



(86) Date de dépôt PCT/PCT Filing Date: 2014/06/05
(87) Date publication PCT/PCT Publication Date: 2014/12/24
(45) Date de délivrance/Issue Date: 2022/04/12
(85) Entrée phase nationale/National Entry: 2015/11/26
(86) N° demande PCT/PCT Application No.: US 2014/041014
(87) N° publication PCT/PCT Publication No.: 2014/204655
(30) Priorité/Priority: 2013/06/21 (US13/923,917)

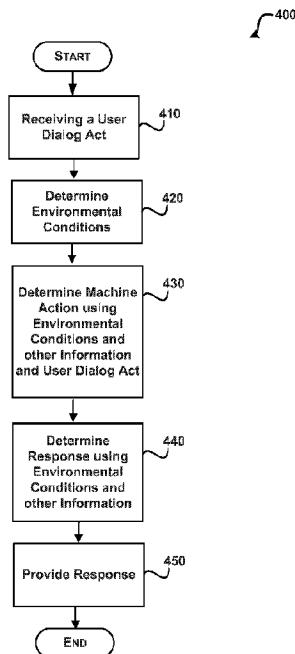
(51) Cl.Int./Int.Cl. G06F 40/35 (2020.01),
G10L 15/22 (2006.01)

(72) Inventeurs/Inventors:
BOIES, DANIEL, US;
HECK, LARRY, US;
ANASTASAKOS, TASOS, US;
SARIKAYA, RUHI, US

(73) Propriétaire/Owner:
MICROSOFT TECHNOLOGY LICENSING, LLC, US

(74) Agent: SMART & BIGGAR LLP

(54) Titre : POLITIQUES DE DIALOGUES ET GENERATION DE REPONSES SENSIBLES A L'ENVIRONNEMENT
(54) Title: ENVIRONMENTALLY AWARE DIALOG POLICIES AND RESPONSE GENERATION



(57) Abrégé/Abstract:

Environmental conditions, along with other information, are used to adjust a response of a conversational dialog system. The environmental conditions may be used at different times within the conversational dialog system. For example, the environmental conditions can be used to adjust the dialog manager's output (e.g., the machine action). The dialog state information that is used by the dialog manager includes environmental conditions for the current turn in the dialog as well as environmental conditions for one or more past turns in the dialog. The environmental conditions can also be used after receiving the machine action to adjust the response that is provided to the user. For example, the environmental conditions may affect the machine action that is determined as well as how the action is provided to the user. The dialog manager and the response generation components in the conversational dialog system each use the available environmental conditions.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(10) International Publication Number

WO 2014/204655 A1

(43) International Publication Date
24 December 2014 (24.12.2014)

(51) International Patent Classification:
G06F 17/27 (2006.01)

(72) Inventors: **BOIES, Daniel**; c/o Microsoft Corporation, LCA - International Patents (8/1172), Redmond, Washington 98052-6399 (US). **HECK, Larry**; c/o Microsoft Corporation, LCA - International Patents (8/1172), Redmond, Washington 98052-6399 (US). **ANASTASAKOS, Tasos**; c/o Microsoft Corporation, LCA - International Patents (8/1172), Redmond, Washington 98052-6399 (US). **SARIKAYA, Ruhi**; c/o Microsoft Corporation, LCA - International Patents (8/1172), Redmond, Washington 98052-6399 (US).

(21) International Application Number:
PCT/US2014/041014

(74) Agent: **LEONARD, Christopher J.**; (USOC - Merchant & Gould P.C.), Microsoft Corporation, LCA - International Patents (8/1172), Redmond, Washington 98052-6399 (US).

(22) International Filing Date:
5 June 2014 (05.06.2014)

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

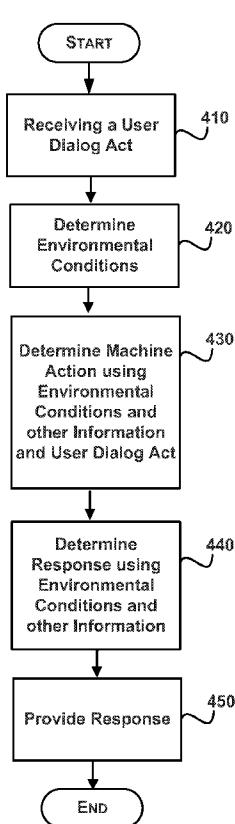
(25) Filing Language:
English(26) Publication Language:
English

(30) Priority Data:
13/923,917 21 June 2013 (21.06.2013) US

(71) Applicant: **MICROSOFT CORPORATION** [US/US];
One Microsoft Way, Redmond, Washington 98052-6399 (US).

[Continued on next page]

(54) Title: ENVIRONMENTALLY AWARE DIALOG POLICIES AND RESPONSE GENERATION



(57) **Abstract:** Environmental conditions, along with other information, are used to adjust a response of a conversational dialog system. The environmental conditions may be used at different times within the conversational dialog system. For example, the environmental conditions can be used to adjust the dialog manager's output (e.g., the machine action). The dialog state information that is used by the dialog manager includes environmental conditions for the current turn in the dialog as well as environmental conditions for one or more past turns in the dialog. The environmental conditions can also be used after receiving the machine action to adjust the response that is provided to the user. For example, the environmental conditions may affect the machine action that is determined as well as how the action is provided to the user. The dialog manager and the response generation components in the conversational dialog system each use the available environmental conditions.

WO 2014/204655 A1



EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR, OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

Published:

- with international search report (Art. 21(3))

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, VN, ZA, ZM, ZW).

**ENVIRONMENTALLY AWARE DIALOG POLICIES AND RESPONSE
GENERATION**

BACKGROUND

[0001] Conversational dialog systems are used for a variety of different applications. For 5 example, conversational dialog systems have been developed to allow users to search for content, buy items, obtain directions, and the like. These conversational dialog systems are continually being improved to allow for more natural interactions to occur between a computer and a user.

SUMMARY

10 [0002] This Summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used as an aid in determining the scope of the claimed subject matter.

[0003] Environmental conditions, along with other information, are used to adjust 15 different operations of a conversational dialog system. The environmental conditions may be used at different times within the conversational dialog system. The dialog manager and the response generation components in the conversational dialog system each use the available environmental conditions along with the other information in the dialog system (e.g., previous turn language understanding results in the dialog, speech recognition results and confidences, 20 knowledge results, previous machine actions, session context, client context ...) when performing their operations. For example, the environmental conditions can be used to adjust the dialog manager's output (e.g., the machine action). The dialog state information that is used by the dialog manager includes environmental conditions along with other dialog information for the current turn in the dialog as well as environmental conditions for one or 25 more past turns in the dialog. The environmental conditions can also be used after receiving the machine action to adjust the response that is provided to the user. For example, the response generator may use the environmental conditions to affect how the machine action is provided to the user (e.g., speech, visual).

[0003a] According to one aspect of the present invention, there is provided a computer-implemented method, performed by at least one processor, for using environmental conditions in addition to other dialog state information in a conversational dialog system, comprising: receiving a user dialog act requesting a computer system to perform a task; based on the user dialog act, determining a first machine action to perform the task; accessing environmental conditions that are associated with the user dialog act; based on the environmental conditions and the user dialog act, determining a second machine action other than performing the task, wherein the second machine action at least temporarily prevents completion of the first machine action; and performing the second machine action to provide a response to a user.

5

[0003b] According to another aspect of the present invention, there is provided a computer-readable storage device storing computer-executable instructions for using environmental conditions in addition to other dialog state information in a conversational dialog system, comprising: receiving a user dialog act requesting a computer system to perform a task; based on the user dialog act, determining a first machine action to perform the task; accessing environmental conditions that are associated with the user dialog act; based on the environmental conditions and the user dialog act, determining a second machine action other than performing the task, wherein the second machine action at least temporarily prevents completion of the first machine action; and performing the second machine action to provide a response to a user.

10

[0003c] According to still another aspect of the present invention, there is provided a system for using environmental conditions in addition to other dialog state information in a conversational dialog system, comprising: a processor and memory; an operating environment, executing using the processor; and a dialog manager that is configured to perform actions comprising: receiving a user dialog act requesting a computer system to perform a task; based on the user dialog act, determining a first machine action to perform the task; accessing environmental conditions that are associated with a user dialog act; determining a second machine action, other than performing the task, using the user dialog act and the environmental conditions, wherein the second machine action at least temporarily prevents completion of the first machine action; and performing the second machine action to provide a response to a user.

15

20

25

30

[0003d] According to yet another aspect of the present invention, there is provided a system comprising: at least one processor; and a memory storing instructions that when executed by the at least one processor perform a set of operations comprising: receiving a user dialog act requesting a computer system to perform a task involving a contact;

5 accessing environmental conditions that are associated with the contact, wherein the environment conditions include at least one of a noise level, a time zone, or a temperature; based on the environmental conditions and the user dialog act, determining a machine action to display a suggestion to send a text-based message to the user; and performing the machine action to display the suggestion to the user.

10 **[0003e]** According to a further aspect of the present invention, there is provided a computer-implemented method, performed by at least one processor, for using environmental conditions of a contact in a conversational dialog system, comprising: receiving a user dialog act requesting a computer system to perform a task involving the contact; accessing environmental conditions that are associated with the contact; determining from the 15 environmental conditions that the contact is in at least one of the following conditions: a roaming service area or a time zone where the time is night; based on the determination from the environmental conditions, determining a machine action; and performing the machine action to provide a response to the user.

[0003f] According to yet a further aspect of the present invention, there is provided a 20 storage device storing computer-executable instructions for using environmental conditions of a contact in a conversational dialog system, the instructions comprising: receiving, from a user, a user dialog act requesting a computer system to call the contact; accessing environmental conditions that are associated with the contact; based on the environmental conditions and the user dialog act, determine a suggested action other than calling the contact; 25 providing the suggested action to the user, wherein the suggested action is sending a text-based message the contact; receiving further input from the user; and based on the further input, sending a text-based message to the contact.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] FIGURE 1 illustrates using environmental conditions to adjust different operations performed by a conversational dialog system;

[0005] FIGURE 2 shows interactions between an application and a dialog service for
5 configuring permissions;

[0006] FIGURE 3 illustrates an exemplary conversational dialog system using environmental conditions;

[0007] FIGURE 4 illustrates a process for using environmental conditions in a conversational dialog system;

5 [0008] FIGURE 5 illustrates an exemplary online system that uses environmental conditions in determining a machine action and how to deliver a response in a dialog; and

[0009] FIGURES 6, 7A, 7B and 8 and the associated descriptions provide a discussion of a variety of operating environments in which embodiments of the invention may be practiced.

