



(12) 发明专利申请

(10) 申请公布号 CN 114584660 A

(43) 申请公布日 2022. 06. 03

(21) 申请号 202210085898.1

(51) Int.Cl.

(22) 申请日 2017.06.13

H04M 3/58 (2006.01)

(30) 优先权数据

H04M 3/527 (2006.01)

62/349,396 2016.06.13 US

G10L 25/63 (2013.01)

(62) 分案原申请数据

201710444727.2 2017.06.13

(71) 申请人 谷歌有限责任公司

地址 美国加利福尼亚州

(72) 发明人 埃亚·塞加利斯

丹尼尔·瓦莱夫斯基

雅尼夫·利维坦 约西·马蒂亚斯

(74) 专利代理机构 中原信达知识产权代理有限
责任公司 11219

专利代理师 周亚荣 邓聪惠

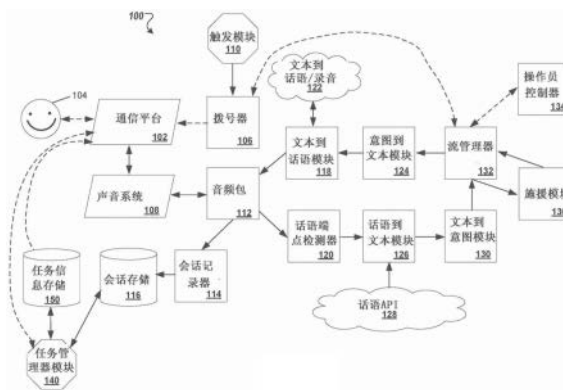
权利要求书7页 说明书28页 附图16页

(54) 发明名称

向人类操作员的升级

(57) 摘要

本公开涉及向人类操作员的升级。方法、系统和装置,包括在计算机存储介质上编码的计算机程序,涉及合成呼叫发起和施援。在一些实施方式中,一种方法,包括:由呼叫发起系统分析在电话呼叫的第一端上的第一人与在电话呼叫的第二端上的机器人程序之间的电话呼叫期间在第一人和机器人程序之间的实时对话。呼叫发起系统能够基于对实时对话的分析确定电话呼叫是否应当从机器人程序转移到电话呼叫的第二端上的第二人。响应于确定电话呼叫应当被转移到电话呼叫的第二端上的第二人,呼叫发起系统将电话呼叫从机器人程序转移到第二人。



1. 一种计算机实现的方法,包括:

由被配置为代表第一人在计算设备的机器人程序与人之间进行电话对话的所述计算设备接收来自第二人的传入电话;

由所述计算设备的所述机器人程序使得所述机器人程序接听来自所述第二人的所述传入电话;

由所述计算设备的所述机器人程序使得代表所述第一人与所述第二人进行对应的电话对话;

由所述计算设备在所述对应的电话对话期间接收来自所述第二人的非结构化交谈;

由所述计算设备生成所述非结构化交谈的转录;

基于所述非结构化交谈的所述转录,由所述计算设备确定所述第二人的意图;

基于所述第二人的所述意图,由所述计算设备生成口头响应;和

由所述计算设备的所述机器人程序使得在与所述第二人相关联的附加计算设备处可听见地渲染所述口头响应。

2. 根据权利要求1所述的方法,包括:

由所述计算设备访问与所述非结构化交谈相关联的数据,

其中,所述口头响应是基于与所述非结构化交谈相关联的所述数据。

3. 根据权利要求2所述的方法,其中,与所述非结构化交谈相关联的所述数据是所述第一人的日历数据。

4. 根据权利要求2所述的方法,其中,与所述非结构化交谈相关联的所述数据是所述第一人针对所述机器人程序执行任务的请求。

5. 根据权利要求1所述的方法,其中,生成所述口头响应包括:

基于所述人的所述意图,由所述计算设备生成所述口头响应的附加转录;和

向话语合成器提供所述口头响应的所述附加转录。

6. 根据权利要求1所述的方法,包括:

由所述计算设备接收来自所述人的附加非结构化交谈;

由所述计算设备生成所述附加非结构化交谈的附加转录;和

基于所述附加转录,由所述计算设备的所述机器人程序生成附加口头响应。

7. 根据权利要求1所述的方法,包括:

提供指示来自所述第二人的所述非结构化交谈的概要以及在所述对应的电话对话期间在与所述第二人相关联的所述附加计算设备处可听见地渲染的所述口头响应的数据以用于呈现给所述第一人。

8. 一种计算机实现的方法,包括:

由用于向组织发起电话呼叫并且在呼叫发起系统的机器人程序与所述组织的人类代表之间进行电话对话的所述呼叫发起系统的呼叫触发模块接收来自人类用户的请求,以通过向组织发起电话呼叫并且在所述呼叫发起系统的所述机器人程序与组织的人类代表之间进行电话对话来执行任务;

响应于来自所述人类用户的执行所述任务的所述请求,访问指示所述组织什么时间能够接收电话的存储数据;

基于访问指示所述组织什么时间能够接收电话的所述存储数据,确定所述组织什么时

间能够接收电话；

确定当前时间与所述组织能够接收电话是不相关联的；

至少基于确定所述当前时间与所述组织能够接收电话是不相关联的，延迟 (i) 向所述组织发起所述电话呼叫以及 (ii) 在所述呼叫发起系统的所述机器人程序与所述组织的所述人类代表之间进行所述电话对话；

在延迟 (i) 向所述组织发起所述电话呼叫以及 (ii) 在所述呼叫发起系统的所述机器人程序与所述组织的所述人类代表之间进行所述电话对话之后，确定所述当前时间与所述组织能够接收电话是相关联的；和

至少基于确定所述当前时间与所述组织能够接收电话是相关联的，向所述组织发起所述电话呼叫并且在所述呼叫发起系统的所述机器人程序与所述组织的所述人类代表之间进行所述电话对话。

9. 根据权利要求8所述的方法，其中，指示所述组织什么时间能够接收电话的所述存储数据包括指示所述组织什么时间是不想要电话呼叫的数据。

10. 根据权利要求8所述的方法，其中，指示所述组织什么时间能够接收电话的所述存储数据包括指定所述组织的开放时间的数据。

11. 根据权利要求8所述的方法，还包括：

确定所述任务是将不立即发起电话呼叫的类型。

12. 根据权利要求8所述的方法，还包括：

确定所述任务是将应用默认排程参数的类型。

13. 根据权利要求8所述的方法，还包括：

在所述电话对话已发起之前，提供 (i) 所述任务和 (ii) 指示所述组织什么时间能够接收电话的所述数据的视觉表示，以用于向所述人类用户输出。

14. 一种计算机实现的方法，包括：

由被配置为在呼叫发起系统的机器人程序和组织的人类代表之间拨打电话呼叫并且进行电话对话的所述呼叫发起系统从人类用户接收执行任务的请求；

响应于接收到执行所述任务的所述请求，由所述呼叫发起系统确定在所述呼叫发起系统的机器人程序和所述组织的人类代表之间拨打电话呼叫并且进行电话对话以执行所述任务；

响应于确定在所述呼叫发起系统的机器人程序和所述组织的人类代表之间拨打电话呼叫并且进行电话对话以执行所述任务，在所述呼叫发起系统的机器人程序和所述组织的人类代表之间拨打电话呼叫并且进行电话对话；

基于在所述电话呼叫期间在所述呼叫发起系统的机器人程序和所述组织的人类代表之间发生的交谈，由所述呼叫发起系统确定所述任务的状态已经改变；

基于确定所述任务的所述状态已经改变，基于在所述电话呼叫期间在所述呼叫发起系统的机器人程序和所述组织的人类代表之间发生的所述交谈来生成所述电话对话的概要；以及

向所述人类用户提供所述电话对话的所述概要以用于输出。

15. 根据权利要求14所述的方法，其中，所述概要包括所述对话的状态的单行文本概要。

16. 根据权利要求14所述的方法,其中,所述概要概括在所述电话呼叫期间在所述机器人程序和所述人类代表之间讨论的内容的方面。

17. 根据权利要求14所述的方法,其中,所述概要概括在所述机器人程序和所述人类代表之间的所述对话的当前状态。

18. 根据权利要求14所述的方法,其中,所述人类用户不是所述机器人程序与所述人类代表之间的所述电话呼叫的一方。

19. 根据权利要求14所述的方法,其中,所述概要包括在所述电话呼叫期间收集的信息。

20. 根据权利要求14所述的方法,其中,所述概要识别所述人类呼叫者在所述对话期间的意外评论。

21. 根据权利要求14所述的方法,还包括:

确定所述任务的状态已经改变包括确定所述任务完成,以及生成所述电话对话的所述概要基于确定所述任务完成。

22. 根据权利要求14所述的方法,还包括:

确定所述任务的状态已经改变包括确定在所述呼叫发起系统的机器人程序和所述组织的人类代表之间至少应拨打附加电话呼叫并且进行附加电话对话以完成所述任务,以及生成所述电话对话的所述概要基于确定在所述呼叫发起系统的机器人程序和所述组织的人类代表之间至少应拨打附加电话呼叫并且进行附加电话对话以完成所述任务。

23. 一种方法,包括:

由自动电话呼叫发起系统的自动电话呼叫触发模块接收指示第一事件的数据,所述自动电话呼叫发起系统被配置为在所述自动电话呼叫发起系统的机器人程序和实体的人类代表之间发起电话呼叫并且进行电话对话;

由所述自动电话呼叫触发模块使用指示所述第一事件的所述数据确定所述第一事件是特定触发事件,所述特定触发事件(i)是检测到与第一数据源相关联的字段的价值和与第二数据源相关联的所述字段的对应值的不一致,以及(ii)触发以发起电话呼叫开始的所述自动电话呼叫发起系统的工作流;

基于所述特定触发事件,从多个可能工作流中选择特定工作流,所述特定工作流与所述特定触发事件相对应;和

响应于选择所述特定工作流,通过以下来执行所述特定工作流:(i)由所述自动电话呼叫发起系统自动发起对所述特定工作流指定的实体的电话呼叫和(ii)通过在所述自动电话呼叫发起系统的机器人程序和所述特定工作流指定的所述实体的人类代表之间进行电话对话来确定所述字段的正确值。

24. 根据权利要求23所述的方法,其中,指示所述第一事件的所述数据由用户提供。

25. 根据权利要求23所述的方法,其中,所述特定触发事件包括用户请求。

26. 根据权利要求23所述的方法,其中,所述特定触发事件包括特定类型的事件,所述特定类型的事件是以下中的一个:天气事件、体育事件、娱乐事件或季节性事件。

27. 根据权利要求23所述的方法,其中,所述特定触发事件包括在提交给搜索引擎的搜索请求中检测到的趋势。

28. 根据权利要求23所述的方法,其中,所述特定触发事件包括经过预定时间段。

29. 一种方法,包括:

由自动电话呼叫发起系统接收执行任务的请求,所述自动电话呼叫发起系统(i)包括任务管理器模块并且(ii)被配置为在所述自动电话呼叫发起系统的机器人程序与所述电话呼叫的接收者之间发起电话呼叫并且进行电话对话,所述任务需要在所述自动电话呼叫发起系统的机器人程序与所述电话呼叫的接收者之间发起电话呼叫并且进行电话对话;

