのシーケンスとアライメントさせている。一般的なフレームワークは、ゼロ以上の数のカードをそれぞれのトレース内のそれぞれの「規則的な構造」に対して割り当てる機械学習モデルを構築するというものであり、且つ、カードは、入力として（そのタイプとしてスキーマを有する）なんらかの数の入力を取得し、且つ、別のスキーマタイプのオブジェクトを「生成」している。

【0046】

例えば、システムが、対話型のHTMLフォーム（HTMLの<form>タグによって示されているもの）、或いは、ウェブページ内のセマンティック要素を恐らくは通知するドキュメント構造（例えば、Kushmerick, Nicholas, “Wrapper induction: Efficiency and expressiveness.” in Proc. Fifteenth International Joint Conference on Artificial Intelligence, Nagoya, Japan, August 23-29, 1997におけるラッパ誘導、手作業で規定されたパターン、或いは、なんらかのその他の手段を通じて識別されたもの）を見出した際に、システムは、そのウェブページの更なる分析によって継続する。ウェブページのこのような部分は、「規則的な構造」と呼称されている。更に正確には、規則的な構造は、システムが解釈しうるウェブページのサブパートであり、このケースにおいては、ラッパ誘導によって識別された抽出された項目と関連するHTMLフォーム又は1つの（又は、複数の）HTMLノードである。それぞれの規則的な構造は、HTMLの<input>要素及びラッパ誘導システムによって識別されたセマンティックフィールドである追加要素を含むことができる。

【0047】

一例として、航空会社のウェブサイト上におけるフライトの予約を検討してみよう。通常のプロセスは、以下のようなものである。

1. 出発と帰着の日付、出発と到着の場所、乗員の数、及び、エコノミーやファーストクラスなどのチケットの「クラス」を入力する。
2. 選択肢のリストから旅程を選択する。
3. 旅行保険などのような更なる選択肢を選択する。
4. （その他の選択肢のいくつかのページが存在しうる。）
5. 旅行者及び支払情報の入力と、フライトのチケットの購入と、がユーザーに対して求められる。

## 【 0 0 4 8 】

これらのステップのそれぞれは、ゼロ以上の数のテンプレートに対応している。システムの仕事は、いずれのテンプレートがそれぞれのステップに対応しているのかを推定するというものである。好適な実施形態においては、ウェブページのそれぞれの規則的な構造は、正確に1つのテンプレートに対応しており、且つ、シーケンス内のそれぞれのページ用のテンプレートは、次のページと関連するテンプレート用の入力のうち少なくとも1つを生成するべく、必要とされている。

## 【 0 0 4 9 】

この例を継続すれば、フライトを予約するプロセスは、いくつかの別個のテンプレートに分解することができる。まず、末尾から始めれば、システムは、「Acme Airlinesのフライトの購入」という値を表すテンプレートを生成する。このテンプレートは、一人又は複数人の旅行者の連絡先情報、ユーザーの支払情報、及び規則的な構造の以前の組の完了を表す特別なトークンタイプという、いくつかの入力を取得する。以前のページ用のテンプレートが、同様に導出され、先行しているページ用のトークンに加えて、入力として使用されるページ上において記述されたなんらかの数の入力タイプが存在している。

## 【 0 0 5 0 】

システムは、様々な方法により、この基本的な手順を実行することにより、テンプレートを生成することができよう。好適な一実施形態においては、それぞれのテンプレートは、異なるウェブサイトからの関係するテンプレートをグループ化するべく使用されるグループテンプレートに基づいている（例えば、1つのウェブサイト上におけるフライトの選択対別のウェブサイト上におけるフライトの選択）。グループテンプレートは、名称、実引数のリスト、及び出力タイプを定義している。この例においては、「select-flight-itinerary」及び「search-flights」がテンプレートであってもよいであろう。ウェブサイトを使用してグループテンプレートから判定される特定のウェブサイト用のテンプレートは、一般に、実引数タイプのサブセットを有しており、且つ、出力タイプは、グループテンプレートのタイプのサブタイプであってもよい。グループテンプレートの組は、既知のタイプのデータベースに由来するものであってもよく、或いは、これは、データから自動的に推定されてもよく、或いは、これら2つの混合であってもよく、この場合には、いくつかのグループテンプレートが、既知のデータベースによって予め入力され、且つ、いくつかは、手作業によって推定される。これらのグループテンプレートは、少なくとも2つの理由から、有用である。第1に、これらは、システムが、見出されたウェブページ内の規則的な構造を解釈することを支援する。第2に、これらは、1つのカードから学習された振る舞いが別のものに「変換」（或いは、さもないければ、マッピング）されうるように、対話システムが関係するテンプレートをクラスタ化することを許容している。好適な実施形態においては、対話システムは、由来するグループテンプレートを通知するそれぞれの推定されたテンプレートごとに機械学習機能（インジケータ機能のみ）を追加することにより、これを実行している。