10

DETAILED DESCRIPTION

[0010] Referring now to the drawings, in which like numerals represent like elements, various embodiment will be described.

[0011] FIGURE 1 illustrates using environmental conditions to adjust different operations performed by a conversational dialog system.

15

[0012] As illustrated, system 100 includes dialog service 105, dialog manager 26, permissions and settings 110, machine action generator 120, dialog policy 130, response generator 140, dialog state information 145, environmental conditions 150, understanding model(s) 160, application 110 and touch screen input device 115.

20

[0013] Generally, conversational dialog system 100 uses environmental signals (e.g., language understanding and knowledge results, speech recognition results and confidences ...) for policy and response generation in addition to the use of other information or signals typically available in the dialog system. The environmental signals augment (not replace) the signals typically used within a dialog system. Dialog manager 26 uses environmental conditions 150 in addition to dialog state information 145 to adjust different operations of a conversational dialog system. The environmental conditions may be used at different times within the conversational dialog system 100. The dialog manager 26, machine action generator 120, and the response generator 140 components in the conversational dialog system 100 can each use the available environmental conditions 150 when performing their operations.

25

[0014] Machine action generator 120 can use environmental conditions 150 to adjust the machine action that is determined along with using dialog state information 145. Generally, the environmental conditions 150 augment the dialog state information 145 that is used within the dialog system. For example, dialog state information 145 includes information, such as, but not limited to: e.g., previous turn language understanding results

in the dialog, speech recognition results and confidences, knowledge results, previous machine actions, session context, client context, and the like. Each interaction between the user and the computing device (i.e., the machine) is a turn. The current turn is the current interaction between the user and the machine for which a machine action is being

5 determined. A previous turn is a previous interaction between the user and the machine. According to an embodiment, the dialog state information 145 includes information relating to one or more previous turns. For example, language understanding results that associate a meaning to a natural language sentence are stored in dialog state information 145. These previous language understanding results may be used to provide a context for

10 the current turn in the dialog. The dialog state information may also include knowledge results that are the results returned from a knowledge source, such as a database or data source. For example, if a user requests movies in a dialog turn, the knowledge results include a list of movies. The machine action that is determined using the environmental conditions may be different than the machine action that is determined without using the

15 environmental conditions. For example, the machine action that is generated by machine action generator 120 may indicate to ask the driver if they are sure they would like to make a call while driving when environmental conditions 150 indicate that the user is currently driving. Without the use of environmental conditions 150 to augment the machine action, the machine action generator 120 would likely generate a machine action

20 to automatically initiate the call while the user is currently driving instead of asking the driver if they are sure they would like to make a call while driving. Machine action generator 120 uses dialog policy 130 when determining the machine action. The dialog policy 130 includes different rules, including rules that use environmental conditions 150 and other dialog state information, to adjust the machine action that is generated.

25 [0015] Environmental conditions 150 can also be used by response generator 140 after receiving the machine action to adjust the response that is provided to the user. Environmental conditions 150 may affect how the machine action is provided to the user (e.g., speech, visual ...). For example, the response generated by response generator 140 may be a visual response when environmental conditions 150 indicate that the user's

30 environment is noisy. The response generated by response generator 140 may be an auditory response when environmental conditions 150 indicate that the user's environment is very bright and it is unlikely that a display may be seen clearly.

[0016] Environmental conditions 150 may include a variety of different conditions, such as, but not limited to: weather conditions, traffic conditions, current activity conditions, a

current computing device being used, and the like. The current activity conditions may include information, such as, but not limited to: a current speed, a current location, and the like. Generally, the environmental conditions relate to any environmental conditions that are associated with a user providing the input (the user dialog act). According to an embodiment, environmental conditions may also be associated with another user 154 (e.g., a contact that is being called). According to an embodiment, the environmental conditions that are collected and used to adjust the operation of the conversational dialog system are specified and authorized by the user.

[0017] In order to facilitate communication with the dialog manager 26, one or more callback routines, may be implemented. According to an embodiment, application program 110 is a multimodal application that is configured to receive speech input and input from a touch-sensitive input device 115 or other input devices. For example, voice input, keyboard input (e.g., a physical keyboard and/or SIP), video based input, and the like. Application program 110 may also provide multimodal output (e.g., speech, graphics, vibrations, sounds, ...). Dialog manager 26 may provide information to/from application 110 in response to user input (e.g., speech, gesture). For example, a user may say a phrase to identify a task to perform by application 110 (e.g., selecting a movie, buying an item, identifying a product, ...). Gestures may include, but are not limited to: a pinch gesture; a stretch gesture; a select gesture (e.g., a tap action on a displayed element); a select and hold gesture (e.g., a tap and hold gesture received on a displayed element); a swiping action and/or dragging action; and the like. Gestures may also include optically captured gestures, such as: a wave gesture, a scroll gesture, a guide gesture, and the like. For example, a device, such as MICROSOFT KINECTTM may be used to detect gestures.

[0018] System 100 as illustrated comprises a touch screen input device 115 that detects when a touch input has been received (e.g., a finger touching or nearly touching the touch screen). Any type of touch screen may be utilized that detects a user's touch input.

[0019] A natural user interface (NUI) and/or some other interfaces may be used to interact with dialog service 105. For example, a combination of a natural language dialog and other non-verbal modalities of expressing intent (e.g., gestures such as the gestures described above, touch, gaze, images, videos, spoken prosody, etc.) may be used to interact with dialog service 105. Dialog manager 26 may use an understanding model (e.g., a Language Understanding (LU) model or a multimodal understanding model). As illustrated, dialog manager 26 is part of a dialog service 105 that receives speech utterances and is configured to have a dialog with a user.

[0020] Generally, Natural Language Understanding (NLU) in goal-oriented dialog systems is directed at identifying the domain(s) and intent(s) of the user, as expressed in natural language (NL), and to extract associated arguments or slots. For example, in an airline domain, users often request flight information (e.g., “I want to fly to Boston from 5 New York next week”). More details are provided below.

[0021] FIGURE 2 shows interactions between an application and a dialog service for configuring permissions.

[0022] As illustrated, FIGURE 2 shows dialog service 105 including settings 210 and log(s) 212 and application 110 that includes settings user interface 222, audit user 10 interface 226, location user interface 228 and local settings 224.

[0023] A user may configure different criteria, such as permissions, filters and settings that are related to the operation of the application 110 and to the use and collection of environmental conditions that are associated with one or more users. For example, an authorized user may determine what environmental conditions may be collected and used 15 by dialog service 105.

[0024] The user may set permissions that specify what environmental conditions may be collected. For example, a user may opt in to allow different types of environmental conditions to be collected and used by dialog service 105.

[0025] A user may also configure one or more filters that filter the information using 20 various criteria (e.g., time of day, requesting device, requesting location, and the like). For example, a user may allow certain environmental conditions to be collected during certain times and not allow environmental conditions to be collected during other times. A user may configure a filter to limit the accuracy of one or more of the environmental conditions. For example, location information that is associated with a user may be 25 configured to provide a general location as compared to detailed location. A user may also set local settings 224 such as application customization settings (e.g., color schemes, icons and the like) for use when interacting with application 110. For example, the settings 224 may change an appearance of a User Interface, such as UI 228.

[0026] Audit user interface 226 may be used by an authorized user to view and interact 30 with information relating to environmental conditions that are stored in log 212. Access to the log related data is limited to an authorized user who has provided the correct security credentials. The data is stored securely within log(s) 212.

[0027] FIGURE 3 illustrates an exemplary conversational dialog system using 5 environmental conditions.

[0028] As illustrated, system 300 includes input manager 310, language understanding component 320, dialog engine 330, response generator 140, and environment engine 350.

[0029] Input manager 310 is configured to receive a user input (the user dialog act). The user dialog act may be different types of user input. For example, the user dialog act may be speech input (e.g., an utterance), touch input, gesture input, text input, as well as other types of input. Input manager 310 may include different components, such as an Automatic Speech Recognizer (ASR), a gesture recognizer, a handwriting recognizer, and the like. When the user dialog act is speech, the speech is recognized by input manager 310 using ASR. Input manager 310 outputs text that is delivered to language understanding component 320.

[0030] Language understanding component 320 is configured to receive and analyze the text. Generally, the language understanding component 320 associates a meaning to the words of the text. For example, the language understanding component 320 may tag the text, perform a syntactic analysis of the text, and the like. The language understanding component 320 is traditionally a rules based component that converts natural language sentences into semantic representations.

[0031] Dialog engine 330 manages the dialog by determining a machine action in response to the user dialog act. Dialog engine 330 may be a rules based model or a statistical model, or some combination of rules and statistical models. Dialog engine 330 computes possible dialog states using dialog state computation component 332. Dialog state computation component 332 uses environmental conditions obtained from environment engine 350 in addition to other dialog state information when calculating the possible dialog states. The environmental conditions that are used by the dialog system are in addition to other dialog information generally used within a dialog system, such as but not limited to: previous turn language understanding results, speech recognition results and confidences, knowledge results, previous machine actions, session context, client context, and the like. The environmental conditions may be associated with one or more users. Dialog manager 26 determines the machine action or machine actions. The dialog state information that is used by dialog manager 26 includes environmental conditions for the current turn in the dialog as well as environmental conditions for one or more past turns in the dialog. Dialog engine 330 provides the determined machine action to response generator 140.

[0032] Response generator 140 uses environmental conditions that are obtained from environment engine in addition to the other dialog information to determine what type of

response to provide the user. For example, the response generator 140 may determine to provide a spoken response to the user or a visual response to the user. For example, a Text-to-Speech ("TTS") component may output the response as synthesized speech when determined. According to an embodiment, the environmental conditions are updated after 5 dialog manager 26 determines the machine action and before response generator 140 receives the environmental conditions.

[0033] FIGURE 4 illustrates a process for using environmental conditions in a conversational dialog system. When reading the discussion of the routines presented herein, it should be appreciated that the logical operations of various embodiments are 10 implemented (1) as a sequence of computer implemented acts or program modules running on a computing system and/or (2) as interconnected machine logic circuits or circuit modules within the computing system. The implementation is a matter of choice dependent on the performance requirements of the computing system implementing the invention. Accordingly, the logical operations illustrated and making up the embodiments 15 described herein are referred to variously as operations, structural devices, acts or modules. These operations, structural devices, acts and modules may be implemented in software, in firmware, in special purpose digital logic, and any combination thereof.

[0034] After a start operation, the process 400 moves to operation 410, where a user 20 dialog act is received. The user dialog act may be different types of user input, such as, but not limited to: speech input (e.g., an utterance), touch input, gesture input, text input, as well as other types of input. Generally, the user dialog act is part of a conversation with a dialog system.