由所述自动电话呼叫发起系统的任务管理器模块确定已发生触发事件以提供需要在所述自动电话呼叫发起系统的机器人程序与所述电话呼叫的接收者之间发起所述电话呼叫并且进行所述电话对话的所述任务的当前状态;

响应于所述触发事件,由所述自动电话呼叫发起系统的所述任务管理器模块确定需要在所述自动电话呼叫发起系统的机器人程序与所述电话呼叫的接收者之间发起所述电话呼叫并且进行所述电话对话的所述任务的所述当前状态;

由所述自动电话呼叫发起系统生成需要在所述自动电话呼叫发起系统的机器人程序与所述电话呼叫的接收者之间发起所述电话呼叫并且进行所述电话对话的所述任务的所述当前状态的表示;以及

提供需要在所述自动电话呼叫发起系统的机器人程序与所述电话呼叫的接收者之间发起所述电话呼叫并且进行所述电话对话的所述任务的当前状态的所述表示,以用于由自动电话呼叫发起系统输出。

30. 根据权利要求29所述的方法,其中,所述触发事件是针对状态的用户请求。

31. 根据权利要求29所述的方法,其中,所述触发事件是操作员交互以在操作员已查看与执行所述任务的所述请求相关联的会话信息之后向用户提供状态。

32. 根据权利要求29所述的方法,其中,所述触发事件是状态更新事件。

33. 根据权利要求29所述的方法,其中,所述当前状态的所述表示是视觉表示。

34. 根据权利要求29所述的方法,其中,所述当前状态的所述表示是可听表示。

35. 根据权利要求29所述的方法,其中,向所述用户提供需要发起电话呼叫的所述任务的当前状态的所述表示包括:

确定用于将所述当前状态传递给所述用户的方便时间和方法。

36. 根据权利要求29所述的方法,还包括:

通知所述用户在处理执行所述任务的所述请求时将存在延迟。

37. 根据权利要求29所述的方法,其中,确定所述触发事件已发生包括:

确定所述电话呼叫的所述发起已延迟。

38. 根据权利要求29所述的方法,其中,确定所述触发事件已发生包括:

确定所述自动电话发起系统已经在所述自动电话呼叫发起系统的机器人程序与所述电话呼叫的接收者之间发起所述电话呼叫、进行所述电话对话并且完成了需要发起所述电话呼叫的所述任务。

39. 一种计算机实现的方法,包括:

由自动呼叫发起系统分析第一人与第二人或所述自动电话呼叫发起系统的机器人程序之间的电话呼叫;

基于分析所述第一人与所述第二人或所述自动电话呼叫发起系统的机器人程序之间的所述电话呼叫,由所述自动呼叫发起系统确定将所述自动呼叫发起系统的模式(i)从在

所述第一人与所述机器人程序之间进行所述电话呼叫切换到在所述第一人与所述第二人之间,或(ii)从在所述第一人与所述第二人之间进行所述电话呼叫切换到在所述第一人与所述机器人程序之间;以及

基于确定切换所述自动呼叫发起系统的所述模式,由所述自动呼叫发起系统将所述自动呼叫发起系统的所述模式(i)从在所述第一人与所述机器人程序之间进行所述电话呼叫切换到在所述第一人与所述第二人之间,或(ii)从在所述第一人与所述第二人之间进行所述电话呼叫切换到在所述第一人与所述机器人程序之间。

40. 根据权利要求39所述的方法,其中,分析所述第一人与所述第二人或所述自动电话呼叫发起系统的机器人程序之间的所述电话呼叫包括:

基于所述第一人的行为、举止、语调、恼怒程度、语言或者词语选择来确定在所述电话呼叫期间的压力。

41. 根据权利要求39所述的方法,其中,分析所述第一人与所述第二人或所述自动电话呼叫发起系统的机器人程序之间的所述电话呼叫包括:

确定所述机器人程序在所述电话呼叫期间将不会完成请求的任务的置信度。

42. 根据权利要求39所述的方法,其中,分析所述第一人与所述第二人或所述自动电话呼叫发起系统的机器人程序之间的所述电话呼叫包括:

确定所述第一人要求将所述电话呼叫转移到另一人。

43. 根据权利要求39所述的方法,其中,分析所述第一人与所述第二人或所述自动电话呼叫发起系统的机器人程序之间的所述电话呼叫包括:

确定所述第一人已嘲讽所述机器人程序或询问所述机器人程序是否是机器人。

44. 根据权利要求39所述的方法,其中,所述第二人是操作员。

45. 根据权利要求39所述的方法,还包括:

在切换所述自动呼叫发起系统的所述模式花费比预定时间量更长时间时,终止所述电话呼叫。

46. 一种计算机实现的方法,包括:

由用于拨打电话呼叫并且在所述电话呼叫期间在呼叫发起系统的机器人程序和组织的代表之间进行对话的呼叫发起系统,在电话呼叫期间接收组织的代表的言语,其中,所述人类代表的所述言语提示所述机器人程序重复自身;

由所述呼叫发起系统的意图到文本模块生成文本回复,所述文本回复(i)以不同方式重述先前输出的回复并且(ii)根据与正由所述呼叫发起系统针对人类用户执行的任务相关联的预定义工作流;

更改对所述组织的所述人类代表的口头言语的所述文本回复,以包括话语不流利;以及

提供所更改的文本回复的音频表示,以用于由所述呼叫发起系统的文本到话语模块输出。

47. 根据权利要求46所述的方法,其中,所述话语不流利包括非词汇言语。

48. 根据权利要求46所述的方法,其中,所述话语不流利包括补白词语。

49. 根据权利要求46所述的方法,其中,更改所述文本回复包括将所述话语不流利添加到所述文本回复。

50. 根据权利要求46所述的方法,其中,更改所述文本回复包括将所述文本回复的特定词项的拼写附加到所述文本回复。

51. 根据权利要求46所述的方法,其中,所述组织的人类代表的所述言语包括不完整言语,并且还包括:

更改所述文本回复以包括所述话语不流利包括向所述文本回复引入状态检查提示。

52. 一种计算机实现的方法,包括:

由用于拨打电话呼叫并且在所述电话呼叫期间在呼叫发起系统的机器人程序和组织的人类代表之间进行对话的呼叫发起系统访问表示预定义的、理想化的对话工作流的一组链接节点,每个节点都被一组一个或多个相应意图链接;

访问包括通过与来自各种组织的一组人类代表的相应人类代表对话来完成任务的相应人类操作员的音频样本的训练数据;

针对与响应于所述相应操作员的组织的所述相应代表相对应的一个或多个音频样本,通过以下来更新所述预定义的、理想化的对话工作流:将所述音频样本的表示与在所述预定义的、理想化的对话工作流中的所述意图中的一个意图相关联或使用以新方式定义的意图以新方式连接所述预定义的、理想化的对话工作流的节点;

由所述呼叫发起系统通过发起与不包括在来自所述各种组织的所述一组人类代表中的组织的人类代表发起电话呼叫来为人类用户执行所述任务;以及

使用所更新的对话工作流,由所述呼叫发起系统的所述机器人程序与不包括在来自所述各种组织的所述一组人类代表中的所述组织的所述人类代表进行电话对话。

53. 根据权利要求52所述的方法,其中,所述音频样本的所述表示包括所述音频样本的转录。

54. 根据权利要求52所述的方法,其中,所述一个或多个音频样本中的每一个是手动注释的。

55. 根据权利要求52所述的方法,包括使用机器学习训练的分类器来注释所述一个或多个音频样本中的每一个。

56. 根据权利要求52所述的方法,包括分配所述一个或多个音频样本与每个相应意图相关联的可能性。

57. 根据权利要求52所述的方法,其中,在所述呼叫发起系统的任何机器人程序向所述各种组织或不包括在所述各种组织中的其他组织的任何人类代表发起电话呼叫以执行所述任务之前,更新所述理想化的对话工作流。

58. 一种由一个或多个处理器实现的方法,所述方法包括:

接收来自人类用户的用户输入以发起使用机器人程序的任务的执行,其中,使用所述机器人程序的所述任务的执行包括:

同时向对应组织发起多个电话呼叫,以及

同时在所述机器人程序的对应实例和所述对应组织中的每一个的对应人类代表之间进行对应电话对话;

基于所述用户输入,识别在进行所述对应电话交谈时利用的对应参数的至少一个值;

基于从所述对应组织中的每一个的所述对应人类代表接收到的对应附加用户输入,生成与和所述对应组织中的每一个的所述对应参数的所述至少一个值相关的信息相关联的

对应输出;以及

在同时进行对应电话对话的同时,使得与和所述对应组织中的每一个的所述对应参数的所述至少一个值相关的所述信息相关联的所述对应输出经由所述人类用户的计算设备被提供以用于呈现给所述人类用户。

59. 根据权利要求58所述的方法,还包括:

响应于使得所述对应输出被提供以用于呈现给所述人类用户:

从所述人类用户接收对与所述对应组织中的给定组织相关联的给定输出的人类用户选择,以及

使得所述机器人程序完成与所述给定组织的所述任务的执行。

60. 根据权利要求59所述的方法,其中,使得所述机器人程序完成与所述给定组织的所述任务的执行包括:

使得所述机器人程序利用所述对应参数的所述至少一个值来促进所述机器人程序的所述对应实例与所述给定组织的所述对应人类代表之间的所述对应电话对话。

61. 根据权利要求60所述的方法,其中,所述任务包括以下中的一个:预约任务、预订任务、库存任务或运输任务。

62. 根据权利要求59所述的方法,还包括:

使得所述机器人程序终止与所述其他对应组织的对应人类代表的所述对应电话对话。

63. 根据权利要求59所述的方法,其中,经由所述人类用户的所述计算设备的图形用户界面提供被提供以用于呈现给所述人类用户的与和所述对应组织中的每一个的所述对应参数的所述至少一个值相关的所述信息相关联的所述对应输出,并且其中,从所述人类用户接收对给定输出的所述人类用户选择包括经由所述计算设备的所述图形用户界面接收所述人类用户选择。

64. 根据权利要求59所述的方法,其中,经由所述人类用户的所述计算设备的一个或多个扬声器提供被提供以用于呈现给所述人类用户的与和所述对应组织中的每一个的所述对应参数的所述至少一个值相关的所述信息相关联的所述对应输出,并且其中,从所述人类用户接收对给定输出的所述人类用户选择包括经由所述计算设备的麦克风接收所述人类用户选择。