## 【 0 0 5 1 】

好適な実施形態においては、実行トレース内において表されているウェブページの規則的な構造は、教師なし生成確率モデルを使用することにより、テンプレートに対してアライメントされている。生成モデルによって生成されたアライメントは、規則的な構造内のそれぞれの要素（HTMLの要素又はラッパ誘導からのセマンティックフィールド）を選択されたグループテンプレートに対する実引数のプロパティの1つと関連付けている。例えば、出荷フォーム内の「都市」用のテキストボックスは、グループテンプレート内の「受取人」連絡先実引数の「都市」プロパティと関連付けられうるのである。好適な実施形態においては、生成モデルは、以下のとおりである。

いくつかの数のグループテンプレートを生成する。

それぞれのグループテンプレートpごとに、

名称 $n_{p_a}$ を有するいくつかの入力実引数タイプ $t_{p_a}$ を生成する。

10

20

30

40

50

それぞれのトレース内のそれぞれの規則的な構造  $w$  ごとに、

グループテンプレート  $p_w$  を選択する。

要素の数  $E_w$  を選択する（規則的な構造内の要素は、規則的な構造のケースにおいて、HTMLフォーム内の要素であり、且つ、さもなければ、ラッパ誘導システムによって識別されたフィールドである）。

それぞれの要素  $e_w$   $E_w$  ごとに、

予め放出された要素を条件として、 $p_c$  タイプ  $t_a$  を有するグループテンプレートから入力実引数  $a_e$  を選択する。

その実引数のプロパティ  $p_e$  を選択する。

そのプロパティのタイプ及び入力実引数を条件として、そのタイプ用の要素を放出する。「input」要素の場合には、フォーム内に入力される値も、同様に生成される。

10

好適な実施形態においては、モデルの変数の分布を構成することができるが、既定においては、これらは、以下のようにパラメータ化されている。

タイプは、再帰的プロセスに従って生成される。

システムのタイプデータベース内において予め入力された「primitive」タイプの組が存在している。これらは、ストリング、数値、場所、電子メールアドレス、日付などを含む。

又、連絡先情報及び支払詳細事項のようなもののための「basic」タイプの組も存在している。これらは、次の黒丸において記述されているレコードタイプである。

20

「list」タイプが存在しており、これは、 $N$  個の要素を有し、それぞれが同一のタイプである（タイプ生成器によって再帰的に選択される）。

それぞれが名称及びタイプを有するいくつかのフィールド  $F$ （幾何学的に、或いは、別の分布を通じて、選択されたもの）から構成された「record」タイプが存在している。フィールドのタイプは、生成器を再帰的に呼び出すことによって選択される。

タイプを生成するべく、システムは、primitiveタイプ、new recordタイプ、又はlistタイプを生成するように選択している。

分布は、構造の再使用を促進するべく、アダプタグラム又は類似の数学的オブジェクトを使用することにより、記憶することができる。

タイプの「プロパティ」の組は、タイプの構造から推定されたツリーの葉であり、連絡先のような基本的なタイプがプロパティを有する。

30

システムがグループテンプレートの予め定義された組のみを使用するように構成されている場合には、タイプの組は、グループテンプレート内のタイプの組として固定されており、且つ、再帰的なプロセスは使用されない。

入力実引数  $a_e$  の選択は、ある程度の確率により、同一の実引数を保持する、且つ、さもなければ、別の実引数を均等な方式によってランダムに選択する、分布を有する、入力実引数に対する離散時間マルコフモデルであってもよい。これは、モデルが、実引数から実引数にランダムにジャンプするのではなく、行内の1つの実引数について情報を生成することを選択するようにする。