[0035] Transitioning to operation 420, environmental conditions are determined. The 25 environmental conditions are associated with one or more users. For example, the environmental conditions may be current environmental conditions that are associated with a user making the user dialog act. The environmental conditions may also be current environmental conditions that are associated with another user that is a part of the user dialog act. For example, the user dialog act may be an act to contact another user. The environmental conditions may include a variety of different conditions, such as, but not 30 limited to: weather conditions (e.g., temperature, pressure ...), traffic conditions, current activity conditions, a current computing device being used, and the like. Generally, any environmental condition may be determined and used by the conversational dialog system. The current activity conditions may include a variety of different environmental conditions, such as, a location of the user, a velocity of the user, and the like.

[0036] Moving to operation 430, the machine action is determined using the environmental conditions, other dialog state information, and the user dialog act.

According to an embodiment, the environmental conditions are used in addition to the other dialog state information by the conversational dialog system when calculating the

5 possible dialog states. For example, the other dialog state information may include but is not limited to: previous turn language understanding results, speech recognition results and confidences, knowledge results, previous machine actions, session context, client context, and the like. The dialog state information includes environmental conditions for the current turn in the dialog as well as environmental conditions for one or more past turns in the dialog.

10 [0037] The following examples illustrate adjusting the machine action using environmental conditions and the other information in the dialog system to determine the machine action. The examples are for illustration purposes and are not intended to be limiting.

15 [0038] The following is an example dialog while a user is driving. 1) User: "Call Paul at home"; 2) Dialog System: "Are you sure you want to make the call while driving?" As can be seen, instead of the machine action being to make the call, the machine action was adjusted to request whether the user wanted to make the call.

20 [0039] The next example is a dialog where a user attempts to call a contact in a time zone where the contact is likely sleeping. 1) User: "Call Paul"; 2) Dialog System: "Paul is likely sleeping right now. Would you like to send a message to Paul instead?" In this example, the environmental conditions that are associated with the contact changed the machine action.

25 [0040] The next example is a dialog where a user attempts to place a call while user is in a roaming mode or the contact is in a roaming mode. 1) User: "Call Paul"; Dialog System: "Calling Paul will cost Paul two dollars a minute since he is currently roaming. Would you like to message Paul instead?" In this example, the environmental conditions that changed the machine action are associated with the current activity of the contact.

30 [0041] The next example is a dialog where a user attempts to call a contact while the contact is likely driving. 1) User: "Call Paul"; 2) Dialog System: "Paul may be driving. Would you like to send a message instead?" In this example, the environmental conditions that changed the machine action are associated with the current activity of the contact.

[0042] The next example is a dialog where a user attempts to call a contact while the contact is in a noisy environment. 1) User: "Call Paul"; 2) Dialog System "Paul might not

be able to hear you well. Would you like to send a message instead?" In this example, the environmental conditions that changed the machine action are associated with the current noise conditions of the contact.

[0043] Flowing to operation 440, a response for the machine action is determined using 5 the environmental conditions. Many different types of responses may be provided (e.g., visual, auditory, haptic ...). A spoken response may be provided to the user or a visual response may be provided to the user. For example, a user may be in a noisy environment, and the response is adjusted to be louder than normal. The type of response may also be changed. For example, the response may be changed from an auditory response to a visual 10 response.

[0044] Moving to operation 450, the response is provided to the user.

[0045] The process then flows to an end operation and returns to processing other actions.

[0046] FIGURE 5 illustrates an exemplary online system that uses environmental 15 conditions in determining a machine action and how to deliver a response in a dialog. As illustrated, system 1000 includes service 1010, data store 1045, environmental conditions 150, touch screen input device 1050 (e.g., a slate), smart phone 1030 and display device 1080.

[0047] As illustrated, service 1010 is a cloud based and/or enterprise based service that 20 may be configured to provide services, such as dialog services as described herein. The service may be interacted with using different types of input and output. For example, a user may use speech input, touch input, hardware based input, and the like. Functionality of one or more of the services provided by service 1010 may also be configured as a client/server based application.

[0048] As illustrated, service 1010 is a multi-tenant service that provides resources 1015 and services to any number of tenants (e.g., Tenants 1-N). Multi-tenant service 1010 is a cloud based service that provides resources 1015 to tenants subscribed to the service and maintains each tenant's data separately and protected from other tenant data.

[0049] System 1000 as illustrated comprises a touch screen input device 1050 (e.g., a 30 slate device) and smart phone 1030 that detects when a touch input has been received (e.g., a finger touching or nearly touching the touch screen). Any type of touch screen may be utilized that detects a user's touch input. For example, the touch screen may include one or more layers of capacitive material that detects the touch input. Other sensors may be used in addition to or in place of the capacitive material. For example, Infrared (IR)

sensors may be used. According to an embodiment, the touch screen is configured to detect objects that in contact with or above a touchable surface. Although the term "above" is used in this description, it should be understood that the orientation of the touch panel system is irrelevant. The term "above" is intended to be applicable to all such orientations.

5 The touch screen may be configured to determine locations of where touch input is received (e.g., a starting point, intermediate points and an ending point). Actual contact between the touchable surface and the object may be detected by any suitable means, including, for example, by a vibration sensor or microphone coupled to the touch panel. A non-exhaustive list of examples for sensors to detect contact includes pressure-based
10 mechanisms, micro-machined accelerometers, piezoelectric devices, capacitive sensors, resistive sensors, inductive sensors, laser vibrometers, and LED vibrometers.

[0050] According to an embodiment, smart phone 1030, touch screen input device 1050, and device 1080 are configured with multimodal input and output and each include an application (1031, 1051, 1081) that interact with service 1010.

15 [0051] As illustrated, touch screen input device 1050, smart phone 1030, and display device 1080 shows exemplary displays 1052, 1032, and 1082 showing the use of an application. Data may be stored on a device (e.g., smart phone 1030, touch screen input device 1050 and/or at some other location (e.g., network data store 1045). Data store 1045, or some other store, may be used to store an understanding model, as well as other data.
20 The applications used by the devices may be client based applications, server based applications, cloud based applications or some combination. According to an embodiment, display device 1080 is a device such as a MICROSOFT XBOXTM coupled to a display.

[0052] Dialog manager 26 is configured to perform operations relating to processes as described herein. Dialog manager 26 is configured to access environmental conditions
25 150. While manager 26 is shown within service 1010, the functionality of the manager may be included in other locations (e.g., on smart phone 1030 and/or touch screen input device 1050 and/or device 1080).

[0053] The embodiments and functionalities described herein may operate via a multitude of computing systems including, without limitation, desktop computer systems, 30 wired and wireless computing systems, mobile computing systems (e.g., mobile telephones, netbooks, tablet or slate type computers, notebook computers, and laptop computers), hand-held devices, multiprocessor systems, microprocessor-based or programmable consumer electronics, minicomputers, and mainframe computers.

[0054] In addition, the embodiments and functionalities described herein may operate over distributed systems (e.g., cloud-based computing systems), where application functionality, memory, data storage and retrieval and various processing functions may be operated remotely from each other over a distributed computing network, such as the

5 Internet or an intranet. User interfaces and information of various types may be displayed via on-board computing device displays or via remote display units associated with one or more computing devices. For example user interfaces and information of various types may be displayed and interacted with on a wall surface onto which user interfaces and information of various types are projected. Interaction with the multitude of computing

10 systems with which embodiments of the invention may be practiced include, keystroke entry, touch screen entry, voice or other audio entry, gesture entry where an associated computing device is equipped with detection (e.g., camera) functionality for capturing and interpreting user gestures for controlling the functionality of the computing device, and the like.

15 [0055] FIGURES 6-8 and the associated descriptions provide a discussion of a variety of operating environments in which embodiments of the invention may be practiced. However, the devices and systems illustrated and discussed with respect to FIGURES 6-8 are for purposes of example and illustration and are not limiting of a vast number of computing device configurations that may be utilized for practicing embodiments of the

20 invention, described herein.

[0056] FIGURE 6 is a block diagram illustrating physical components (i.e., hardware) of a computing device 1100 with which embodiments of the invention may be practiced. The computing device components described below may be suitable for the computing devices described above. In a basic configuration, the computing device 1100 may include at least one processing unit 1102 and a system memory 1104. Depending on the configuration and type of computing device, the system memory 1104 may comprise, but is not limited to, volatile storage (e.g., random access memory), non-volatile storage (e.g., read-only memory), flash memory, or any combination of such memories. The system memory 1104 may include an operating system 1105 and one or more program modules 1106 suitable

25 for running software applications 1120 such as the dialog manager 26. The operating system 1105, for example, may be suitable for controlling the operation of the computing device 1100. Furthermore, embodiments of the invention may be practiced in conjunction with a graphics library, other operating systems, or any other application program and is not limited to any particular application or system. This basic configuration is illustrated in

FIGURE 6 by those components within a dashed line 1108. The computing device 1100 may have additional features or functionality. For example, the computing device 1100 may also include additional data storage devices (removable and/or non-removable) such as, for example, magnetic disks, optical disks, or tape. Such additional storage is

5 illustrated in FIGURE 6 by a removable storage device 1109 and a non-removable storage device 1110.

[0057] As stated above, a number of program modules and data files may be stored in the system memory 1104. While executing on the processing unit 1102, the program modules 1106 (e.g., the dialog manager 26) may perform processes including, but not

10 limited to, one or more of the stages of the methods and processes illustrated in the figures. Other program modules that may be used in accordance with embodiments of the present invention may include electronic mail and contacts applications, word processing applications, spreadsheet applications, database applications, slide presentation applications, drawing or computer-aided application programs, etc.

15 [0058] Furthermore, embodiments of the invention may be practiced in an electrical circuit comprising discrete electronic elements, packaged or integrated electronic chips containing logic gates, a circuit utilizing a microprocessor, or on a single chip containing electronic elements or microprocessors. For example, embodiments of the invention may be practiced via a system-on-a-chip (SOC) where each or many of the components

20 illustrated in FIGURE 6 may be integrated onto a single integrated circuit. Such an SOC device may include one or more processing units, graphics units, communications units, system virtualization units and various application functionality all of which are integrated (or “burned”) onto the chip substrate as a single integrated circuit. When operating via an SOC, the functionality, described herein, with respect to the dialog manager 26 may be

25 operated via application-specific logic integrated with other components of the computing device 1100 on the single integrated circuit (chip). Embodiments of the invention may also be practiced using other technologies capable of performing logical operations such as, for example, AND, OR, and NOT, including but not limited to mechanical, optical, fluidic, and quantum technologies. In addition, embodiments of the invention may be practiced

30 within a general purpose computer or in any other circuits or systems.