65. 根据权利要求58所述的方法,其中,在向所述对应组织发起所述多个电话呼叫之前,基于从所述人类用户接收到的进一步的附加用户输入来识别所述对应组织。

66. 根据权利要求58所述的方法,其中,所述对应组织是基于所述人类用户的所述计算设备的当前位置来识别的。

67. 根据权利要求58所述的方法,其中,从所述人类用户接收到的用以发起所述任务的执行的所述用户输入是经由所述人类用户的所述计算设备的搜索界面接收的。

向人类操作员的升级

[0001] 分案说明

[0002] 本申请属于申请日为2017年6月13日的中国发明专利申请201710444727.2的分案申请。

[0003] 相关申请的交叉引用

[0004] 本申请要求2016年6月13日提交的美国临时专利申请No.62/349,396的权益,其通过引用被整体并入本文。

技术领域

[0005] 本说明书涉及自然语言处理。

背景技术

[0006] 用户可能需要收集多种类型的信息,在不与人互动的情况下这些信息不易被获取。例如,为了验证或收集来自多个营业场所或组织的数据,用户可能需要呼叫每个商家或组织来收集信息。虽然web搜索引擎可以通过提供服务或商家的联系信息来帮助用户执行这样的任务,但用户仍然必须自己来呼叫服务或商家来自己完成任务。

[0007] 为了维护从多个营业场所或组织收集到的信息的数据库,人类操作员可以发起对大量商家的自动呼叫以收集数据,但是选择被呼叫者(例如,提供相同美食的特定城镇的所有餐馆)并且手动进行拨打呼叫可能是耗时的。此外,确定何时以及是否拨打呼叫通常需要对现有数据进行人类分析,以便识别对验证、更新或补充信息的需求。

[0008] 用户还可能希望执行诸如预约或雇用服务等任务。然而,一般来说,用户必须与一个人交互以完成所需任务。例如,为了对没有网站的小餐馆进行预订,用户可能需要打电话并与女店主交谈。在某些情况下,即使用户自己拨打电话,他们也可能会遇到自动电话树,而自动电话树通常只接受有限的用户响应。

发明内容

[0009] 系统可以通过从系统接收的数据确定是否对特定号码发起呼叫来帮助用户完成各种任务,这些任务涉及通过电话呼叫与人通信或与通过电话(例如IVR)操作的自动系统通信。一旦拨打呼叫,系统可以获得信息,向第三方提供信息,执行动作,例如代表用户等等。在某些示例中,系统代表用户参与和人的交谈。该交谈可以经由系统与人之间的电话连接发生。在某些示例中,该系统可以包括遵循工作流程来操作搜索引擎的一部分,或者形成搜索引擎的一部分,该工作流程与提交涉及要完成的任务的查询的搜索引擎用户的意图相关联。系统可以通过至少一个自主或半自主的软件代理(“机器人程序”)操作来执行用户的任务。

[0010] 在一个一般方面,一种方法包括:通过呼叫发起系统的呼叫触发模块接收指示第一事件的数据,该呼叫发起系统用于拨打呼叫并且在呼叫期间在呼叫发起系统的机器人程序和被呼叫人之间进行对话;通过所述呼叫触发模块和使用指示所述第一事件的数据来确

定所述第一事件是多个可能的触发事件中的特定触发事件,所述特定触发事件触发开始于发起电话呼叫的所述呼叫发起系统的工作流;基于所确定的触发事件,从多个可能的工作流中选择特定工作流,所述特定工作流与所确定的触发事件相对应;并且响应于所述选择, i) 发起对由所述特定工作流指定的被呼叫者的电话呼叫,以及 ii) 将所述工作流作为所述机器人程序和被呼叫者之间的双向对话来执行。

[0011] 实施方式可以包括以下特征中的一个或多个。例如,所确定的触发事件是与第一数据源相关联的值和与第二数据源相关联的对应值的不一致性。指示第一事件的数据可以由用户提供。所确定的触发事件可以是用户请求。所确定的触发事件可以是以下之一的特定类型的事件:天气事件、娱乐事件或季节性事件。所确定的触发事件可以是在提交给搜索引擎的搜索请求中检测到的趋势。所确定的触发事件可以是经过了预定时间段。

[0012] 在另一个一般方面,一种方法包括:由任务管理器模块确定触发事件已经发生以提供用户呼叫请求的当前状态;由所述任务管理器模块确定所述用户呼叫请求的当前状态;生成用户呼叫请求的当前状态的表示;以及向用户提供所生成的用户呼叫请求的当前状态的表示。

[0013] 实施方式可以包括以下特征中的一个或多个。例如,所确定的触发事件可以是对状态的用户请求。所确定的触发事件可以是在操作员已经查看了与用户呼叫请求相关联的会话信息之后向用户提供状态的操作员交互。所确定的触发事件可以是状态更新事件。当前状态的表示可以是视觉表示。当前状态的表示可以是口头表示。向用户提供生成的用户呼叫请求的当前状态的表示可以包括确定用于将当前状态递送给用户的方便的时间和方法。

[0014] 在另一个一般方面,一种将电话呼叫转移离开机器人程序的方法包括:通过呼叫发起系统分析在电话呼叫的第一端上的第一人与电话呼叫的第二端上的机器人程序之间的电话呼叫期间在第一人与机器人程序之间的实时对话;由所述呼叫发起系统基于对所述实时对话的分析,确定是否应该将所述电话呼叫从所述机器人程序转移到所述电话呼叫的第二端上的第二人;并且响应于确定应该将所述电话呼叫转移到所述电话呼叫的第二端的第二人,所述呼叫发起系统将所述电话呼叫从机器人程序转移到第二人。

[0015] 实施方式可以包括以下特征中的一个或多个。例如,分析在电话呼叫期间在第一人和机器人程序之间的实时对话可以包括:基于第一人的行为、举止、语调、恼怒程度、语言或词语选择来确定电话呼叫期间的压力(strain)。该方法可以包括当机器人程序重复自身、道歉或要求澄清时确定电话呼叫期间压力增加。该方法可以包括当人纠正机器人程序或抱怨呼叫质量时确定电话呼叫期间压力增加。该方法可以包括当机器人程序适当地响应第一人的交谈时确定电话呼叫期间压力减小。分析在电话呼叫期间在第一人和机器人程序之间的实时对话可以包括确定呼叫发起系统的置信度等级,在该呼叫发起系统中电话呼叫的任务将由机器人程序完成。分析在电话呼叫期间在第一人和机器人程序之间的实时对话可以包括确定第一人要求将电话呼叫转移到另一人。分析在电话呼叫期间在第一人和机器人程序之间的实时对话可以包括确定第一人已经嘲讽机器人程序或询问机器人程序是否是机器人。确定是否应该将所述电话呼叫从机器人程序转换到第二人可以包括:确定压力高于预定义阈值;以及响应于所述压力高于预定义阈值,确定应当将所述电话呼叫从所述机器人程序转移到所述第二人。分析在电话呼叫期间在第一人与机器人程序之间的实时对

话可以包括跟踪对话中的一个或多个事件。确定是否应该将电话呼叫从机器人转移到第二人可以包括：使用基于特征的规则集，其确定对话中的一个或多个事件是否符合规则准则；以及响应于确定对话中的一个或多个事件满足规则准则，确定应该将电话呼叫从机器人程序转移到第二人。

[0016] 分析在电话呼叫期间在第一人和机器人程序之间的实时对话可以包括从对话中识别意图并且识别先前会话的历史意图和历史结果。确定是否应该将电话呼叫从机器人转移到第二人可以包括：将对话的意图、历史意图或历史结果发送到一个或多个机器学习模型；并基于意图、历史意图或历史结果确定是否应该转移电话呼叫。第二人可以是人类操作员。机器人程序可以使用与人类操作员相同的话音，使得从机器人程序到第二人的转移对于第一人是透明的。第二人可以是机器人程序正在为其进行电话呼叫的用户。该方法可以包括当将电话呼叫从机器人程序转移到第二人花费了超过预定时间量的时间时，终止电话呼叫。该方法可以包括终止电话呼叫而不是将电话呼叫转移到人。

[0017] 这个和其它方面的其他实施方式包括被配置为执行编码在计算机存储设备上的方法的动作的对应的方法、装置和计算机程序。一个或多个计算机程序可以被如此配置，凭借这样的指令，当由数据处理装置执行该指令时，使得该装置执行动作。

[0018] 可以实现本文档中描述的主题的特定实施例，以便实现以下一个或多个优点。因为仅存储确认数据的一个集合而不是存储未确认数据的多个集合，所以减少了各种数据源所需的数据存储量。例如，不用存储特定杂货店的未确认的营业时间的三个不同集合（例如，从店面收集的一个集合，从商店的网站收集的一个集合，以及从商店的应答机收集的一个集合），数据源可以存储从与杂货店的人类代表的通话中获得的确认的商店营业时间的一个集合。

[0019] 通过自动检测向呼叫发起系统指示呼叫将被发起的触发事件，减少了执行诸如从被呼叫者收集数据、排程预约、或向第三方提供信息等操作所需的人类输入量。另外，因为呼叫仅在触发事件发生时发起，所以维护信息数据库所需的计算机资源量由于所拨打的呼叫的减少而减少。系统会自动将呼叫发送到特定被呼叫者或被呼叫者的集合，减少人类必须执行的分析量和人类必须监视的数据量。

[0020] 此外，系统代表人类用户进行对话，进一步减少执行特定任务所需的人类输入量。呼叫发起系统可以同时协调多个呼叫。例如，用户可能希望将来预约三十分钟。系统可以呼叫用户指定的每个餐厅，并与其它线路的代表进行对话。呼叫的第一家餐厅的员工可能会建议进行预约，但用餐者必须坐在酒吧间。呼叫的第二家餐厅的员工可能会建议等待二十分钟，但呼叫的第三家餐厅的员工可能会通知系统第三家餐厅要求用餐者在一个小时内完成就餐，因此桌子会在一小时内准备好。该系统可以并行地呼叫三家餐厅中的每一家，通过呈现他的选项并接收响应来咨询用户，并在拒绝所有其他预约的情况下基于他的响应对最适合用户的餐厅进行预约。自动呼叫发起系统比人类更高效，因为自动系统可以一次完成这些呼叫。人类助理不能容易地并行进行所有这些餐厅的呼叫。

[0021] 在附图和下面的描述中阐述了本说明书中描述的主题的一个或多个实施例的细节。通过描述、附图和权利要求，主题的其他特征、方面和优点将变得显而易见。

附图说明

- [0022] 图1A示出呼叫发起系统的示例框图,该呼叫发起系统拨打呼叫并且在呼叫期间在呼叫发起系统的机器人程序与人之间进行对话。
- [0023] 图1B示出呼叫发起系统的示例框图,该呼叫发起系统拨打呼叫并且在呼叫期间在呼叫发起系统的机器人程序与人之间进行对话。
- [0024] 图1C示出用户可以通过其输入关于请求的更多细节的示例用户界面。
- [0025] 图1D示出用户向机器人程序说话以便做出请求的示例。
- [0026] 图2A是示出完成由用户指配的任务的过程的示例的流程图。
- [0027] 图2B是示出用于完成由用户指配的任务的过程的另一示例的流程图。
- [0028] 图3示出由系统执行的过程的示例 workflows。
- [0029] 图4是触发模块的框图。
- [0030] 图5是示出用于发起电话呼叫的过程的示例的流程图。
- [0031] 图6是系统的任务管理器模块的框图。
- [0032] 图7A图示示出关于现有任务的进度的信息的操作员控制板。
- [0033] 图7B图示用于查核用户请求的任务中的一个任务的操作员查核屏幕。
- [0034] 图8是示出提供任务状态的过程的示例的流程图。
- [0035] 图9A示出在进行预约排程时图1B的理发预约请求的视觉状态。
- [0036] 图9B示出了一旦预约被成功排程,图1B的理发预约请求的视觉状态。
- [0037] 图10A图示图1C的餐厅预约请求的口头状态请求和更新。
- [0038] 图10B图示由系统提供而无需由用户提示的用于图1C的餐厅预约请求的口头状态更新。
- [0039] 图11示出用于将电话呼叫从机器人程序转移到人类的示例过程1100。
- [0040] 图12是示出计算设备和移动计算设备的示例的示意图。
- [0041] 各附图中的相同附图标记和标号表示相同的元件。