明示的に言及されていないその他の分布も、幾何分布によって付与されるカウントを除いて、その個々の出力空間上における多項分布であるものと解釈される。

40

【0052】

当然のことながら、その他の分布も、コンポーネントの任意のものについて可能である。

【0053】

この生成モデルを使用することにより、（好適な実施形態において使用されている）メトロポリス-ヘイスティングス（Metropolis-Hastings）に基づいたMCMCサンブラを含む任意の妥当な推定技法を使用し、規則的な構造とグループテンプレートとの間のアライメントを推定することができる。変数  $p_e$  は、カードの構造と規則的な構造との間のアライメントを定義している。このアライメントは、フォームに入力す

50

るべく、且つ、ラッパ誘導によって見出された規則的な構造のオブジェクト情報を抽出するべく、システムが配備される際に使用される。

【0054】

それぞれの規則的な構造のアライメント変数が見出されたら、システムは、アライメント変数において使用されている入力すべてを含む選択されたテンプレートのインスタンスを生成し、且つ、カードは、適切な情報をフィールドに入力すると共にリンクをクリックする又はフォームをサブミットすることにより、その結果を生成する。

【0055】

ウェブサイト用のテンプレートが抽出され、且つ、対話システムがそれらを使用するようにトレーニングされたら、残りの主なステップは、カードに提供された入力を使用してウェブフォームに実際に入力し、且つ、ページ上において既に提供されている実引数の情報を抽出するというものである。上述のように、これは、テンプレートの入力内のプロパティとウェブページ内の要素との間のアライメントを保持することにより、実現することができる。次いで、Seleniumのようなプログラムを使用することにより、適切なデータと共に、任意のフォームをプログラムによってサブミットすることができる。

【0056】

システムは、特定のサイト用のテンプレートのその抽出されたインスタンスをどれだけ確かに有しているのかに関するコンフィデンス尺度を有する。好適な実施形態においては、この尺度は、単に、推定されたカード(induced cards)が付与された場合に、規則的な構造を観察する生成モデルによって演算される対数尤度である。この尺度が、特定のウェブサイト用の所与の閾値を超過していない場合には、ウェブサイトは、システムの拡張された集合体に含まれてはいない。これらは、破棄されてもよく、或いは、人間による分析のためにフラグが付与されてもよい。この閾値は、ビジネスニーズに従って設定されてもよく、且つ、これは、ドメインに伴って変化してもよく、人々は、スポーツの得点を見出すことよりも、銀行サイトとのやり取りの場合に、相対的に高いコンフィデンスを要求することになる。

【0057】

時折、ユーザーは、システムがこれまで参照してはいないサイトの訪問を要求する場合がある。このケースにおいては、システムは、オンデマンドでウェブページをカードに変換するべく試みるように選択することができる。これを実行するために、システムは、いずれのタイプが予想され、且つ、どの種類のカードがページによって表されているのか、を判定するべく、テンプレートアライメントアルゴリズムの1つのバージョンをウェブページに対して実行する。カードがマッチングした場合には、システムは、そのカードを使用する。そうではない場合には、システムの構成に応じて、システムは、ユーザーにエラーを報告してもよく、或いは、人間のエージェントに任せてもよい。

【0058】

いずれにしても、図1に示されているシステムは、テンプレートの組(即ち、インターフェイステンプレート135)によって構成されており、これらは、ユーザー110との間における言語的なやり取りのその分析の一部として、対話システム120によって呼び出すことができる。推定されたスクリプトと、ウェブページの構造(例えば、フォームスロット)との間におけるカードへの入力の関連付けと、が、入力をテンプレートの出力に対してマッピングするためのメカニズムを提供している。

【0059】

図4を再度参照すれば、第2技法は、新しいウェブに基づいたアプリケーションの発見のためのアクセス可能なウェブに基づいたアプリケーションの「クローリング」を伴っている。基本的に、新しいウェブサイト用の実行トレース416は、どのタイプの入力が適切なタイプの応答を導出するのかを理解するための、自動的なプロービングシステム470によるフォームの入力可能な部分に対する自動化されたプロービングに基づいたものであってもよい。