[0059] The computing device 1100 may also have one or more input device(s) 1112 such as a keyboard, a mouse, a pen, a sound input device, a touch input device, etc. The output device(s) 1114 such as a display, speakers, a printer, etc. may also be included. The aforementioned devices are examples and others may be used. The computing device 1100

may include one or more communication connections 1116 allowing communications with other computing devices 1118. Examples of suitable communication connections 1116 include, but are not limited to, RF transmitter, receiver, and/or transceiver circuitry; universal serial bus (USB), parallel, and/or serial ports.

5 [0060] The term computer readable media as used herein may include computer storage media. Computer storage media may include volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information, such as computer readable instructions, data structures, or program modules. The system memory 1104, the removable storage device 1109, and the non-removable storage device 10 1110 are all computer storage media examples (i.e., memory storage). Computer storage media may include RAM, ROM, electrically erasable read-only memory (EEPROM), flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other article of manufacture which can be used to store 15 information and which can be accessed by the computing device 1100. Any such computer storage media may be part of the computing device 1100. Computer storage media does not include a carrier wave or other propagated or modulated data signal.

20 [0061] Communication media may be embodied by computer readable instructions, data structures, program modules, or other data in a modulated data signal, such as a carrier wave or other transport mechanism, and includes any information delivery media. The term "modulated data signal" may describe a signal that has one or more characteristics set or changed in such a manner as to encode information in the signal. By way of example, and not limitation, communication media may include wired media such as a wired network or direct-wired connection, and wireless media such as acoustic, radio frequency 25 (RF), infrared, and other wireless media.

[0062] FIGURES 7A and 7B illustrate a mobile computing device 1200, for example, a mobile telephone, a smart phone, a tablet personal computer, a laptop computer, and the like, with which embodiments of the invention may be practiced. With reference to FIGURE 7A, one embodiment of a mobile computing device 1200 for implementing the 30 embodiments is illustrated. In a basic configuration, the mobile computing device 1200 is a handheld computer having both input elements and output elements. The mobile computing device 1200 typically includes a display 1205 and one or more input buttons 1210 that allow the user to enter information into the mobile computing device 1200. The display 1205 of the mobile computing device 1200 may also function as an input device

(e.g., a touch screen display). If included, an optional side input element 1215 allows further user input. The side input element 1215 may be a rotary switch, a button, or any other type of manual input element. In alternative embodiments, mobile computing device 1200 may incorporate more or less input elements. For example, the display 1205 may not be a touch screen in some embodiments. In yet another alternative embodiment, the mobile computing device 1200 is a portable phone system, such as a cellular phone. The mobile computing device 1200 may also include an optional keypad 1235. Optional keypad 1235 may be a physical keypad or a “soft” keypad generated on the touch screen display. In various embodiments, the output elements include the display 1205 for showing a graphical user interface (GUI), a visual indicator 1220 (e.g., a light emitting diode), and/or an audio transducer 1225 (e.g., a speaker). In some embodiments, the mobile computing device 1200 incorporates a vibration transducer for providing the user with tactile feedback. In yet another embodiment, the mobile computing device 1200 incorporates input and/or output ports, such as an audio input (e.g., a microphone jack), an audio output (e.g., a headphone jack), and a video output (e.g., a HDMI port) for sending signals to or receiving signals from an external device.

[0063] FIGURE 7B is a block diagram illustrating the architecture of one embodiment of a mobile computing device. That is, the mobile computing device 1200 can incorporate a system 1202 (i.e., an architecture) to implement some embodiments. In one embodiment, the system 1202 is implemented as a “smart phone” capable of running one or more applications (e.g., browser, e-mail, calendaring, contact managers, messaging clients, games, and media clients, players). In some embodiments, the system 1202 is integrated as a computing device, such as an integrated personal digital assistant (PDA) and wireless phone.

[0064] One or more application programs 1266 may be loaded into the memory 1262 and run on or in association with the operating system 1264. Examples of the application programs include phone dialer programs, e-mail programs, personal information management (PIM) programs, word processing programs, spreadsheet programs, Internet browser programs, messaging programs, and so forth. The system 1202 also includes a non-volatile storage area 1268 within the memory 1262. The non-volatile storage area 1268 may be used to store persistent information that should not be lost if the system 1202 is powered down. The application programs 1266 may use and store information in the non-volatile storage area 1268, such as e-mail or other messages used by an e-mail application, and the like. A synchronization application (not shown) also resides on the

system 1202 and is programmed to interact with a corresponding synchronization application resident on a host computer to keep the information stored in the non-volatile storage area 1268 synchronized with corresponding information stored at the host computer. As should be appreciated, other applications may be loaded into the memory

5 1262 and run on the mobile computing device 1200, including the dialog manager 26 as described herein.

[0065] The system 1202 has a power supply 1270, which may be implemented as one or more batteries. The power supply 1270 might further include an external power source, such as an AC adapter or a powered docking cradle that supplements or recharges the

10 batteries.

[0066] The system 1202 may also include a radio 1272 that performs the function of transmitting and receiving radio frequency communications. The radio 1272 facilitates wireless connectivity between the system 1202 and the “outside world”, via a

communications carrier or service provider. Transmissions to and from the radio 1272 are conducted under control of the operating system 1264. In other words, communications received by the radio 1272 may be disseminated to the application programs 1266 via the operating system 1264, and vice versa.

[0067] The visual indicator 1220 may be used to provide visual notifications, and/or an audio interface 1274 may be used for producing audible notifications via the audio

20 transducer 1225. In the illustrated embodiment, the visual indicator 1220 is a light emitting diode (LED) and the audio transducer 1225 is a speaker. These devices may be directly coupled to the power supply 1270 so that when activated, they remain on for a duration dictated by the notification mechanism even though the processor 1260 and other

25 components might shut down for conserving battery power. The LED may be programmed to remain on indefinitely until the user takes action to indicate the powered-on status of the device. The audio interface 1274 is used to provide audible signals to and receive audible signals from the user. For example, in addition to being coupled to the audio transducer 1225, the audio interface 1274 may also be coupled to a microphone to receive audible

input, such as to facilitate a telephone conversation. In accordance with embodiments of 30 the present invention, the microphone may also serve as an audio sensor to facilitate control of notifications, as will be described below. The system 1202 may further include a video interface 1276 that enables an operation of an on-board camera to record still images, video stream, and the like.

[0068] A mobile computing device 1200 implementing the system 1202 may have additional features or functionality. For example, the mobile computing device 1200 may also include additional data storage devices (removable and/or non-removable) such as, magnetic disks, optical disks, or tape. Such additional storage is illustrated in FIGURE 7B

5 by the non-volatile storage area 1268. Mobile computing device 1200 may also include peripheral device port 1230.

[0069] Data/information generated or captured by the mobile computing device 1200 and stored via the system 1202 may be stored locally on the mobile computing device 1200, as described above, or the data may be stored on any number of storage media that 10 may be accessed by the device via the radio 1272 or via a wired connection between the mobile computing device 1200 and a separate computing device associated with the mobile computing device 1200, for example, a server computer in a distributed computing network, such as the Internet. As should be appreciated such data/information may be accessed via the mobile computing device 1200 via the radio 1272 or via a distributed 15 computing network. Similarly, such data/information may be readily transferred between computing devices for storage and use according to well-known data/information transfer and storage means, including electronic mail and collaborative data/information sharing systems.

[0070] FIGURE 8 illustrates an embodiment of an architecture of an exemplary system, 20 as described above. Content developed, interacted with, or edited in association with the dialog manager 26 may be stored in different communication channels or other storage types. For example, various documents may be stored using a directory service 1322, a web portal 1324, a mailbox service 1326, an instant messaging store 1328, or a social networking site 1330. The dialog manager 26 may use any of these types of systems or the 25 like for enabling data utilization, as described herein. A server 1320 may provide the dialog manager 26 to clients. As one example, the server 1320 may be a web server providing the dialog manager 26 over the web. The server 1320 may provide the dialog manager 26 over the web to clients through a network 1315. By way of example, the client computing device may be implemented as the computing device 1100 and embodied in a 30 personal computer, a tablet computing device 1310 and/or a mobile computing device 1200 (e.g., a smart phone). Any of these embodiments of the client computing device 1100, 1310, and 1200 may obtain content from the store 1316.

[0071] Embodiments of the present invention, for example, are described above with reference to block diagrams and/or operational illustrations of methods, systems, and

computer program products according to embodiments of the invention. The functions/acts noted in the blocks may occur out of the order as shown in any flowchart. For example, two blocks shown in succession may in fact be executed substantially concurrently or the blocks may sometimes be executed in the reverse order, depending upon the functionality/acts involved.

[0072] The description and illustration of one or more embodiments provided in this application are not intended to limit or restrict the scope of the invention as claimed in any way. The embodiments, examples, and details provided in this application are considered sufficient to convey possession and enable others to make and use the best mode of claimed invention. The claimed invention should not be construed as being limited to any embodiment, example, or detail provided in this application. Regardless of whether shown and described in combination or separately, the various features (both structural and methodological) are intended to be selectively included or omitted to produce an embodiment with a particular set of features. Having been provided with the description and illustration of the present application, one skilled in the art may envision variations, modifications, and alternate embodiments falling within the broader aspects of the general inventive concept embodied in this application that do not depart from the broader scope of the claimed invention.

CLAIMS:

1. A computer-implemented method, performed by at least one processor, for using environmental conditions in addition to other dialog state information in a conversational dialog system, comprising:

5 receiving a user dialog act requesting a computer system to perform a task;

based on the user dialog act, determining a first machine action to perform the task;

accessing environmental conditions that are associated with the user dialog act;

10 based on the environmental conditions and the user dialog act, determining a second machine action other than performing the task, wherein the second machine action at least temporarily prevents completion of the first machine action; and

performing the second machine action to provide a response to a user.

2. The method of claim 1, further comprising determining a response type for
15 the second machine action based on the environmental conditions.

3. The method of claim 1, wherein accessing the environmental conditions that are associated with the input comprises accessing at least one of: a location; a noise level; a velocity; a temperature; a current activity; or a current computing device.

4. The method of claim 1, wherein the environmental conditions comprise
20 environmental conditions associated with another user.

5. The method of claim 1, wherein determining the machine actions based on the environmental conditions and the user dialog act comprises using at least one of a rules based model or a statistical model.

6. The method of claim 2, wherein determining the response type comprises using the environmental conditions to determine a type of response from: an audible response or a visual presentation of the response.