具体实施方式

[0042] 本公开描述了一种允许在本文中被称为“机器人程序(bot)”的自动化或半自动化系统通过在呼叫期间拨打呼叫并独立地与人类进行对话来与人类通信的技术。机器人程序接收并监视数据以检测指示应该发起呼叫的触发事件。机器人程序可以通过预定义的工作流或可重复的操作模式的序列来工作,每个操作都由要执行的操作的抽象描述或意图进行链接。基本上,机器人程序能够使用这些工作流来确定如何反应和对人类说什么,以便为用户执行有用的任务。

[0043] 系统处理作为查询接收的各种任务,诸如“book a table for two at Yves Saint Thomas Restaurant for Thursday(在伊夫圣托马斯餐厅预订周四的两张桌子)”、“my sink is leaking and I need a plumber!It’s after 10p.m.!(我的水槽漏水,我需要水管工!晚上10点之后!)”等。

[0044] 希望排程预约、购买物品、请求服务等用户可能需要在完成他们想要达成的任务之前执行多次搜索并且拨打许多呼叫。在第一个用例中,在餐厅预订餐桌,用户可以在搜索引擎上搜索餐厅。在一些示例中,如果餐厅恰好在网站或应用上,则可以在网站或应用

(或通过与网站或应用的集成)执行查询,如果不是,则用户可以呼叫餐厅并协商预约。

[0045] 作为示例,系统可以用于为用户拨打呼叫。系统与商家和其他服务进行通信,以完成用户请求的任务。在一些示例中,机器人程序执行大部分通信。在一些示例中,人类操作员可以审查并验证由机器人程序执行的操作的成功。在一些示例中,人类操作员执行动作,并且机器人程序从人类操作员的通信中学习,以提高他们的自动化通信技能。

[0046] 在第二用例中,用户希望在正常营业时间之外找到水管工,这样的查询可能更难处理。例如,如果用户手动地搜索水管工,他可能在搜索引擎上搜索水管工并向他们中的几个给予呼叫。用户可能必须向每个水管工解释时间限制,他们位于何处以及问题的性质,并获得价格估计。这可能非常耗时。

[0047] 类似地,对于第三用例,检查本地商店在库存中是否有产品,用户可能必须搜索本地商店并且呼叫每一个本地商店以确定商店是否有他们正在寻找的特定商品或产品。

[0048] 除了帮助用户进行特定任务之外,系统还可以更新信息索引,例如营业时间、提供的服务等。响应于检测到丢失、老化、不一致的数据等,系统可以自动触发以更新数据。通常,为了获得这样的信息,用户可能需要单独检查每个商家或数据源。

[0049] 该系统提供许多优点,包括减少完成涉及发起电话呼叫的某些任务所需的人类输入的量。例如,系统可以基于确定已经满足某些触发条件来自动地发起电话呼叫,满足某些触发条件诸如由沙龙提供的服务与第三方预约网站上列出的服务之间的不一致性。该系统可以减少交易查询的摩擦,例如,通过检测电话呼叫的一端的人的沮丧或不适和结束通话或改变进行对话的方式。该系统可以将发展中国家的用户与诸如运输或教育服务的服务联系起来。该系统还可以将用户连接到没有网站或数字存在的低科技行业。此外,即使与最大的聚合器相比,该系统对于不同的应用是可扩展的。

[0050] 图1A示出用于呼叫发起系统的系统的示例框图,该呼叫发起系统拨打呼叫并且在呼叫期间呼叫发起系统的机器人程序和人类104和之间进行对话。下面将详细描述框图100中所示的每个组件。

[0051] 系统100包括各种组件和子系统,它们一起工作以允许机器人程序与人类104有效地通信。系统100可以包括:通信框架或平台102、拨号器106、声音系统108、呼叫触发模块或触发模块110、音频包112、会话记录器114、会话存储116、文本到话语模块118、话语端点检测器120、已存储的文本到话语结果或录音122、意图到文本模块124、话语到文本模块126、话语应用程序接口(API)128、文本到意图模块130、流管理器132、操作员控制器134和施援(bailout)模块136。在一些实施方式中,系统包括所有模块。在一些实施方式中,系统包括这些模块的组合。例如,在一个实施方式中,文本到意图模块是不必要的,并且意图直接地给到话语合成模块。

[0052] 图1B示出用于呼叫发起系统的系统的备选示例框图,该呼叫发起系统拨打呼叫并且在呼叫期间在呼叫发起系统的机器人程序和人类之间进行对话。在此示例中,通信平台102被替换为用于用户请求的客户入口点和用于其他请求(即来自商家的入站呼叫)的电话信令服务器。该系统将两种类型的请求发送到电话服务器(196),该电话服务器(196)与从另一端进行呼叫的机器人程序服务(195)进行呼叫。在一些实施方式中,机器人程序服务(195)包括对话模型(198)和语言模型(199),以使得机器人程序服务能够进行类似人类的电话对话。电话服务器(196)可以包括TTS模型(197)。话语识别器(191)和/或音频合成器

(194)可以为电话服务器(196)提供信息,以便在电话呼叫(190)的另一端理解并响应于人类。操作员(134)使用任务用户界面(160)和综合处理用户界面(170)来监视呼叫。操作员(134)可以审查来自录音工作室和评估TTS(114)的所记录的呼叫。呼叫播放器(162)将向操作员(134)重播呼叫。操作员可以使用本地代理(175)来排程呼叫,以通过电话服务器(196)发起电话呼叫。

[0053] 在图1A的实施方式中,通信平台102允许机器人程序通过执行诸如拨打呼叫、接收来自商家或用户(104,144)的入站呼叫或联系目标商家的任务来联系外部角色。通信平台102还允许机器人程序接收来自用户的请求以代表用户进行呼叫。

[0054] 在一些实施方式中,用户通过与用户界面的交互或者通过话语请求来请求对另一用户或商家的呼叫。这些用户请求可以用于辅助类型的任务,诸如:预订预约、进行餐厅预约、找到遛狗者、或者找出哪个商店有用户想要购买的物品。

[0055] 图1C示出了用户可以通过其输入关于该请求的更多细节的示例用户界面。用户可以通过点击“book(预订)”按钮或以某种其他方式与用户界面交互来发起请求。例如,如果用户想要进行理发预约,则用户可以与用户想要理发的沙龙相关联的网站进行交互。用户可以替选地与包括在搜索结果中作为结果的沙龙的搜索结果列表或者与在地图上示出沙龙的用户界面进行交互。这些界面中的任何一个可以允许用户请求呼叫。用户可以输入用户请求的细节,诸如:用户想要找的专业形象设计师、用户想要完成的服务的种类以及服务的日期和时间。如图1B所示,用户可以点击“continue with booking(继续预订)”按钮或采取其他动作来指示正在进行理发预约的请求。

[0056] 图1D示出用户为了进行请求而对机器人程序说话的示例。在用户向机器人程序说出请求之后,机器人程序可以肯定应答该请求。机器人程序还可以请求关于来自用户的请求的附加信息。例如,如果用户向机器人程序说出关于进行理发的请求,则机器人程序可以询问关于用户想要的理发预约的位置、应该排程的理发预约的日期和用户想要排程哪种类型的理发服务的信息。

[0057] 用户可以在任务不能执行时向系统进行任务请求。例如,在所有美发沙龙都关闭的晚上11点,用户可以请求呼叫以排程理发预约。因此,系统可以将请求存储在任务信息存储中,以在之后的日期或时间,诸如在由系统100以其他方式确定或获得的沙龙的营业时间期间发起和完成。

[0058] 在一些实施方式中,系统向用户提供在处理该请求中将存在延迟的初始反馈。例如,当用户在沙龙关闭的晚上11点请求呼叫以排程理发预约时,系统向用户提供视觉、音频或其他指示,其指示因为沙龙已经关闭,所以完成任务将存在延迟,直到一旦沙龙营业系统就能够联系上沙龙为止。

[0059] 在一些实施方式中,任务信息存储150存储关于每个任务的信息,诸如:请求该任务的用户的姓名、待呼叫的一个或多个人员或场所、所请求的任务的类型、进行任务请求的方法、关于类型特定的任务的细节、关于为了完成任务已经进行的活动的细节、任务的开始日期、任务的完成日期、对于进行请求的用户最后状态更新的时间、对该呼叫任务进行再次确认的操作员、用户对于该任务请求的结束日期、以及该任务的当前状态。

[0060] 在一些实施方式中,任务管理器模块140确定何时向人员或者商家排程呼叫。任务管理器模块140监视来自任务信息存储150的任务,并且确定对所接收到的任务进行排程的

适当时间。一些任务立即被排程，而其他任务则在发生特定触发事件后被排程。

[0061] 在一些情形下，在由系统100拨打的呼叫的另一端将存在人类，诸如人类104。该人类104可以是机器人程序尝试联系的组织的代表。在一些示例中，为了呼叫商家而使用通信平台。本系统100可以与通信平台集成。例如，本系统100可以使用用于测试web应用的框架 (framework) 来以编程方式操作web浏览器，并且使用基于web的电话会议服务。系统100可以创建并使用多个通信平台账号。在一些示例中，系统100可以自动地在通信平台账号之间交替以避免呼叫速度节流。

[0062] 拨号器106促进发起或拨打机器人程序进行的呼叫。拨号器106可通信地连接到通信平台102。拨号器106向通信平台102提供指令以向由拨号器106选择的特定被呼叫者发起电话呼叫。例如拨号器106可以播放与电话号码的数位相对应的音频音。一旦拨打了呼叫，则系统100可以与线路另一端的人类被呼叫者进行对话。

[0063] 拨号器106可以接收向特定被呼叫者发起呼叫的指令。例如拨号器106可以从系统100内的其他模块接收包含指令的数据，该其他模块诸如是触发器模块110或流管理器132。

[0064] 触发器模块110检测触发事件，或指示系统100应当向特定被呼叫者发起呼叫的特定事件。触发事件可以是预定类型的事件。例如系统100的用户可以指定特定类型的触发事件。触发事件可以包括由系统100的用户执行的明确的动作、在被提供至触发器模块110的数据中检测到的模式 (pattern)、从发生特定事件起经过的预定时间段、以及各种其他类型的事件。响应于检测到触发事件，触发器模块110向拨号器106提供指令以向特定被呼叫者发起呼叫，或者向流管理器132提供指令以选择特定工作流的节点或向拨号器提供指令。