【0060】

10

20

30

40

50

クローラーによって発見されたそれぞれのウェブサイトごとに、システムは、例えば、図2及び図3において示されているページ内などの、そのページ内の規則的な構造の組を抽出している。規則的な構造が適切に識別されたら、システムは、規則的な構造内の対話型要素を操作することにより、実行トレースを収集する。具体的には、システムは、ダミーデータを使用してフォームに入力するべく試み、ラジオボックスがチェックされ、テキストフィールドに適切なタイプのデータ（例えば、ダミー名称、ダミー日付）などが入力される。システムは、Seleniumなどの自動化されたウェブサイトスクリプタを使用することにより、このデータをサブミットする。又、システムは、規則的な構造のHTMLノード内のリンクをクリックすることもできる。システムが応答ページを受け取った際に、システムは、エラーとして、更に多くの情報を要求する同一のタスクに関連する更なるページとして、或いは、「末尾」として、そのページを分類する。この分類は、教師あり学習クラシファイア（サポートベクトル機械など）又は手書きの規則を使用することにより、実行することができる。いずれにしても、システムは、HTTP状態コード及びキーワード（「error」又は「success」など）を含む、返されたページからの特徴を使用する。好適な実施形態においては、システムは、サポートベクトル機械を使用している。これに加えて、システムは、クレジットカード又は（PayPalやGoogle Walletなどのような）その他の支払情報に対する要求に遭遇した際に、停止し、これにより、このようなページを成功として解釈するように構成することができる。

10

#### 【0061】

20

システムは、タスクを完了させる例示用のシーケンスの組を構築するべく、異なるデータにより、同一のフォームを反復的に呼び出す。次いで、システムは、これらのトレースを記録し、これらは、規則的な構造（並びに、これらを含むウェブページ）及びフォームに入力された情報のシーケンスに等しい。

#### 【0062】

このプロセスは、ランダムに実行されてもよく、或いは、好適な実施形態においては、Q学習又はディープQ学習などの強化学習アルゴリズムを使用することによって実行されてもよい。強化学習においては、システムは、それぞれのステップにおいて可能な動作の組から1つの動作を取得し、且つ、システムは、定期的なインターバルにおいて報酬信号を受け取る。このセットアップにおける動作は、異なるフォーム入力における入力と、フォームのサブミットと、である。報酬信号は、肯定的又は否定的なものであってもよく、システムが、サーバーから、エラーコード又は予め規定された条件にマッチングしている（例えば、返されたページが「error」という単語を含む）メッセージを受け取った際には常に、システムは、否定的報酬を取得する。その一方において、フォームがエラーなしに進行した場合には、システムは、（システムが無限ループを辿ることを阻止するべく）わずかに否定的な報酬又はゼロの報酬を受け取る。正常な完了ページに到達した場合には、システムは、大きな肯定的な報酬を受け取る。これらの値は、構成することができる。

30

#### 【0063】

正常なトレースが、この自動化されたプロセスによって得られたら、システムは、ウェブに基づいたアプリケーションの対話型の使用の際に収集されるトレースの場合と基本的に同一の方法により、これらを使用する。即ち、トレースは、カードテンプレートのインスタンスに対してアライメントされ、且つ、カードテンプレートは、入力からカードの出力へのマッピングをもたらずウェブのやり取りを実行するように構成されている。

40

#### 【0064】

言語的インターフェイスを使用したウェブサイトとのやり取りの方法の自動的な発見のための上述のものにそれなりに関係する方式においては、プロセスは、（例えば、電話オペレータを介して）グラフィカルなインターフェイスを音声に基づいた人間のインターフェイスに提供するべく、或いは、音声に基づいた人間のインターフェイスとやり取りするべくシステムによって仲介された言語的インターフェイス（例えば、テキスト又は音声入

50

力)を提供するべく、多少、「反転」されてもよい。即ち、システムは、異なる言語的インターフェイスを介して、ユーザーのみならず、「アプリケーション」とインターフェイスしてもよく、且つ、基本的に、やり取りを変換又は仲介してもよい。例えば、現在、「オンライン」ツールのみを使用することにより、ヘアカットの約束を予約することは、或いは、近所の花屋に花を注文することは、困難でありうる。その代わりに、これらの種類のサービスは、通常、電話を介して調達されている。

#### 【0065】

実際に、電話を使用して約束をスケジュールリングする且つサービスを注文する能力は、人間のアシスタントのかなり一般的な責務である。従って、人間のアシスタントの能力の多くを複製することを狙いとした汎用対話システムは、そのユーザーの代わりに、このよ