7. The method of claim 1, further comprising updating the environmental
5 conditions after determining the second machine action.

8. The method of claim 1, further comprising receiving a selection of environmental conditions that may be used for a particular user.

9. The method of claim 1, further comprising updating a dialog policy to include rules relating to different environmental conditions.

10 10. The method of claim 1, wherein the response comprises a prompt to the user, wherein the prompt requests an additional response from the user based on the environmental conditions and the requested task.

11. A computer-readable storage device storing computer-executable instructions for using environmental conditions in addition to other dialog state information
15 in a conversational dialog system, comprising:

receiving a user dialog act requesting a computer system to perform a task;

based on the user dialog act, determining a first machine action to perform the task;

accessing environmental conditions that are associated with the user dialog
20 act;

based on the environmental conditions and the user dialog act, determining a second machine action other than performing the task, wherein the second machine action at least temporarily prevents completion of the first machine action; and

performing the second machine action to provide a response to a user.

12. The computer-readable storage device of claim 11, wherein accessing the environmental conditions that are associated with the input comprises accessing at least one of: a current activity being performed when the user dialog act is received or a current computing device.

5 13. The computer-readable storage device of claim 11, wherein the environmental conditions comprise environmental conditions that are associated with a current contact.

10 14. The computer-readable storage device of claim 11, wherein determining the machine actions comprises using a rules based model that include rules for different environmental conditions.

15. The computer-readable storage device of claim 11, further comprising instruction for using the environmental conditions to determine a type of response from: an audible response or a visual presentation of the response.

16. The computer-readable storage device of claim 11, further comprising updating the environmental conditions after determining the second machine action using the environmental conditions.

17. A system for using environmental conditions in addition to other dialog state information in a conversational dialog system, comprising:

20 a processor and memory;
an operating environment, executing using the processor; and
a dialog manager that is configured to perform actions comprising:
receiving a user dialog act requesting a computer system to perform a task;
based on the user dialog act, determining a first machine action to perform the task;

accessing environmental conditions that are associated with a user dialog act;

5 determining a second machine action, other than performing the task, using the user dialog act and the environmental conditions, wherein the second machine action at least temporarily prevents completion of the first machine action; and

performing the second machine action to provide a response to a user.

18. The system of claim 17, wherein accessing the environmental conditions comprises accessing at least one of: a current activity being performed when the user dialog act is received or a current computing device when authorized.

10 19. The system of claim 17, wherein the dialog manager is further configured to determine, based on the environmental conditions, a type of response from: an audible response or a visual presentation of the response.

15 20. The system of claim 17, wherein the dialog manager is further configured to update the environmental conditions after determining the second machine action using the environmental conditions.

21. A system comprising:

at least one processor; and

a memory storing instructions that when executed by the at least one processor perform a set of operations comprising:

20 receiving a user dialog act requesting a computer system to perform a task involving a contact;

accessing environmental conditions that are associated with the contact, wherein the environment conditions include at least one of a noise level, a time zone, or a temperature;

based on the environmental conditions and the user dialog act, determining a machine action to display a suggestion to send a text-based message to the user; and

performing the machine action to display the suggestion to the user.

22. The system of claim 21, wherein the task is calling the contact.

5 23. The system of claim 21, wherein the machine action is displaying a suggested action other than the task.

24. The system of claim 23, wherein the suggested action is messaging the contact.

25. The system of claim 21, wherein the environmental conditions that are 10 associated with the contact include a noise level.

26. The system of claim 21, wherein determining the machine action is further based on dialog state information, wherein the dialog state information includes at least one of:

15 previous turn language understanding results, speech recognition results and confidences, knowledge results, previous machine actions, session context, or client context.

27. The system of claim 21, wherein the environmental conditions indicate that the contact is in at least one of the following conditions: in a roaming service area or in a noisy environment.

28. A computer-implemented method, performed by at least one processor, for 20 using environmental conditions of a contact in a conversational dialog system, comprising:

receiving a user dialog act requesting a computer system to perform a task involving the contact;

accessing environmental conditions that are associated with the contact;

determining from the environmental conditions that the contact is in at least one of the following conditions: a roaming service area or a time zone where the time is night;

based on the determination from the environmental conditions, determining

5 a machine action; and

performing the machine action to provide a response to the user.

29. The method of claim 28, wherein the task is calling the contact.

30. The method of claim 28, wherein the machine action is displaying a suggested action other than the task.

10 31. The method of claim 30, wherein the suggested action is messaging the contact.

32. The method of claim 28, wherein the environmental conditions that are associated with the contact include at least one of: a location; a noise level; a velocity; a temperature; a current activity; or a current computing device.

15 33. The method of claim 28, wherein determining the machine action is further based on dialog state information, wherein the dialog state information includes at least one of: previous turn language understanding results, speech recognition results and confidences, knowledge results, previous machine actions, session context, or client context.

20 34. The method of claim 28, wherein the environmental conditions further indicate that the contact is in a noisy environment.

35. A storage device storing computer-executable instructions for using environmental conditions of a contact in a conversational dialog system, the instructions comprising:

receiving, from a user, a user dialog act requesting a computer system to call the contact;

accessing environmental conditions that are associated with the contact;

based on the environmental conditions and the user dialog act, determine a

5 suggested action other than calling the contact;

providing the suggested action to the user, wherein the suggested action is sending a text-based message to the contact;

receiving further input from the user; and

based on the further input, sending a text-based message to the contact.

10 36. The storage device of claim 35, wherein the environmental conditions that are associated with the contact include at least a noise level.

37. The storage device of claim 35, wherein the environmental conditions that are associated with the contact include at least one of: a location; a noise level; a velocity; a temperature; a current activity; or a current computing device.

15 38. The storage device of claim 35, wherein the environmental conditions indicate that the contact is in at least one of the following conditions: in a roaming service area, driving, or in a noisy environment.

20 39. The storage device of claim 35, wherein determining the suggested action is further based on dialog state information, wherein the dialog state information includes at least one of: previous turn language understanding results, speech recognition results and confidences, knowledge results, previous machine actions, session context, or client context.

40. The storage device of claim 35, wherein the instructions further comprise:

accessing environmental conditions that are associated with the user; and

based on the environmental conditions that are associated with the user, determining an output type for providing the suggested action to the user, wherein the output type is at least one of audible or visual.

1/9

100

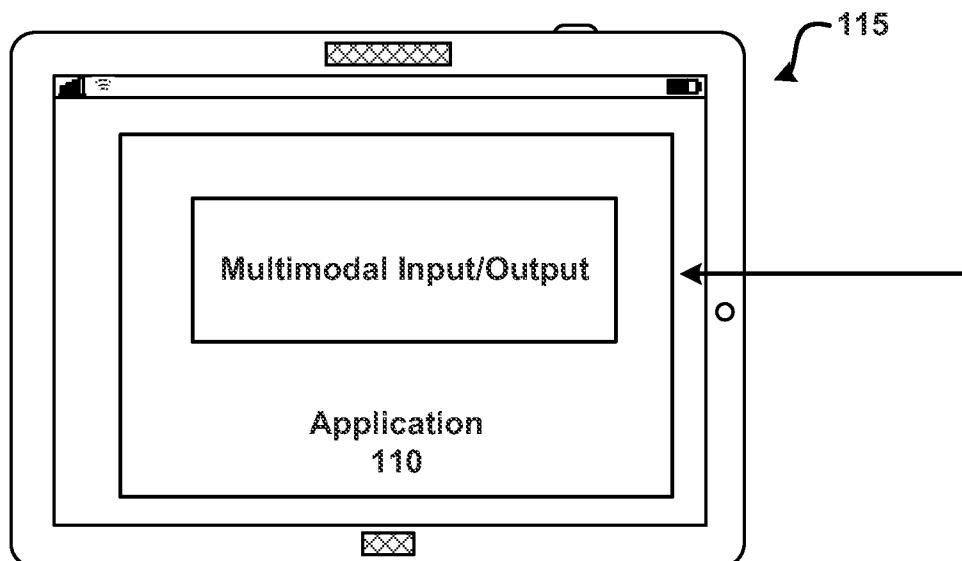
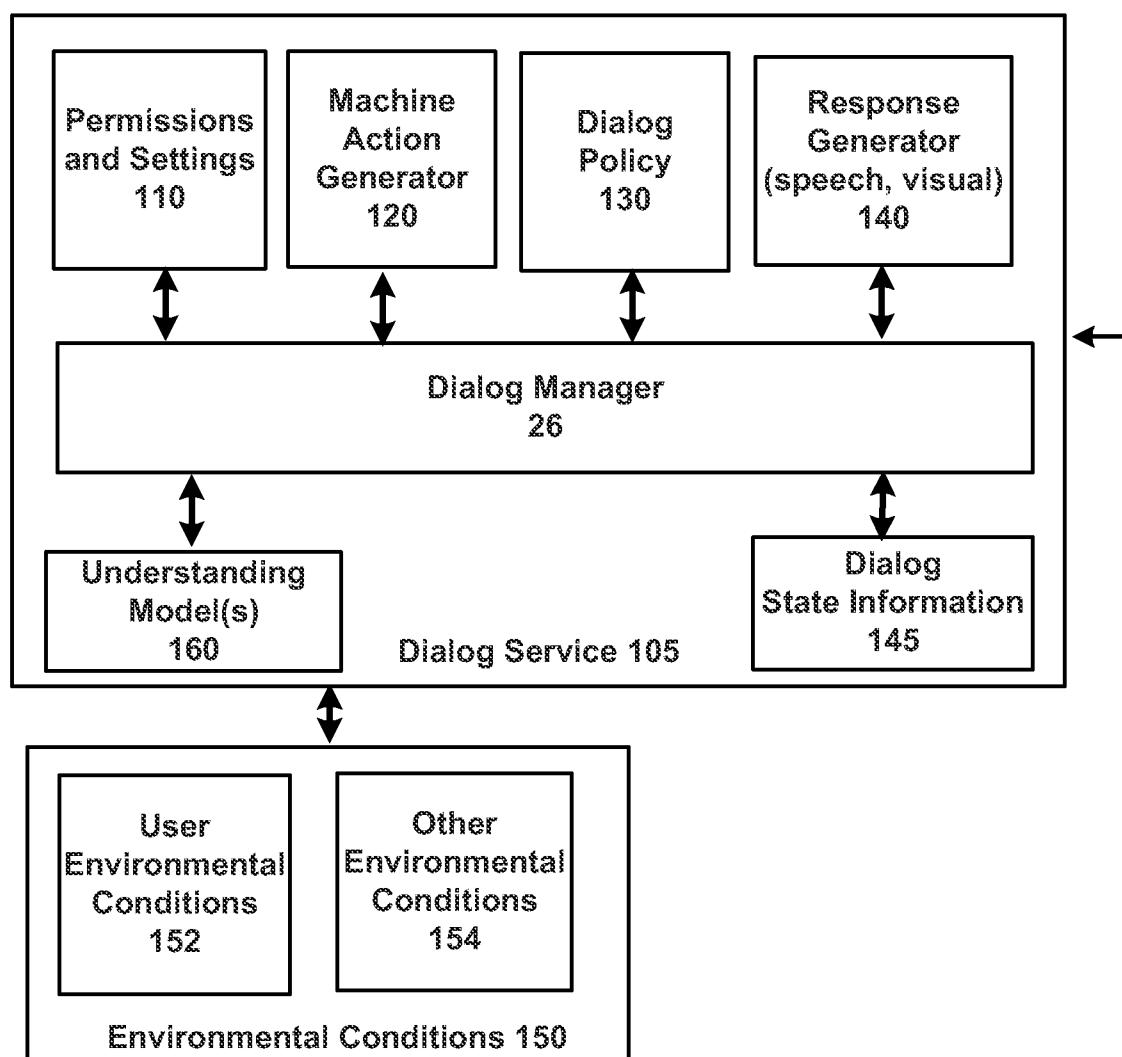