[0065] 声音系统108用于记录和播放音频。在一些示例中，建立了三个虚拟流并将其用于记录整个呼叫：(a) 从电话或电话会议服务到系统100的传入音频；(b) 从系统100回到通信平台的传出音频；(c) 组合a和b的混合流。声音系统108使用音频包112来执行通过通信平台102的通信。

[0066] 音频包112被用于与声音系统108通信。在一些示例中，本系统100包括将音频包112卷包 (wrap) 并且处置传入音频分组的连续流的音频模块。模块还记录每个传入分组，并且允许回放预先记录的音频文件。本系统100使用各种位深、采样频率、分组大小等等。

[0067] 系统100可以记录由机器人程序进行的传入和传出对话。音频包112可以使得系统100能够使用会话记录器114记录特定会话或呼叫。在一些示例中，会话记录器114可以通过在机器人程序的话语生成时记录它，来记录由机器人程序进行的对话的一部分。在其他示例中，会话记录器114可以通过在机器人程序的话语通信系统102被输出到人类104时外部地记录它，来记录由机器人程序进行的对话的一部分。会话记录器114也可以记录人类104的响应。

[0068] 会话记录器114在会话存储116中存储所记录的会话数据。所记录的会话数据可以被存储为音频数据或者存储为表示音频数据的特征数据。例如，所记录的会话数据可以被存储为矢量，该矢量存储会话音频数据的特定特征的值。会话存储116可以是本地数据库、远程服务器、系统100内的物理存储器、或任何各种其他类型的存储器。

[0069] 话语端点检测器120简化机器人程序和线路另一侧的人类之间的对话。为了简化对话，该对话被分段为单独的句子，在人类和机器人程序之间分立地切换。话语端点检测器120负责从音频包112接收连续输入音频流并且将其转换为分立的句子。

[0070] 话语端点检测器120检测话语的端点。在一个实施方式中,话语端点检测器120以两种状态操作:等待话语;以及等待沉默(silence)。话语端点检测器120如下在这些状态之间交替:通过将每个音频分组的均方根偏差(RMSD)与预定义的阈值比较来检查该音频分组。如果单个音频分组的RMSD低于该阈值,则该音频分组被认为是“沉默”。无论何时接收到非沉默的分组,模块将从“等待话语”状态切换为“等待沉默”状态。模块只有在接收到持续了预定义时间段的连续沉默分组的时段后,才向回切换,该预定义时间段取决于整个系统的状态。

[0071] 在一些实施方式中,在“等待声音”时段期间,话语端点检测器120建构纯沉默分组(每十个真实分组建构一个建构的包)并且将其传送至话语到文本模块126。建构的分组可以避免从话语API 128断开。在“等待沉默”时段期间,话语端点检测器120传送来自流的沉默分组达预定义的时间段(对于基线噪声估计有用),并且然后传送所有音频分组。

[0072] 在一些实施方式中,话语端点检测器120使用被训练为观察语调和语境(language context)的机器学习、神经网络、或一些形式的深度学习来找到端点。

[0073] 在一些示例中,话语端点检测器120在确定如何解析音频输入的特定流时考虑说了什么、说话者的语调等。例如,话语端点检测器120可以确定特定人类被呼叫者104倾向于通过低音调变化来结束句子,并且话语端点检测器120可以在检测到音调变化下降时预测由被呼叫者104说出的句子的结束。话语端点检测器120可以基于在时间帧中的信噪比而在呼叫期间动态地调整阈值。

[0074] 话语到文本模块126将由话语端点检测器120解析的音频数据转换为文本,能够针对被用于选择机器人程序的下一个响应的意图来分析该文本。话语到文本模块126的输出是话语选项的有序列表,并且在一些情况下,提供了最佳选项的置信度。话语识别过程包括两个主要组成部分:声学模块和语言模块。对于声学模块,系统可以使用从直接对着电话交谈的人的记录训练的模型。该模型可以使用神经网络,并且在一些示例中,神经网络的第一层可以被再训练以虑及在电话呼叫中存在的声码器。声码器是从话语输入的分析产生声音的话音编解码器。神经网络也可以被再训练以虑及在对商家的呼叫和个人电话呼叫之间不同的背景噪声。语言模块可以使用基于系统的过往经验而对语言模块加偏置的系统来建立。在一些示例中,偏置可以被自动配置。在一些示例中,偏置可以被手动配置。在一些示例中,语言偏置配置在垂直元(vertical)之间改变。

[0075] 话语到文本模块126使用呼叫的场境(context)以便基于预测来自对话另一侧的人要说的话来对语言模块加偏置。例如,系统的机器人程序问“Are you open on Tuesday?(你们周二营业么?)”基于该问题,在对话另一侧的人将以诸如“no,we are closed(不,我们不营业)”或者“yeah,sure(是啊,当然)”的答案响应的可能性很高。机器人程序基于过往呼叫学习很可能的响应并且使用预测来理解传入音频。机器人程序可以预测完整句子响应,但是机器人程序也可以预测短语。例如,在机器人程序说“we have seven people in our party(我们的聚会有七个人)”之后,机器人程序可以预期短语“you said seven people?(你说七个人?)”机器人程序也可以预期短语“you said eleven?(你说十一个?)”,因为seven(七)和eleven(十一)听上去相似。机器人程序也可以预期更低概率的响应,诸如“you said two?(你说两个?)”机器人程序可以基于其预测来为每个短语指派概率权重。

[0076] 在一些实施方式中,话语到文本模块126使用话语API 128来将音频数据转换为文

本。在一些示例中，话语API 128使用机器学习来将音频数据转换为文本。例如，话语API 128可以使用接受音频数据作为输入的模型。话语API 128可以使用任何各种模型，诸如决策树、线性回归模型、逻辑回归模型、神经网络、分类器、支持向量机、归纳逻辑编程、模型组合(ensemble of models) (例如，使用诸如装袋、推进、随机森林等)、遗传算法、贝叶斯网络等，并且可以使用各种方式来训练，诸如深度学习、感知器、关联规则、归纳逻辑、聚类、最大熵分类、学习分类等。在一些示例中，话语API 128可以使用有监督学习。在一些示例中，话语API 128使用无监督学习。在一些示例中，话语API 128可以由话语到文本模块126通过网络访问。例如话语API 128可以由远程第三方提供在云服务器上。

[0077] 为了解决交谈的同步，诸如确定正在说话的人所处的场境，以便确定由机器人程序响应的自然时机，系统100可以识别意图(intent)。意图是在由人类或者由机器人程序说的句子中的单个语义含义的书面语言表达。在一些实施方式中，系统100忽略在接收的最后意图和机器人程序的回复之间从人类接收的任何意图，以便机器人程序产生与人类最新近所说句子相关的响应。然而，系统100可以使用先前意图来通知未来响应。例如系统可以将接收到的最新近意图之前接收到的意图标记为ANCIENT、解析ANCIENT意图、并将其存储以用于线下评估。在一些示例中，可以使用各种其他形式的处置逻辑。

[0078] 虽然系统100的大部分是与使用情况无关的，但系统100的一些部分可以针对特定使用情况被手动地配置或者完全编程，即系统垂直元。垂直元本质上由意图和商家逻辑代码的范式(schema)构成。范式中的意图是在由人类或者由机器人程序说的句子中的单个语义含义的内部书面语言表达。例如，在营业时间提取垂直元中，我们有“AreYouOpen {date: tomorrow}”机器人程序意图，以及对应的“WeAreClosed {date_range: Sep}”人类意图。将来自人类的传入音频转变为意图的过程在此被称为意图分辨(Intent Resolution)。相反过程(将机器人程序意图转变为话语)被称为话语意图(Intent to Speech)。虽然每垂直元配置范式，但是学习和分类意图的大部分代码是通用的并且跨垂直元使用。在一些示例中，仅系统的特定于语言的部分处于意图分辨和话语意图配置中。

[0079] 逻辑代码可以每垂直元编程(共享一些公共代码)并且对于由呼叫的场境(输入参数和到目前为止已经发生的事情)定义的每个可能情形以及传入人类意图确定机器人程序行为。在一些实施方式中，话语被改变为文本，该文本随后被解释为人类意图。人类意图被用于确定机器人意图。在一些垂直元中，机器人程序引导对话，而在其他情况下机器人程序主要对人类进行反应。例如：在数据获取类型的垂直元中，机器人程序以从商家提取一些信息为目标。通常，它将试图提出一系列问题直到它得到所有期望的信息。对于机器人程序以例如进行预约为目标的交易类型垂直元，它将主要回答源自人类的问题(“what is your name? (你叫什么名字?)”……“and phone number? (电话号码呢?)”……)。在这种情况下，系统只有在人类突然变得沉默等时才采取主导。编程者可以设计在人类意图和机器人意图之间的流，使得该转化符合逻辑。在一些实施方式中，存在用于流的协议，使得非工程师也能够进行控制以便改变或更新从人类意图到机器人意图的转化。也可以使用机器学习自动地学习流。

[0080] 在一些实施方式中，输入人类意图分辨可以是隐藏层，并且机器学习可以用于直接从输入文本学习输出机器人意图。人类话语输入可以被改变为文本，并且然后机器人意图可以从该文本直接确定。在又一个实施方式中，系统可以从人类话语直接输出意图。所有

这些设计使用机器学习来学习与场景和各个输入对应的机器人意图。

[0081] 文本到意图模块130使用可能的传入意图的范式、每个这样的意图的示例句子以及语言偏置配置来被配置。基本上,文本到意图模块130负责将传入的句子“捕捉(snap)”到预定义(或“未知”)意图列表中,同时虑及不熟悉的语句和话语识别过程中的错误。例如,在一些实施方式中,文本到意图模块130可以识别出句子(如从话语识别模块接收的)“who be open eleven o'clock in the morning till nine o'clock tomorrow nine thirty I'm sorry (谁将在早上十一点钟直到明天九点钟开放,抱歉九点半)”类似于已知的示例“we open at+(TIME,from_time)and close at+(TIME,to_time)i'm sorry”,这是一个意图“WeAreOpen{from_time=11am,to_time=9:30}”的示例。像“from_time”和“to_time”这样的字段是意图参数。

[0082] 文本到意图模块130可以由两个主要部分组成:(1)注释器和(2)注释的文本到意图分类器。在一些实施方式中,系统具有进行参数分类的后分类阶段。例如,对于短语“Monday to Tuesday,sorry Wednesday,we are closed,(周一到周二我们关闭,抱歉周一到周三)”,注释部分会将文本重写为“<DATE:Monday>to<DATE:Tuesday>,sorry<DATE:Wednesday>we are closed”。这个示例示出了使用指定响应文本中的注释的注释器重写短语。注释的文本意图分类会将注释的短语转换为:WeAreClosed{day1:monday,day2:tuesday,day3:wednesday}。后分类阶段将短语重写成:WeAreClosed{from_day:monday,to_day:wednesday,wrong_to_day;tuesday}。