10

#### 【0066】

例えば、対話システムのユーザーが、特定の歯科クリニックにおいて歯のクリーニングを予約するように要求しうるのであろう。ユーザーの要求に応答して、システムは、データベース内のビジネスの連絡先情報をルックアップし、且つ、恐らくは、リストの内容を確認した後にユーザーの要求及びビジネスのタイプにマッチングしているカードを選択する。このケースにおいては、システムは、「book-medical-appointment」カードを選択してもよく、このカードは、クリニック、約束のタイプ(「歯のクリーニング」)を予想している。次いで、システムは、歯科クリニックに電話をかけることになり、且つ、テキスト/発話変換システムを使用することにより任意選択によって自身を自動化されたエージェントであるとして識別することになり、且つ、次いで、システムの言語生成設備を使用してユーザーの要求を伝えることになろう。この時点において、システムは、いまや、通話を受け取った商人のための対話エージェントとして機能しており、システムは、商人に話すことによってやり取りを完了するべく、別のカード(例えば、「book-merchant-medical-appointment」)を使用する。約束を予約するケースにおいては、これには、日程管理用の予め定義されたカードを使用することによってユーザーのカレンダー上において空き時間をスケジュールリングすることが伴いうるのであろう。概念的には、このプロセス全体は、商人に接続すると共に結果を第1対話セッションに返送する異なる対話セッションのインスタンス生成であるものとして考えることができる。

20

30

#### 【0067】

それぞれのドメインごとに、システム構築者は、「user」カード及び「merchant」カードという、2つのカードの開発者による生成をサポートすることを所望している。「user」カードは、要求を開始するべく必要とされる入力(ビジネスの連絡先情報、カレンダーの制約、望ましい物品又はサービスなど)を規定している。そして、ユーザーカードは、「merchant」カードを呼び出している。商人カードは、要求内の情報を伝達するためのスクリプトを規定している。入力として、これは、(ユーザーカードによってコンパイルされた)ユーザーの要求と、提供するべく商人に必要とされるすべての情報と、を取得している。その結果、これは、確認又は失敗通知を返し、これは、次いで、ユーザーカードに返送される。更に多くの情報に対する要求は、システムが電子メールを処理するべく受け取る「message sent」イベントに類似したイベントとして、「user」と「merchant」との間において伝達される。これらのイベントは、カードにより、必要であるものと解釈される。

40

#### 【0068】

テレフォニー対応型システム用のデータ収集は、ウェブに基づいたシステム用のデータ収集手順に類似している。但し、このタスクのために収集される実行トレースは、人間によって実施されたセッションからのみ収集され、この場合には、人間が、実際に、商人に対して電話をかけている(我々は、我々のシステムが商人又はその他の人々にランダムに

50



通話することを所望してはいないことから、我々は、ウェブフォームについて概説されている自動的な発見手順と類似した内容を試みてはいない)。トレーニングを目的として、実行トレースのそれぞれの半分は、別個の対話として取り扱われている。さもないと、トレーニングは、「オズの魔法使い」方式の場合には、先程概説したトレーニングと同一である。

#### 【0069】

配備された際に、対話システムは、いまや、トレーニングフェーズにおいて学習された更なるカードに対するアクセスを有する。ユーザーがカード記述のうちの1つとマッチングした要求を実施した場合に、システムは、ユーザーカードを起動し、要求に関係する更なる情報を収集し(例えば、利用可能な日付及び時刻について問い合わせる、或いは、ユーザーのカレンダーをチェックする)、(利用可能な場合に、ユーザーの予め登録された好みを使用することによって)適切なビジネスを見出し、恐らくは、ユーザーにビジネス情報について確認し、且つ、次いで、業者に電話をかける。

10

#### 【0070】

商人が応答した場合に、システムは、商人との間におけるそのやり取りにおいて商人カードを配備する。システムは、ユーザーの要求を商人に伝達し、且つ、次いで、要求を完了させるべく、カードのその他のステップを実行する。完了したら、システムは、ユーザーに通知する。

#### 【0071】

現実世界との間のやり取りは、(当然のことながら)構造化されたウェブページとの間のやり取りよりも、格段に複雑であることから、システムは、いくつかの異なる方法により、会話を高度化させる能力を有する。商人との間においてやり取りしている対話システムが、(ユーザーのフラストレーションの標準的な尺度を使用して)対話が望ましくない方式によって進行していると判定した場合には、システムは、(例えば、謝罪し、且つ、電話を切ることによって)やり取りを中止するように、(例えば、更なる時刻について問い合わせることによって)ユーザーに更に多くの情報を要求するように、通話を人間のエージェントに転送するように、或いは、通話をユーザーに直接的に転送するように、選択することができる。選択は、機械学習クラシファイアを使用することにより、要求のタイプ及びユーザーの好みによってパラメータ化されている。この高度化の振る舞いは、人間のエージェントのみならず、ユーザーにも高度化されうるという点を除いて、ユーザーとの間のやり取りにおいて対話システムによって使用されうる振る舞いに類似している。