FIG.1

2/9

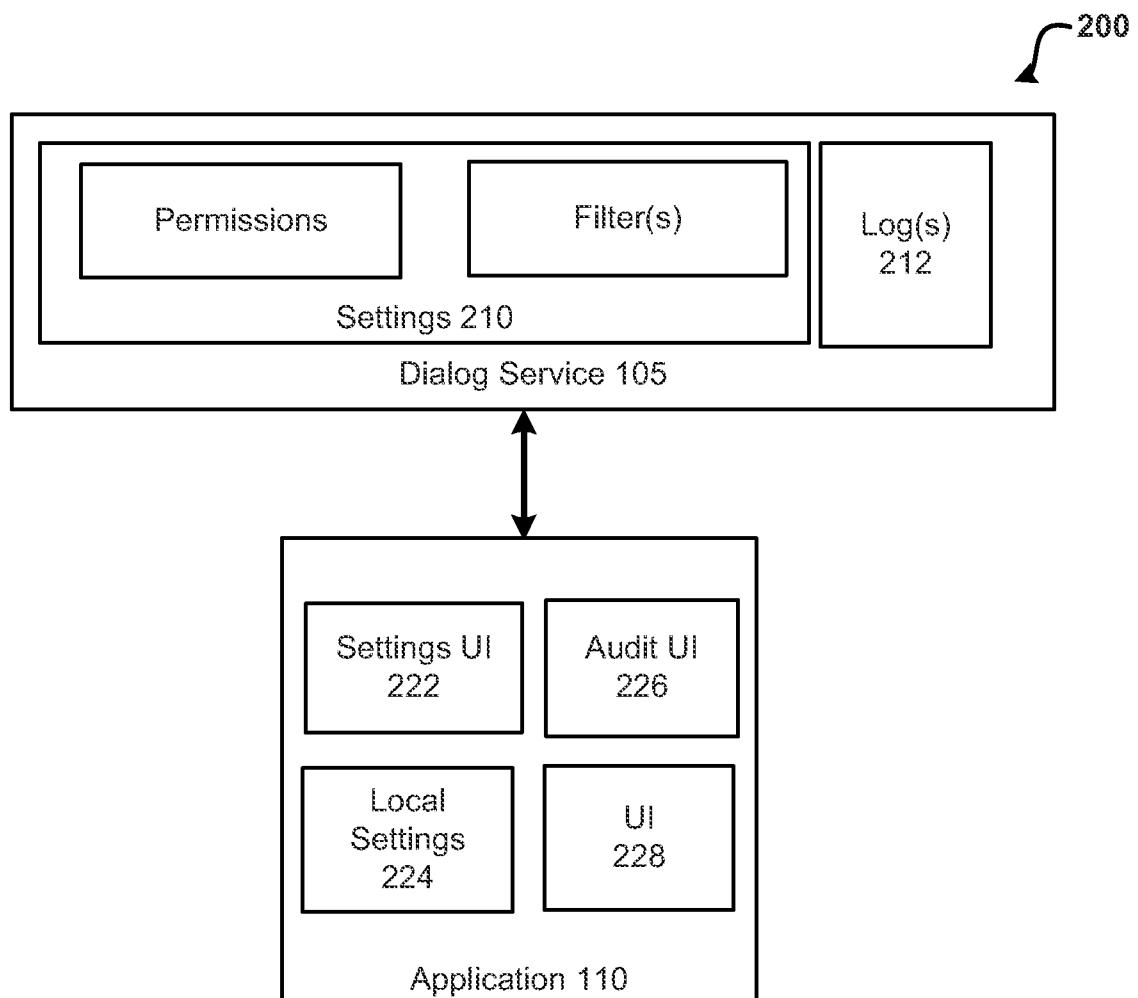


FIG.2

3/9

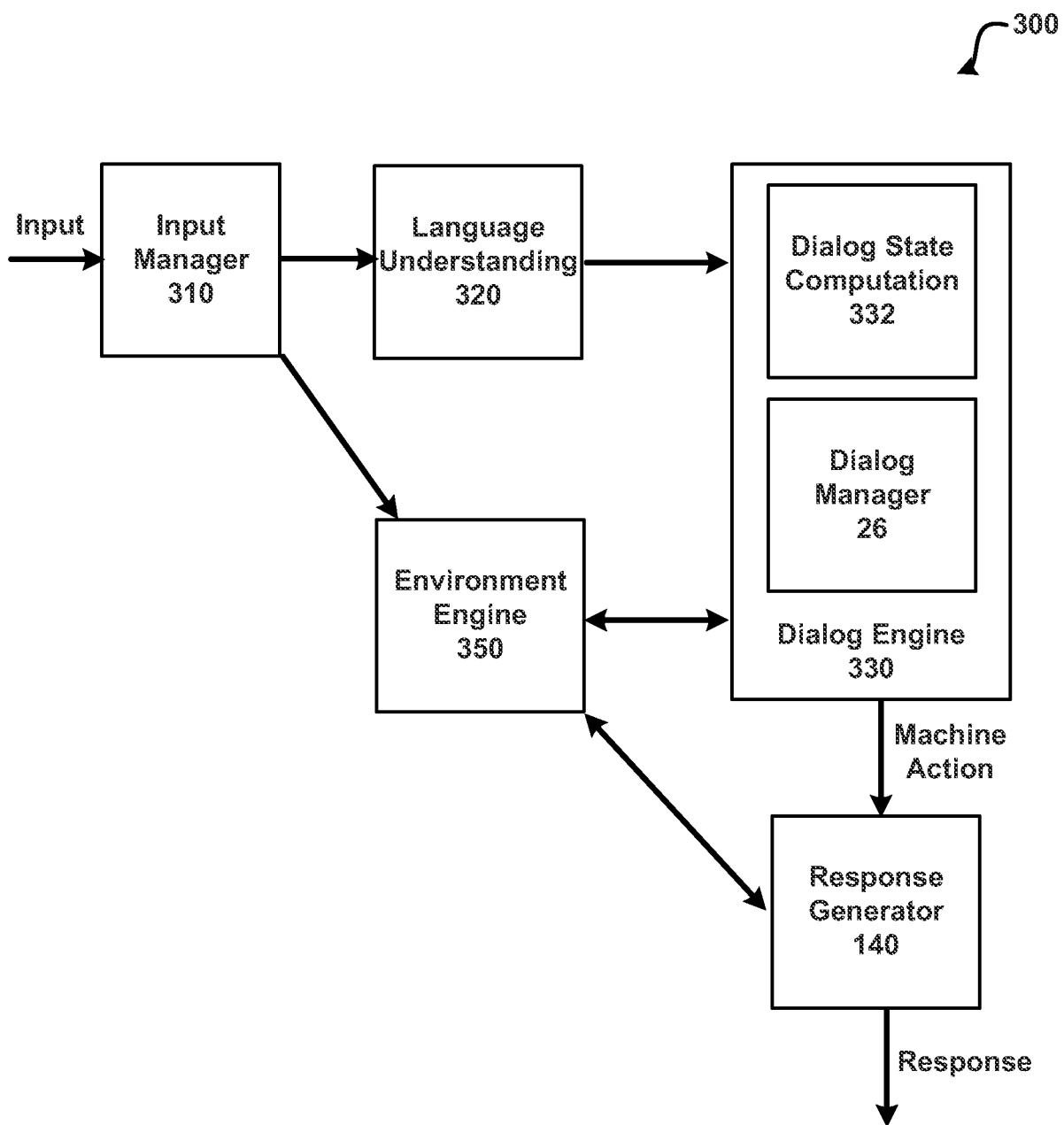


FIG. 3

4/9

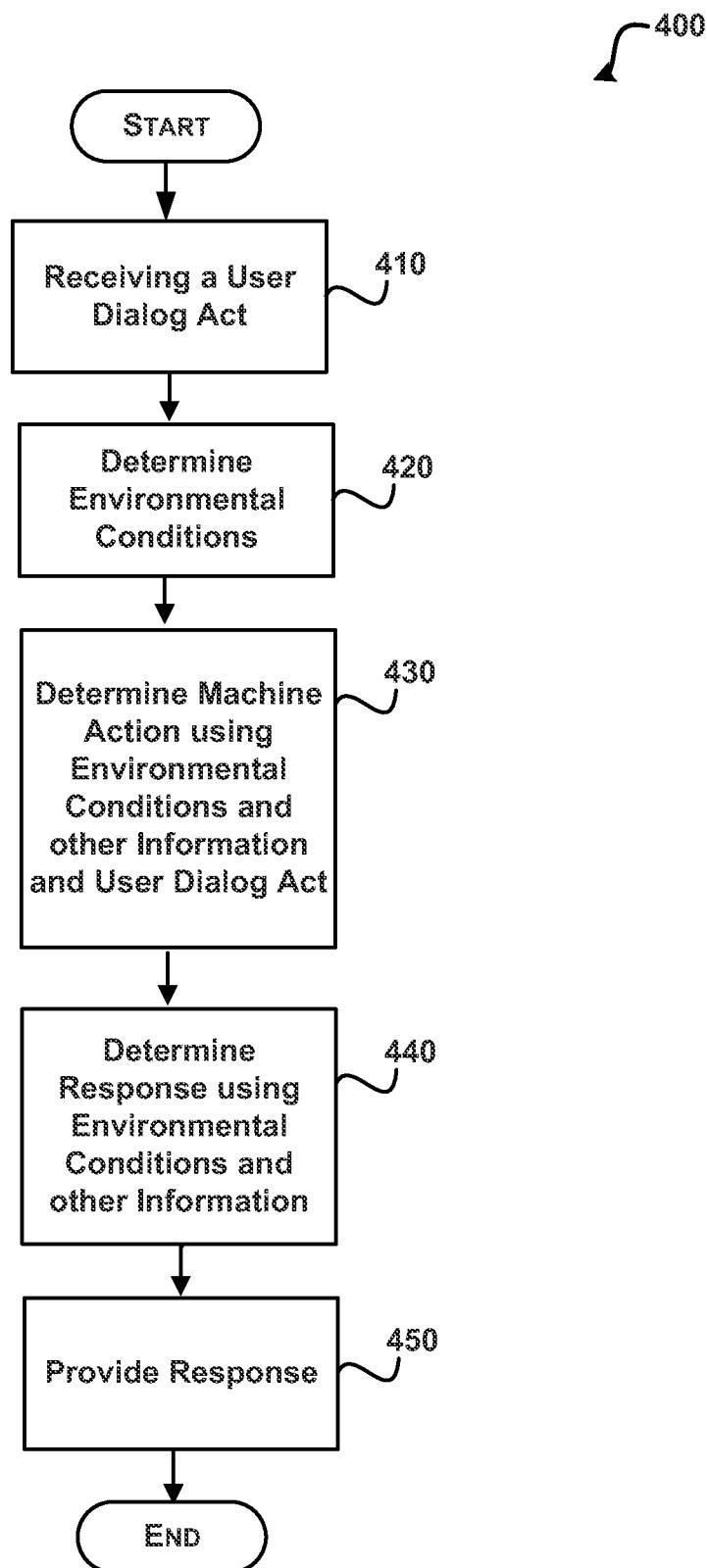