[0083] 一旦系统100接收到话语选项,它就使用文本到意图模块130对其中的每一个注释日期、时间,公用名等。这是为了两个目的:(1)提取逻辑模块的意图参数(例如“time:10am(时间:10am)”);(2)概括文本以简化找到与之前遇到的句子的匹配。文本到意图模块130接收来自话语到文本模块126的输出并且注释话语选项列表。文本到意图模块130然后使用注释来将最可能的选项映射到由流管理器132用来选择特定工作流中的下一动作的意图。

[0084] 为了减少呼叫期间的计算时间,系统100可以预先构建应当被注释的已知文本的库(相对于当前时间和日期)。例如,在2015年9月1日周二,“tomorrow(明天)”,“this Wednesday(本周三)”,“September 2nd(9月2日)”等可以作为注释“DATE:(2015,9,2)”的候选。实时地,系统100对输入句子中的词(经过一些正则化之后)顺序地迭代,并且搜索注释候选的最长匹配。系统100然后用注释替换文本,并返回一个编辑的字符串,其中所有候选从左到右都被注释替换。例如,“we open at 7in the morning(我们在早上7点开放)”将被替换为“we open at@(TIME,7am)”。

[0085] 注释器或文本到意图模块130还可以负责收缩。例如,系统100可以遇到像“let me see...yes,4,4pm(让我想想……是的,4点,下午4点)”这样的句子。文本到意图模块130可以用单个“@(TIME,4pm)”注释替换“4,4pm”。此外,文本到意图模块130可能会收缩小的时间校正,诸如“we close at 10,ah,10:30pm(我们在10点,下午10点30分关闭)”到“we close at@(TIME,10:30pm)”。

[0086] 在其他实施方式中,系统可以使用用于注释文本的其他方法,例如:机器学习算法,其可以基于综合处理的数据学习如何注释文本;可用于注释文本的前缀树;或者可以专门为注释而导出的基于规则的模式。

[0087] 文本到意图模块130在大多数呼叫上解析和注释新的句子,并且话语识别经常扭

曲许多口语词。系统100具有每用例存储的数千个意图,并且基于句子的参数将每个句子分类为具有来自被确定为与句子最相关的存储的意图的意图。例如,基于检测到启示询问呼叫者姓名的问题的特定句子中的词,系统100可以将特定句子分类为具有询问姓名的意图。在一些实施方式中,系统100不识别句子的意图,并且可以将句子分类为具有未知意图。

[0088] 文本到意图模块130使用机器学习算法来处理分类。例如,系统可以使用条件随机字段模块与逻辑回归模块的组合。在一个实施方式中,分类是在句子水平上完成的,即文本的字符串被转换成意图集合或列表。在另一个实施方式中,原始字符串中的每个令牌都被分类为意图,并且意图边界也被分类。例如,在第一个实施方式中,句子“On Monday, we open at seven, let me think... on Tuesday, we open at eight, (周一,我们在七点开放,让我想.....周二,我们在八点开放)”将被分类为包含意图:GiveDailyHours(给出每日时间)+AskToWait(要求等待)。在第二个实施方式中,子字符串“on Monday we open at seven, (周一我们在七点开放)”将被分类为GiveDailyHours意图的边界,子字符串“let me think... (让我想...)”将被分类为类型AskToWait的另一个意图,并且子字符串“on Tuesday, we open at eight, (周二我们在八点开放)”将被分类为类型GiveDailyHours的另一个意图。

[0089] 在一些实施方式中,文本到意图模块130可以不使用机器学习算法,而是针对每个意图使用示例集合,并且然后在每个话语选项以及所有示例之间使用1-最近邻(模式识别算法),其中算法的距离度量是句子的词的归一化编辑距离(一种量化两个例如词的字符串彼此之间多么不相同的方法)的变体。两个词之间的距离更为复杂,且目的在于语音距离的近似。在一些示例中,语义距离可以由注释的文本到意图模块130来确定。

[0090] 在实践中,文本到意图模块130还可以使用交叉话语选项信号(例如,仅在话语选项之一中存在的数字可能是不良解释)。在一些示例中,文本到意图模块130基于系统场境将注释文本的结果偏向意图模块作为先验。最后,文本到意图模块130对于诸如“ComplexOpeningHours(复杂开放时间)”这样的虚拟定义的意图具有一些定制的提取,其中系统能够识别已经给出了开放时间的复杂措辞,但是系统无法正确地提取参数(例如“...dinner is served until 9, you can order desserts for one more hour, and the bar is opened until 2, but we don't accept customers after 1... (...晚餐供应至9点,另外有一小时您可以点甜品,酒吧开放至2点,但我们1点后不接待顾客...)”等)。

[0091] 在一些示例中,用于分类的示例基于综合处理的过去呼叫被自动推断,并且还可以被手动编辑。一般化过程可以用注释替代文本,并省略可疑的综合处理。

[0092] 在一些示例中,人不会说单一的意图,而是一系列意图。例如:“you want a haircut? what time? (你想要理发?什么时候?)”。示例系统支持给定句子中的任何数量的意图。在一些示例中,注释的文本到意图模块将具体的肯定意图和否定意图确定为其他意图的前缀(例如“no sir, we will be closed at this date (不,先生,我们将在此日期关闭)”=>否定+WeAreClosed)。在一些示例中,注释的文本到意图模块支持意图的任何链接。

[0093] 系统包括执行组织逻辑的不同功能的多个模块,包括流管理器132,其包括常识模块133;和施援模块136。

[0094] 流管理器132可以包括跟踪每个呼叫的每个垂直元的定制代码,并且确定如何响应从人接收的每个意图(或长时间的沉默)。然而,在其他实施方式中,流管理器是跨垂直元

通用的。响应是说给人104的合成意图的列表(机器人程序也可以选择保持沉默),且有时是结束呼叫的命令。流管理器132还负责生成呼叫的结果,包括在其期间收集的任何信息。在一些示例中,系统100学习如何对每个输入进行反应是基于最初由人类进行的、并且稍后由“儿童”机器人程序进行的现实呼叫。系统100使逻辑尽可能灵活,以解决在呼叫过程中出现的任何误解。

[0095] 系统100具有多个流,每个流针对特定类型的任务进行定制,例如确定商家的开放时间或为沙龙预约进行预约。系统100可以维护在不同流之间共享的公共库,并且从拨打的呼叫的历史中提取子流,允许系统为每个任务跳转一个新的垂直元。在一些示例中,系统100可以基于手动拨打的呼叫来自动学习针对不同任务的流。

[0096] 人们会在说话时跳过一些重要细节,而不会使他们的对话伙伴困惑。例如,一个人可能会说“we open 10to 4(我们从10点到4点开放)”。机器人程序需要了解该商家是否在上午10点或下午10点开放,类似地,是否在下午4点或凌晨4点关闭。例如,如果这家商店是夜总会,那么机器人程序可能被预期为假定下午10点-凌晨4点;如果商家是餐厅,那么机器人程序可能被预期为假定上午10点-下午4点,等等。

[0097] 流管理器132包括常识模块133,其消除所接收的话语输入中的意图的模糊化。在一些示例中,流管理器132包括多种类型的常识模块,例如,从某些数据集(例如,基线本地数据库)学习统计信息的模块,以及手动编程的模块。第一种类型的模块采用选项(例如开放时间)的集合,并计算每个选项和子选项(例如“凌晨2点至凌晨4点”或“仅凌晨2点”?)的p值。第二种类型的模块使用预定义的规则的集合,其防止系统做出可能驻留在数据集的“常识”错误。每当有多种方式来解释一些变量时,流管理器132可以组合两个分值来确定最可能的选项。在一些示例中,流管理器132得出没有选项是足够可能的结论,并且系统100返回以试图明确地要求人们阐明他们的意思。

[0098] 常识模块133可以使用来自相似的被呼叫者的数据来选择最可能的选项。例如,如果费城的大部分酒吧从晚上8点到凌晨2点营业,则常识模块133可以确定模糊短语“we're open from 10to 2(我们从10点开放到2点)”的最可能的选项是说话者的意思是晚上10点到凌晨2点。在一些示例中,常识模块133可以向流管理器132指示需要进一步的阐明。例如,如果密歇根州杰克逊市的大多数邮局的工作时间为上午10点至下午5点,如果系统100认为被呼叫者答复说他们的时间是“下午2点至下午6点”,这是与典型邮局不同的阈值量,常识模块133可以指示流管理器132要求阐明。

[0099] 有时,呼叫期间有累积的压力,通常是由于高背景噪声、异常场景、重口音、或仅仅是代码中的错误。压力也可能是意外的意图引起的。例如,当打电话给餐厅时,系统可能会遇到意想不到的句子:“so,do you want to make a presentation?(所以,你想做演示吗?)”或“we don't have a TV to show the Super Bowl,just so you know.(我们没有电视来显示超级碗,你知道的)”。系统应该处理其以前没有遇到的意图。为了识别任一方的有问题的情况,机器人程序尝试量化呼叫期间展示的压力量。施援模块136可以模拟监督呼叫的操作者,并选择何时实施手动干预。

[0100] 操作员控制器134可通信地连接到流管理器132,并且操作员控制器134允许人类操作员直接向流管理器132提供指令。在一些示例中,一旦呼叫被传送到人类操作员来处理,操作员控制器134将流管理器132置于保持模式或暂停或关闭流管理器132。

[0101] 一旦流管理器132基于从文本到意图模块130确定的意图来选择特定工作流程中的下一个节点,则流管理器132向意图到文本模块124提供指令。所提供的指令流管理器132包括通过通信平台102传送到被叫者的下一个意图。意图到文本模块124还产生预约话语合成的标记提示,例如在某些词上定义了不同的重音或韵律。意图到文本模块124可以使用手动定义的规则或强化学习来从意图产生新的文本。

[0102] 意图到文本模块124的输出是要转换为音频数据以便在通信平台102输出的文本。文本由文本到话语模块118转换成音频,文本到话语模块118使用先前存储的文本到话语输出和读取122。文本到话语模块118可以从存储的输出/读取122中选择先前存储的输出。在一些实施方式中,系统将在呼叫期间使用文本到话语合成器。例如,如果由流管理器132用于为机器人程序选择的待提供的常见响应是“Great, thank you for your help!(非常好,感谢您的帮助!)”,文本到话语模块118可以选择先前生成的文本到话语输出而不需要在运行时生成该输出。在一些示例中,文本到话语模块118使用类似于话语API 128的通过网络连接访问的第三方API。

[0103] 如上所述,在某些示例中,用户可以通过与提供给用户的搜索(例如,web搜索)结果交互来启动系统100的任务。例如,用户可以查找“book a table for two tonight at a Michelin starred restaurant(今晚在米其林星级餐厅为两人预订桌子)”。任务管理器模块140可以接收任务并将任务信息存储在任务信息存储150中。然后,任务管理器模块140可以确定何时排程任务并设置触发事件。例如,如果用户在米其林星标餐厅开放之前请求预订桌子,则任务管理器模块140可以确定餐厅何时开放并设置该时间的触发事件。如果任务管理器模块140知道由于触发事件而在处理中有延迟,则任务管理器模块140可以通过提供视觉、音频或其他指示来警告用户延迟。在一些实施方式中,任务管理器模块140可以提供关于完成任务所需的时间、何时将排程开始该任务的信息,或提供关于为什么任务被延迟的更多信息。