20

30

#### 【0072】

いくつかのタスクは、物品の配送を必要としうる。例えば、近所の花屋は、その独自の配送サービスを有していない場合がある。又、これらが利用可能である場所において、システムは、クーリエサービス又はTask Rabbitのようなローカルな物流管理又は「micro-task」会社にコンタクトすることもできる。これは、更なる(恐らくは、並行した)電話通話を使用することにより、(本明細書において上述されている)学習されたウェブサービス又はこの目的のためのカスタム作成されたAPIを使用し、完了させることができる。好適な実施形態においては、システムは、マーケットに応じて、これら3つのものの混合を使用している(例えば、Task Rabbitは、すべての場所において利用可能であるわけではない)。

40

#### 【0073】

やり取りアプリケーション用のカードテンプレートが取得されうるトレースを取得する更にその他の方法が存在していることを理解されたい。例えば、グラフィカルウェブクライアント195を介してやり取りしているユーザー190と関連するトレースを収集するのではなく、自動化された音声対話システム120の「オズの魔法使い」シミュレーションが使用されてもよく、この場合には、人間のオペレータが、システムがまだ構成されていないことについて意思決定し、或いは、オペレータが、システムが誤った決定を下すことを観察している。

#### 【0074】

50

又、ウェブに基づいたアプリケーションは、このような方式によって自動化された対話システムからインターフェイスされるアプリケーションの一例に過ぎないことを理解されたい。例えば、アプリケーションが、データネットワーク（例えば、インターネット）上においてアクセス可能である、或いは、任意の特定のプロトコル又はマークアップ言語（例えば、http及びHTML）を使用する、というニーズは存在しておらず、基本的に、カードテンプレートの発見のための同一の手順は、様々な種類のアプリケーションに使用することができる。

【0075】

システムの実装形態は、1つ又は複数のコンピュータを制御する（一時的ではない機械可読媒体上において保存された）命令を含むソフトウェアを使用することができる。例えば、図1に示されている機能は、単一のコンピュータ上において実行されてもよく、或いは、例えば、データネットワーク上において通信している複数のコンピュータ上においてパーティション化及び実行されてもよい。いくつかの複数コンピュータの実装形態においては、特定のコンポーネントは、ユーザーの場所におけるコンピュータ又はその他の演算装置（例えば、発話統合コンポーネント）においてホスティングされてもよく、その他のコンポーネントは、ユーザーから離れた1つ又は複数の場所におけるサーバーコンピュータにおいてホスティングされてもよい。上述の収集されたトレースに従って、トレースを収集し、テンプレートを構成する機能は、更にその他のコンピュータ上においてホスティングされてもよい。

【0076】

上述の説明は、添付の請求項の範囲によって定義されている本発明の範囲を限定するものではなく、例示を目的としていることを理解されたい。その他の実施形態も、添付の請求項の範囲に含まれる。