FIG. 4

5/9

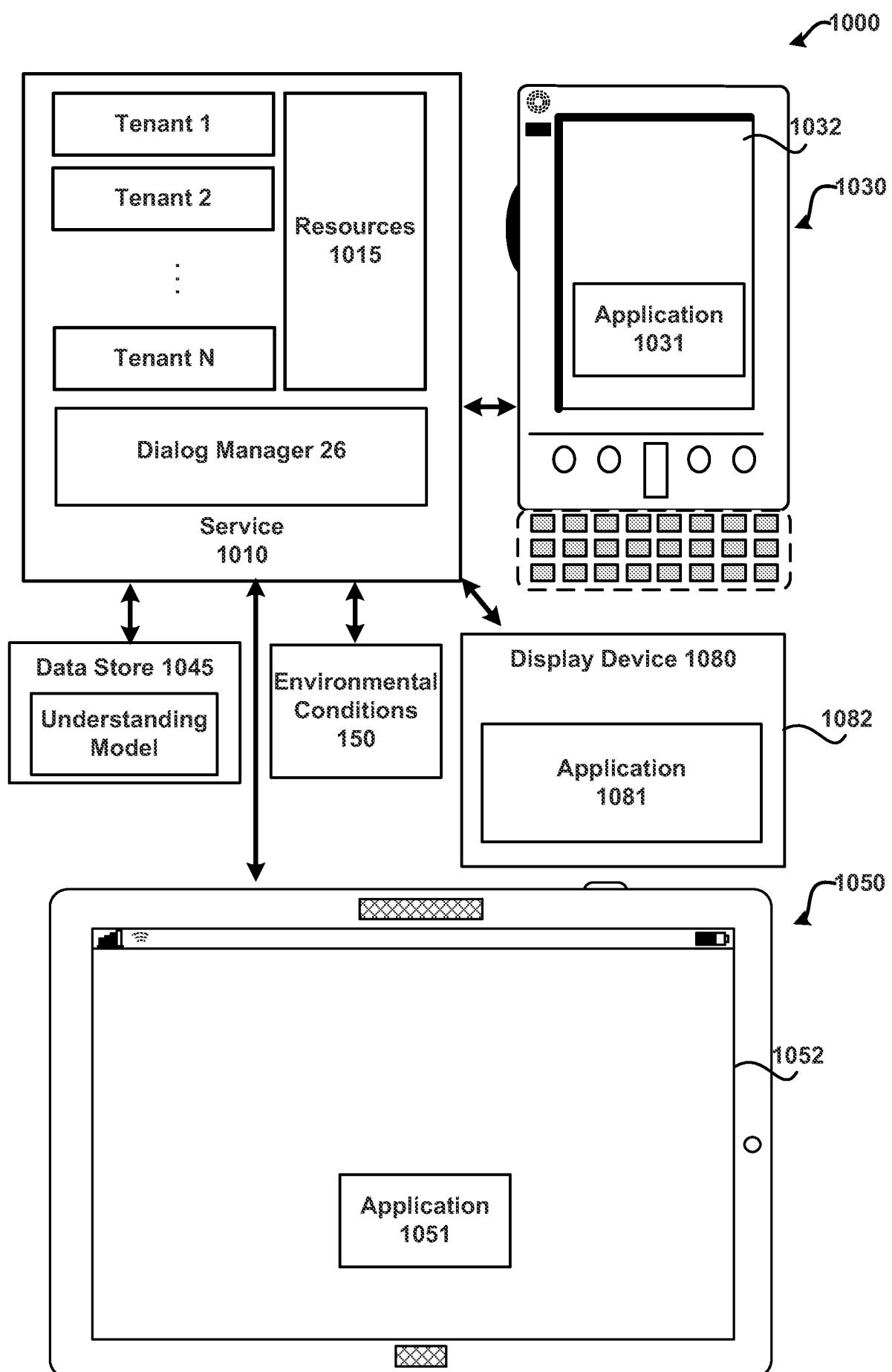


FIG. 5

6/9

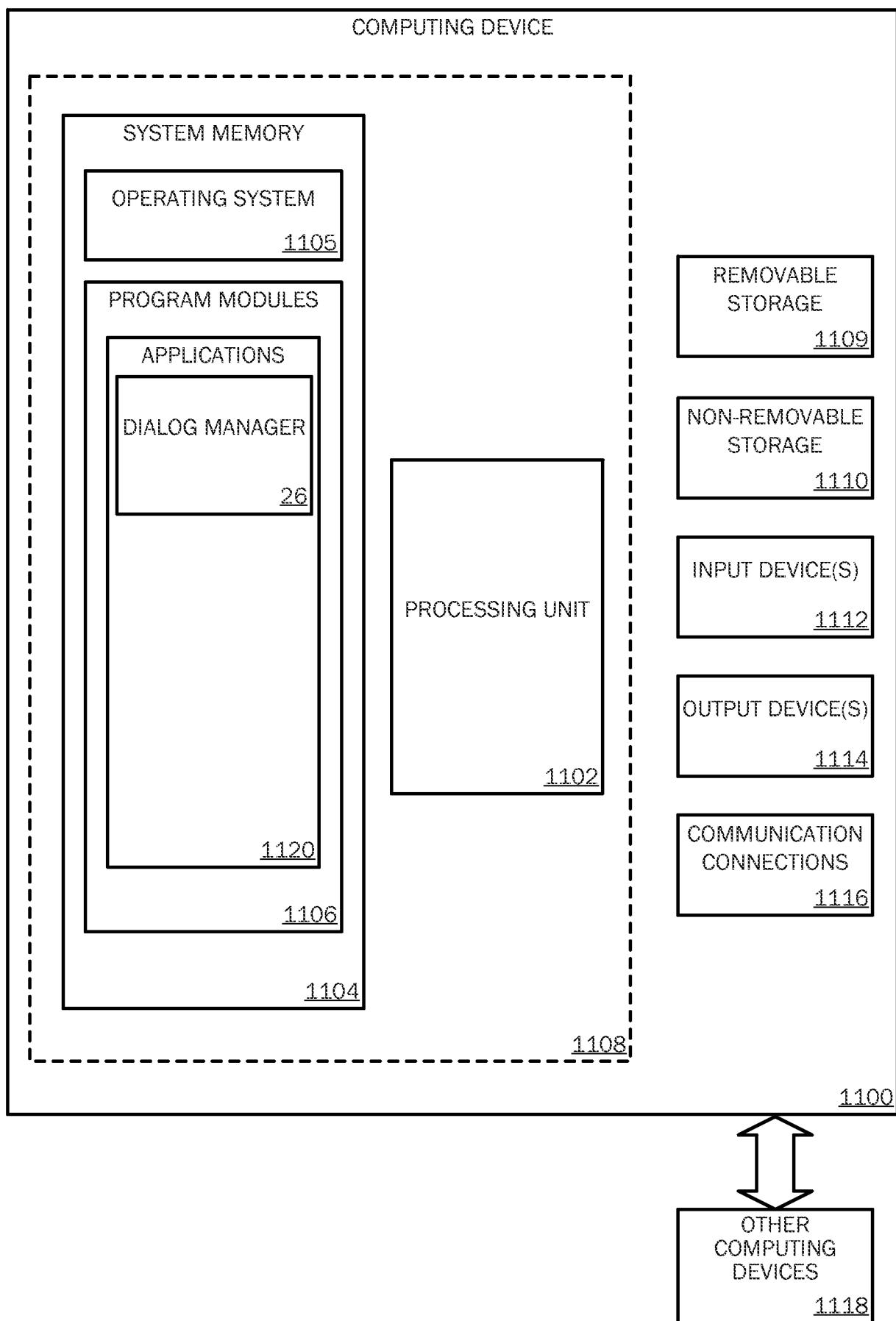
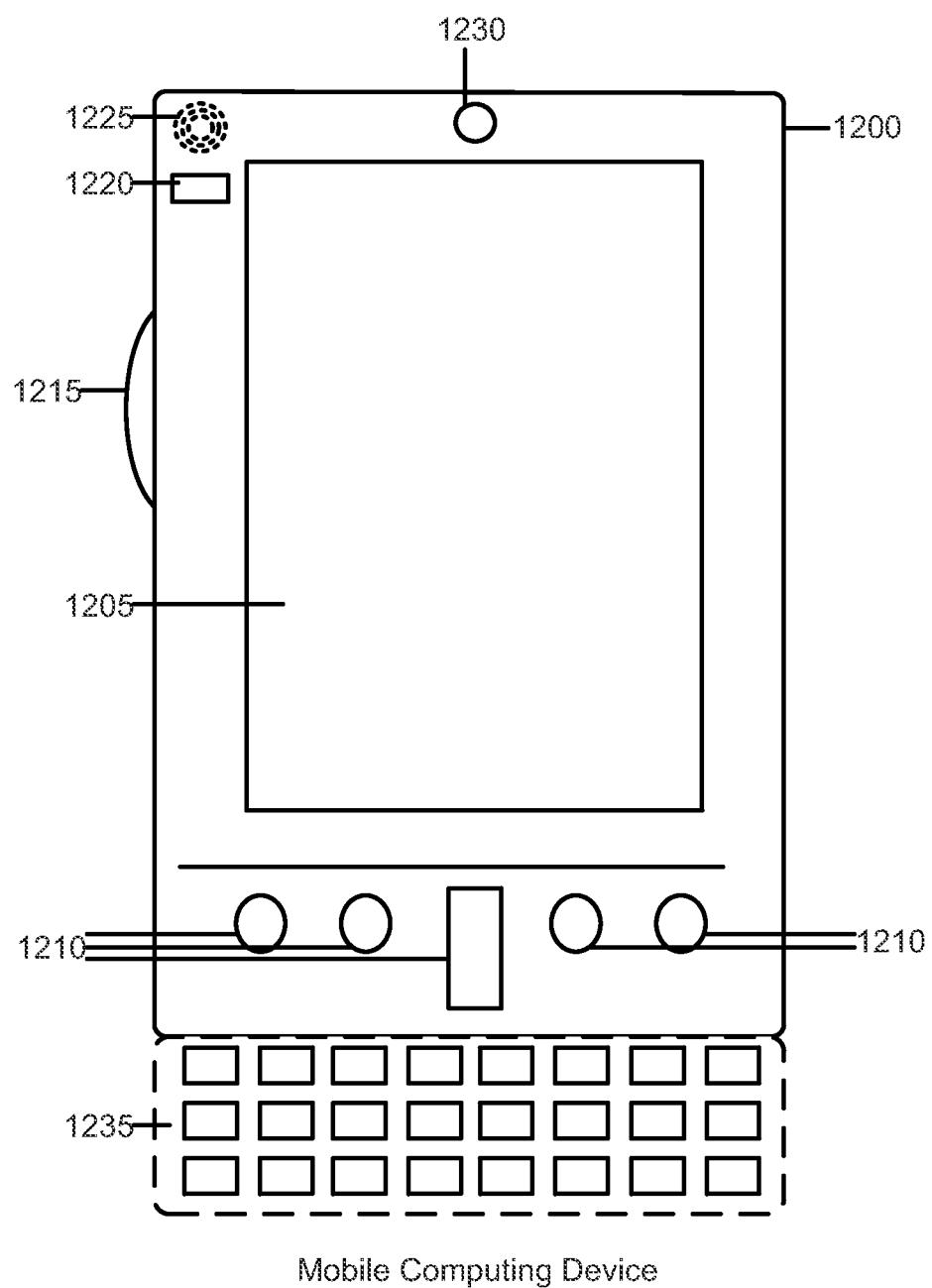