[0104] 触发模块110可以检测到特定的触发事件(在该示例中,餐厅的开放时间)已经发生,并且指示拨号器106拨打呼叫。在一些示例中,系统100可以向用户呈现选择要呼叫的餐厅的选项。在其他示例中,系统100可以自动地向基于特征集合所选择的特定餐厅拨打呼叫。用户可以定义用于为特定任务拨打呼叫的默认首选项。例如,用户可以指定系统100应该选择与用户当前所在位置最接近的餐厅来呼叫,或者系统100应该选择最高评价的餐厅来呼叫。

[0105] 在某些示例中,系统100包括或形成通信应用的一部分或配置为与通信应用通信,通信应用诸如消息传递或聊天app,包括用户界面,用户通过该界面向系统提供对协助任务的请求。例如,用户可能能够使用消息传递应用,在请求中文本输入数字,例如“does Wire City have 20AWG wire in red?(Wire City是否有红色的20AWG线?)”。系统可以接收文本消息,解析请求以确定触发事件已经发生,并发起呼叫以执行适当的动作。例如,系统可以拨打最近的Wire City来询问他们现在库存是否有20规格红线。

[0106] 类似地,在某些示例中,系统100包括或形成虚拟辅助系统的一部分,或配置为与虚拟辅助系统通信,虚拟辅助系统本身是用于辅助用户各种服务或任务的软件代理的集。例如,用户可以向虚拟助理输入(通过语音或文本录入)“is my dry cleaning ready?(我的干洗准备好了吗?)”虚拟助理可以处理该输入,并确定需要与商家的通信以满足该查询,

并因此与系统通信以识别意图,拨打呼叫,并执行相应的工作流。

[0107] 在某些示例中,系统100分别通过与多个人的多个对话自动地执行任务,并且可以对对话进行收集、分析、采取行动和/或呈现对话的个体或累积结果。例如,如果将任务分配给系统100来收集关于在特定区域中的多个餐厅最忙的时间是何时的数据,则系统100可以自动地向每个餐厅拨打呼叫,询问在一段时间内有多少客户就座来分析数据并提供结果。

[0108] 图2A示出完成由用户指派的任务的示例过程200。简而言之,过程200可以包括将对话映射至预定义 workflows 集合中的初始节点,每个预定义 workflow 通过意图来链接(202),基于 workflow 中的当前节点选择传出消息(204),从人类用户接收响应(206),将该响应映射至预定义 workflow 内的意图(208),基于该意图选择下一个节点作为该 workflow 内的当前节点(210),以及重复204-210直至到达预定义 workflow 中的链接的节点集合中的终端节点。该过程200可以由呼叫抑制系统诸如系统100来执行。

[0109] 过程200可以包括将对话映射至预定义 workflows 集合中的初始节点,每个预定义 workflow 通过意图链接(202)。例如,如上关于图1所述的流管理器132可以将对话映射至预定义 workflows 集合中的初始节点,每个预定义 workflow 通过意图链接。在一些示例中,系统100和人类被呼叫者之间的对话可以由用户发起。在一些示例中,该会话包括映射至预定义 workflows 集合中的节点的意图。例如,系统100可以存储具有待被执行的动作的预定义 workflows 集合。在一些示例中,该系统可以基于所识别的意图选择预定义 workflow。多个 workflow 中的每一个可以通过意图链接。在一些示例中,系统100可以向由用户在对话中指定的商家拨打电话呼叫。在一些示例中,商家可以是餐厅、沙龙、医生办公室等。在一些示例中,该系统可以仅在人类接听的情况下认为成功拨打了呼叫,而如果没人接听或者该系统被指向电话树并且没有成功导航电话树,则该系统可以确定没有成功拨打呼叫。

[0110] 过程200可以包括基于 workflow 中的当前节点选择传出消息(204)。例如,如果 workflow 中的当前节点指示用户想要排程理发预约,则流管理器132可以选择说“Hello, I would like to schedule an appointment for a haircut (你好,我想排程理发预约)”的消息。

[0111] 过程200可以包括从人类用户接收响应(206)。例如,系统100可以从在电话呼叫的另一端的人类被呼叫者接收响应,诸如“Sure, and what time and date would you like to schedule this appointment for? (当然可以,您想要在什么日期和时间排程该预约?)”。在一些示例中,系统100可以记录该响应(例如,使用会话记录器114)。在一些示例中,系统100可以为人类操作员再现该响应。在一些示例中,人类操作员可以监视该呼叫(例如,使用操作员控制器134)。

[0112] 过程200可以包括将该响应映射至预定义 workflow 内的意图(208)。该流管理器132可以将该响应映射至预定义 workflow 内的意图。在一些示例中,该系统将所识别的意图与链接每个预定义 workflows 集合的意图相比较。

[0113] 过程200可以包括基于该意图选择下一个节点作为该 workflow 内的当前节点(210)。例如,流管理器132可以使用该意图确定该 workflow 中的下一个节点。流管理器132然后可以规定该下一个节点作为当前节点。过程200可以包括重复204-210直至到达终端节点。因此,在204-210的每个重复周期中使用所规定的当前节点来确定下一个传出消息,直至到达终端节点。

[0114] 图2B示出用于完成由用户分配的任务的示例过程250。简而言之,过程250可以包

括从用户接收与意图相关联的任务 (252), 识别该意图 (254), 基于该意图从通过意图链接的预定义 workflow 集合中选择预定义 workflow (256), 遵循该预定义 workflow (258), 以及完成该任务 (260)。该过程 250 可以由呼叫发起系统诸如系统 100 来执行。

[0115] 过程 250 可以包括从用户接收与意图相关联的任务 (252)。例如, 用户可以通过用户界面向系统 100 提交搜索查询“book an appointment for a haircut (预订理发预约)”。在一些示例中, 该搜索查询可以被触发模块 110 接收, 该触发模块 110 检测到该查询是指示应该向特定被呼叫者拨打呼叫的触发事件。该任务可以是预订预约, 并且该意图可以是理发。在一些示例中, 可以不明确地录入该任务或意图。在一些示例中, 用户可以在不录入搜索查询的情况下提交任务和意图。与意图相关联的任务可以由用于协助任务的系统接收。

[0116] 过程 250 可以包括识别该意图 (254)。例如, 系统 100 可以处理与意图相关联的所接收的任务并且识别该意图。在一些示例中, 该意图可以被明确地输入并且与该任务分离。在一些示例中, 该意图可以是该任务的属性。在一些示例中, 该输入作为话语输入被提供, 并且该话语端点检测器 120 将经解析的输出提供给将文本发送到识别意图的文本到意图模块 130 的话语到文本模块 126。

[0117] 过程 250 可以包括基于该意图从通过意图链接的预定义 workflow 集合中选择预定义 workflow (256)。例如, 系统 100 可以存储具有待被执行的动作的预定义 workflow 集合。在一些示例中, 该系统可以基于从 (254) 所识别的意图选择预定义 workflow。例如, 流管理器 132 可以基于从 (254) 由文本到意图模块 130 所识别的意图选择预定义 workflow。在一些示例中, 该系统将所识别的意图与链接每个预定义 workflow 集合的意图相比较。

[0118] 过程 250 可以包括遵循该预定义 workflow (258)。例如, 系统 100 可以包括遵循在该预定义 workflow 中包括的指令的模块。在一些示例中, 系统 100 的机器人程序可以遵循在该预定义 workflow 中包括的指令。例如, 指令可以包括指令该触发模块 110 将控制数据提供给拨号器 106 来向商家的人类代表拨打呼叫并且与该人类代表对话。

[0119] 过程 250 可以包括完成该任务 (260)。例如, 系统 100 可以完成所指派的整体任务, 诸如支付账单、改变晚餐的预留等。在一些示例中, 系统 100 可以完成该任务的一部分, 诸如呼叫并导航电话树直到到达人类。在一些示例中, 系统 100 可以完成由用户指定的该任务的一部分。例如, 该用户可以指定该系统完成所有任务并且将该呼叫转移给该用户以验证。

[0120] 许多使用案例可以涉及想要从商家购买东西的用户由于交易所需的复杂度、菜单导航、语言挑战、参考知识等而很难进行购买。交易查询可以在卖方侧从人类收集愿意帮助系统成功完成该交易的支持。在一些示例中, 该系统可以在发展中国家以及低科技和服务产业诸如水管、盖顶等提供关键协助。可以采用 workflow 来既协助人类用户成功导航这样的交易又鼓励卖方侧系统协助用户。该系统可扩缩来适应各种使用案例。例如, 餐厅预订应用可以与全世界数以千计的商家合作; 无论所需的规模多大, 本文公开的系统可以被配置为以所需的规模发出餐厅预订。

[0121] 图 3 示出由该系统执行的过程的示例 workflow 300。在该特定示例中, 由系统 100 的机器人程序问简单的布尔问题。应理解的是, 该系统能够回答高复杂度的问题, 并且为了解释的简单起见呈现该 workflow 300。

[0122] 流 300 示出由机器人程序提出的示例问题“Are you open tomorrow? (您明天开业吗?)”。展示了由人类提供的可能响应, 并且提供了对于人类的响应中的每一个的机器人程

序的响应。取决于人类的响应,存在系统100可以被指向的流300的若干阶段。以双边框示出的阶段是系统100退出流300的终端阶段。例如,响应于由机器人程序提出的二元问题,人类被呼叫者可以确认商家明天开业,结束流300。人类被呼叫者可以确认商家明天不开业,结束流300。人类被呼叫者可以要求机器人程序保持通话,因此将机器人程序发送到与流300分离的流,并且结束流300。

[0123] 为了便于用户访问并促进系统100的传播,系统100与现有应用、程序和服务整合。例如,系统100可以与用户的移动设备上的现有搜索引擎或应用整合。与其它服务或垂直元整合允许用户容易提交对于待被完成的任务的请求。例如,系统100可以与搜索引擎知识图整合。

[0124] 在一些使用案例中,可以自动化对人类的实时判断。例如,系统100可以自动检测用户将比理发店预约晚十分钟并在用户到达以前告知理发师。

[0125] 系统100可以基于正进行的对话的场境或关于在知识数据库内存储的特定被呼叫者的数据来选择机器人程序的某些参数。例如,系统100可以基于被呼叫者的口音、位置和其它场境数据来确定该被呼叫者对利用与当前进行的呼叫的语言不同的语言更为舒适。系统100然后可以切换至该机器人程序认为该被呼叫者更为舒适的语言,并且询问该被呼叫者他或她是否更喜欢用新的语言进行呼叫。通过反映该人类被呼叫者的某些话语特性,系统100增加了成功呼叫的概率。系统100减少了由于话语特性在对话内的潜在摩擦源以减少在呼叫期间累积的压力。这些特性可以包括所使用词的平均长度、句子结构的复杂度、短语之间的停顿长度、该被呼叫者最舒适讲的语言、以及各种其它话语特性。