【図1】

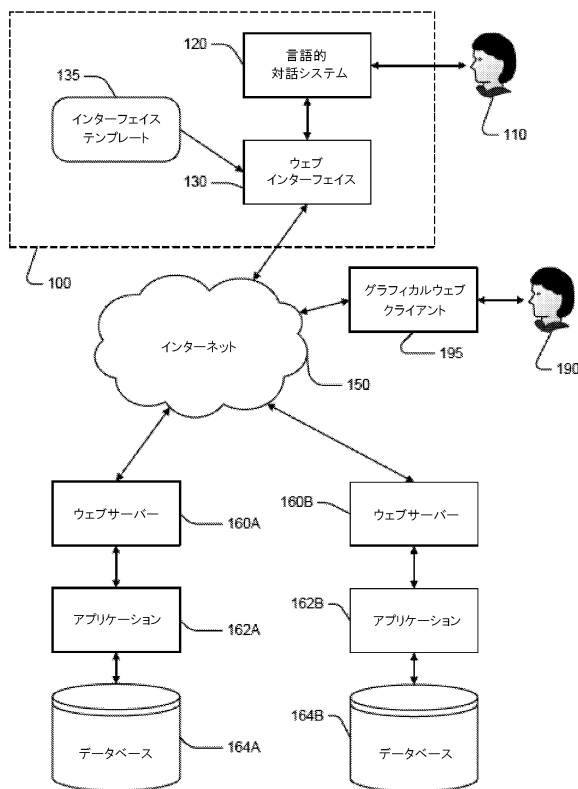


FIG. 1

【図2】

FIG. 2

【図 3】

ACME AIRLINES					
日付	出発	到着	出発地	到着地	価格

FIG. 3

【図 4】

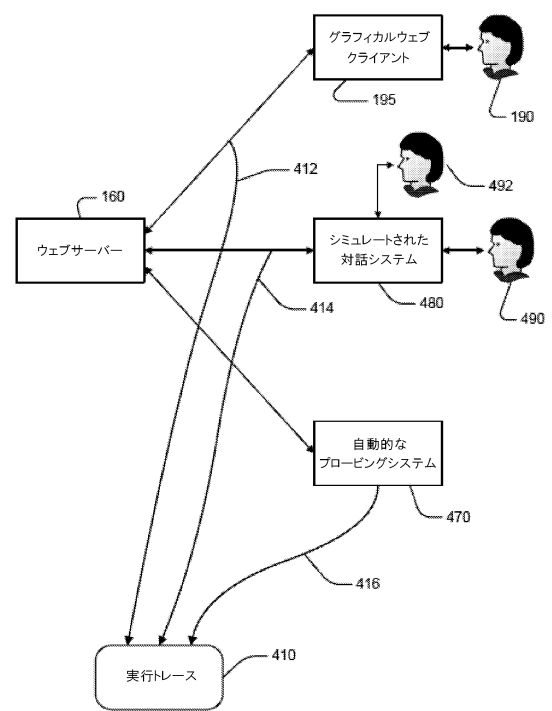


FIG. 4

【図 5】

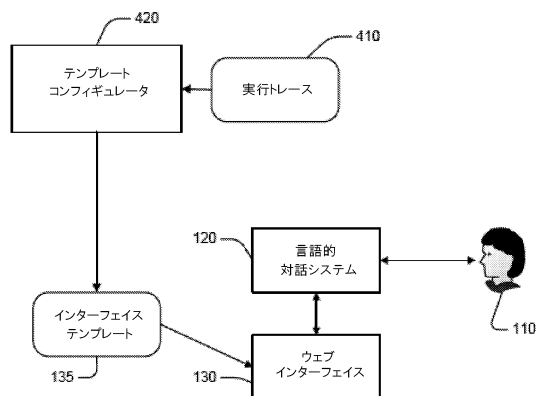


FIG. 5

## フロントページの続き

- (74)代理人 100108213  
弁理士 阿部 豊隆
- (72)発明者 チェン, ペンユ  
アメリカ合衆国 カリフォルニア州 95014 クパチーノ 4 ピアードン ドライブ 10  
228
- (72)発明者 コーエン, ジョーダン アール.  
アメリカ合衆国 ノースカロライナ州 28449 クル ビーチ ノース 4番 アベニュー  
209 エイ
- (72)発明者 ギリック, ローレンス エス.  
アメリカ合衆国 マサチューセッツ州 02459 ニュートン エッツィ ファーム ロード  
55
- (72)発明者 ホール, デビッド レオ ライト  
アメリカ合衆国 カリフォルニア州 94710 バークレー 8番 ストリート 1904
- (72)発明者 クレイン, ダニエル  
アメリカ合衆国 カリフォルニア州 94563 オリンダ フォーリン リーフ テラス 20
- (72)発明者 ポールス, アダム ディー.  
アメリカ合衆国 カリフォルニア州 94704 バークレー アpartment エイ ヘイスト  
ストリート 1920
- (72)発明者 ロス, ダニエル エル.  
アメリカ合衆国 マサチューセッツ州 02459 ニュートン マーシャル ストリート 16
- (72)発明者 ルサック, ジェシー ダニエル エスケス  
アメリカ合衆国 マサチューセッツ州 02143 サマービル ユニット 1 カンブリア ス  
トリート 26

審査官 日比野 可奈子

- (56)参考文献 特開2013-080476(JP,A)  
米国特許出願公開第2010/0299136(US,A1)

## (58)調査した分野(Int.Cl., DB名)

G06F3/01  
G06F3/048-3/0489  
G06F3/16  
G10L15/22