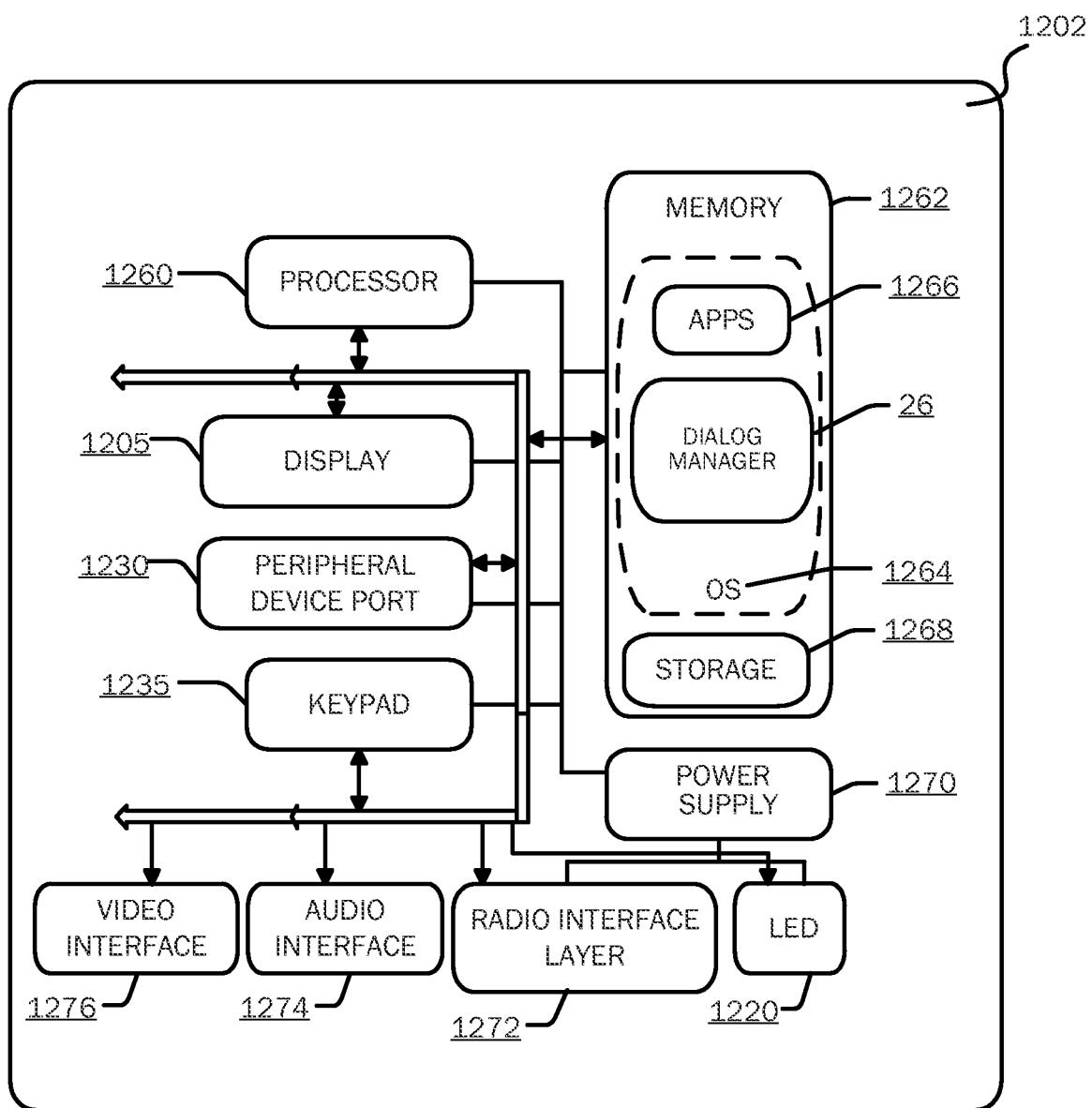


FIG. 6

7/9

**FIG. 7A**

8/9

**FIG. 7B**

9/9

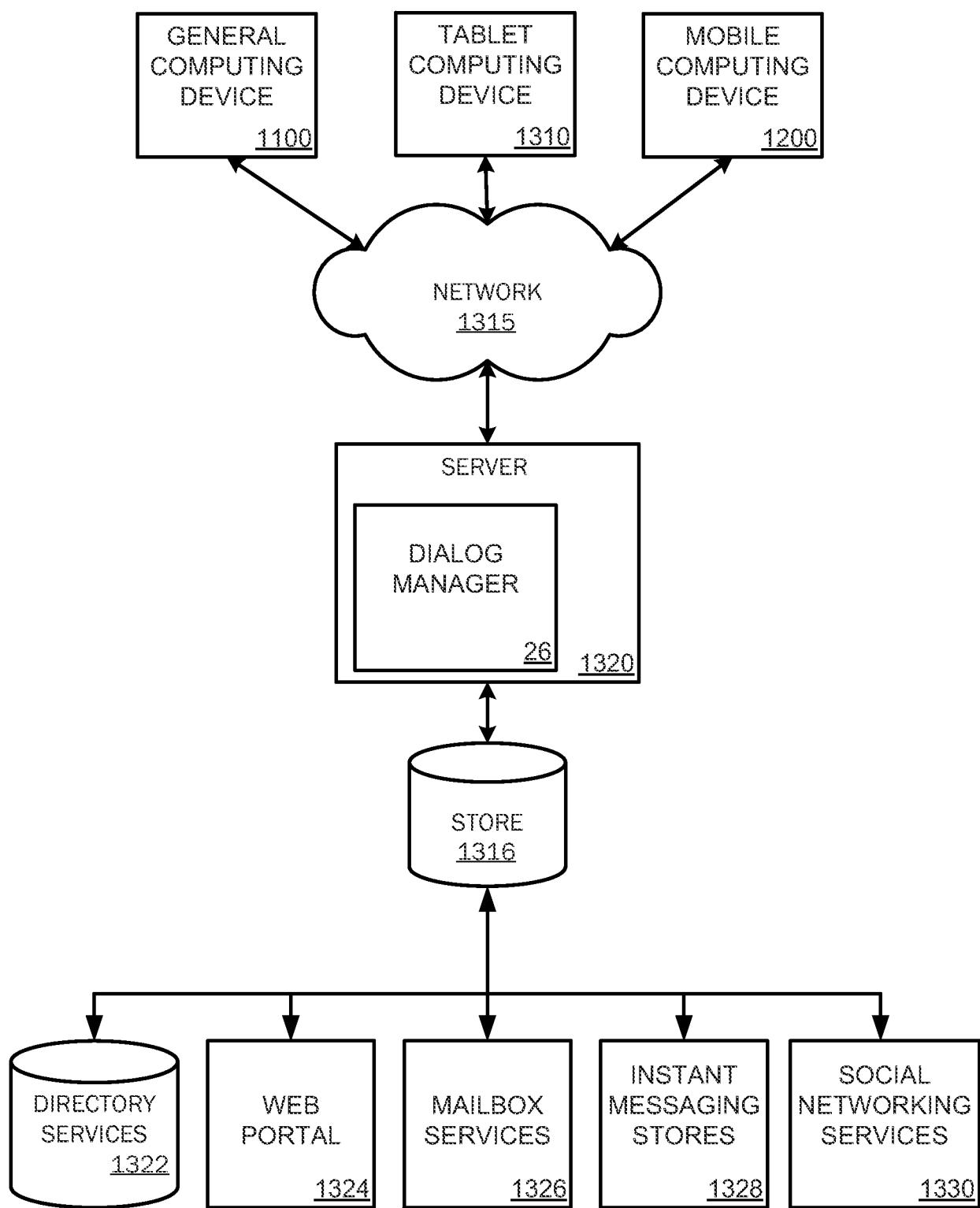


FIG. 8