[0126] 图4是系统100的呼叫触发模块的框图400。触发模块110可通信地连接到拨号器106,并且基于检测触发事件将指令提供给拨号器106以向特定被呼叫者或被呼叫者集合发起呼叫。在一些示例中,触发模块110可以与流管理器132通信以提供流管理器132用来选择特定工作流的节点的触发事件数据或者将指令提供给拨号器106。

[0127] 触发模块110从各个模块接收输入,所述模块包括不一致性检测器402、第三方API 404、趋势检测器406、和事件识别器408。触发模块110也可以从流管理器132接收输入。在一些示例中,模块402-408中的每一个与系统100整合。在其它示例中,模块402-408中的一个或多个远离系统100,并且通过网络连接到触发模块110,所述网络诸如局域网(LAN)、广域网(WAN)、互联网或其组合。网络可以将模块402-408中的一个或多个连接到触发模块,并且可以促进系统100的组件之间(例如,在话语API 128和话语到文本模块126之间)的通信。

[0128] 不一致性检测器402从多个不同源接收数据并且检测在来自第一数据源的数据值和来自第二源的对应数据值之间的不一致性。例如,不一致性检测器402可以接收指示诊所的营业时间的数据并且检测到在该诊所的网站上列出的诊所的营业时间不同于在该诊所的外面张贴的营业时间。不一致性检测器402可以将指示争论的源、检测到不一致性的数据值的类型、冲突的数据值和各种其它特性的数据提供给触发模块110。在一些示例中,不一致性检测器402向触发模块110提供用以向特定被呼叫者发起呼叫的指令。在其它示例中,触发器模块110基于从不一致性检测器402接收的数据确定要联系的特定被呼叫者以及待从该特定被呼叫者收集的数据的字段。

[0129] 触发模块110可以基于由不一致性检测器402提供的数据检测触发事件。触发事件可以包括接收指示矛盾的用户输入。例如,触发模块110可以通过用户界面410接收用户输

入。用户界面410可以是单独应用或程序的界面。例如，用户界面410可以是搜索引擎应用或导航应用的图形用户界面。

[0130] 在一些实施方式中，用户界面410可以提示用户提供信息。例如，用户被检测到在商店广告的闭店时间之后在该商店，该系统可以问该用户该商店是否还开着或者让该用户输入营业时间。该用户可以通过用户界面410输入所请求的数据，并且不一致性检测器402可以确定在通过用户界面410输入的数据和来自第二源诸如知识库412的对应数据之间是否存在矛盾。知识库412可以是存储介质，诸如远程存储设备、本地服务器或任何各种其它类型的存储介质。不一致性检测器402可以确定该用户是否在正常营业时间以外的预先确定的时间量内在该商店（例如，多于20分钟，因为商店可能为比较慢的顾客保持开业一些额外分钟）。

[0131] 在另一个示例性情形下，不一致性检测器402可以确定关于组织的网站的信息是否过时了。例如，不一致性检测器402可以基于来自知识数据库412的数据检测到巴斯钓鱼俱乐部的网站示出在每月的第一个星期三有月度会议，但是该俱乐部的更活跃的社交媒体简档都指示月度会议发生在每月的第二个星期二。不一致性检测器402然后可以向触发模块110输出指示该检测到的不一致性的数据。

[0132] 触发事件可以包括确定数据的特定集合在预定时间量内尚未被更新。例如，系统100的用户可以指定时间量，在该时间量之后不管发生任何其他触发事件，数据都应该被刷新。不一致性检测器可以比较特定数据值的最后更新的时间戳，并且基于该时间戳确定预定的时间量是否已经过去。包括时间戳和数据值本身的特定数据字段的特征可以存储在知识数据库412中。定时器414可以向知识数据库412提供数据以更新已经流逝的时间量。不一致性检测器402可以基于由定时器414提供的定时数据来确定已经流逝了预定时间段。

[0133] 例如，不一致性检测器402可以基于来自知识数据库412的数据确定纽约州伊萨卡的小咖啡店的营业时间在三个月内还没有更新。然后，不一致性检测器402可以向触发模块110提供指示所检测到的事件的输出数据。

[0134] 触发事件可以包括从一个或多个用户接收对发起呼叫的请求。例如，触发模块110可以通过第三方API 404检测来自用户的请求。第三方API 404可通信地连接到诸如用户界面416的用户界面，用户可以通过该用户界面提供指示对发起呼叫的请求。例如，用户界面416可以是应用的图形用户界面，用户可以通过该图形用户界面请求要被排程和执行的呼叫活动。用户可以提供指示特定被呼叫者或被呼叫者的集合的数据，以及被请求以用于提取的特定数据。例如，用户可以请求对维吉尼亚州每家销售牲畜物资的五金商店进行呼叫活动，并询问该五金商店是否提供小鸡饲喂器(chick starter feed)（例如，使得携带物资的位置的索引对以后搜索可用）。

[0135] 在呼叫期间，被呼叫者可以为系统100排程不同的时间来对被呼叫者进行回叫。例如，如果被询问餐厅的菜单是否有任何改变，则人类被呼叫者可以要求系统100在他们有机会察看新菜单之后，在一个小时内或次日对他们进行回叫，以用于进一步的操作。然后，系统100可以对所请求的时间的呼叫进行排程。在一些示例中，触发模块110可以对未来的触发事件进行排程。在其他示例中，流管理器132可以排程要由拨号器106执行以发起呼叫的意图或呼叫事件。

[0136] 触发事件可以包括在知识数据库内的所存储的数据或实时提供的数据中检测到

的趋势或模式。例如，从搜索引擎418接收的搜索数据中检测到的趋势可以是触发事件。搜索引擎418从用户接收搜索请求，并且可以向趋势检测器406提供指示搜索请求的数据。趋势检测器406分析所接收到的数据并在所接收到的数据中检测趋势。例如，如果对北卡罗来纳州阿什维尔的古巴餐馆的搜索在过去一个月中增加了500%，则趋势检测器406可以检测到搜索的增加，并向触发模块110提供指示趋势的数据。

[0137] 趋势检测器406可以基于所识别到的趋势将指示特定的被呼叫者或被呼叫者的集合的数据输出到触发模块110。在一些实施方式中，趋势检测器406提供指示所检测到的趋势的数据，并且触发模块110基于所识别到的趋势来确定特定被呼叫者或被呼叫者的集合。例如，趋势检测器406可以确定针对“内布拉斯加州林肯的龙卷风”的搜索已经增加了40%，并且将搜索的关键词提供给触发模块110。然后，触发模块110可以确定应该向提供应急物资的所有商店拨打呼叫，以查明每个店铺有多少库存必需品，及其营业时间（例如，用于索引和搜索引擎的用户以后的搜索）。

[0138] 触发事件可以包括被识别为对商家、组织、个人等的正常运营具有影响的特定感兴趣事件。事件识别器408从包括第三方数据库420和事件数据库422的各种第三方源接收数据。事件识别器408可以从诸如本地存储器设备或实时数据流的其他源接收数据。事件识别器408识别来自数据库420和422的特定事件，并将指示所识别的事件的数据输出到触发模块110。在一些示例中，触发模块110基于由事件识别器408提供的数据选择特定被呼叫者或被呼叫者的集合以及将要在呼叫期间请求的数据。

[0139] 可以影响商家、组织和个人的运营的特定事件包括极端天气条件、联邦假期、宗教假期、体育赛事、以及各种其他事件。

[0140] 第三方数据库420向事件识别器408提供来自各种第三方数据源的数据，包括气象服务、政府警报等。例如，第三方数据库420可以向趋势识别器408提供风暴警告。事件识别器408然后可以确定冬季风暴正在接近明尼苏达州明尼阿波利斯的东北角，并且可以确定应该对明尼阿波利斯东北角的五金商店拨打呼叫以确定可用的发电机的现有库存。

[0141] 事件数据库422向事件识别器408提供来自各种数据源的数据，并且具体包括指示已知事件的数据。例如，事件数据库422可以提供指示联邦和州假期、宗教节日、游行、体育赛事、展览开放、访问政要以及各种其他事件的数据。

[0142] 例如，如果特定城市正在主办超级碗，则事件数据库422可以向事件识别器408提供数据，事件识别器408将指示该事件的数据提供给触发模块110。触发模块110可以基于关于当前超级碗的已知信息和关于过去超级碗的存储信息，确定应该对该地区的所有酒店拨打呼叫以确认可用性和定价。触发模块110还可以确定应该对体育用品商店拨打呼叫以确定参与超级碗的每个球队的球衣的可用性。在这样的情况下，触发器模块110可以请求的影响商家、组织或个人的运营的其他信息包括：办公楼和学校的关闭、公共交通调度的更改、特殊的餐厅供应或各种其他信息。

[0143] 系统100的各个模块中的一个或多个可以基于从事件识别器408接收到的事件信息来确定所推断的触发事件或要请求的信息。例如，如果时逢亡灵节(Día de Muertos)，则南美餐厅特别是墨西哥餐厅可能会因为庆祝活动而有特别的菜单或营业时间。在这样的示例中，触发模块110可以向拨号器106提供指令以对南美餐馆进行呼叫以更新当天的营业时间和菜单。

[0144] 在一些实施方式中,可以从系统100本身拨打的呼叫中检测触发事件。流程管理器132可以基于系统100进行的会话的部分来确定在会话期间已经表达了建议应该拨打呼叫的意图。例如,如果人类被呼叫者说:“Yes,we are still open until 8p.m.every Thursday,but next week we will be switching over to our summer schedule,and we will be open until 9:30p.m.(是的,我们还是每个星期四营业到晚8点,但下周我们将切换到我们的夏季时间表,我们将营业到晚9:30)”然后流量管理器132可以识别提供有关数据字段的进一步信息的意图。

[0145] 在一些实施方式中,触发事件可以包括从先前拨打的呼叫接收不令人满意的结果。例如,如果机器人程序向商家拨打呼叫以确定该商家是否在独立日假期有特殊的假日营业时间,并且对由该商家的人类代表提供的答案的准确性的置信度没有达到至少阈值量,则系统100可以排程另一个特定日期或时间进行呼叫,例如7月1日,以确定是否将有特殊的假日营业时间。在这样的示例中,触发模块110可以排程触发事件或向流程管理器132提供信息以排程动作。在一些示例中,流管理器132通过排程向拨号器106的指令传输来排程发起回叫。

[0146] 系统100具有常识模块133,其允许流管理器132智能地排程和选择特定工作流的节点。例如,在上述情况下,当在呼叫期间存在正在被请求的信息的有用性的最后期限时,常识模块133还可以确定何时排程呼叫和要请求哪些信息。在一些示例中,常识模块133是流管理器132的组件,如图1所示。在其他示例中,常识模块133是触发模块110的组件,并且便于触发模块110进行关于是否应该发起呼叫的智能确定。

[0147] 图5示出了用于发起电话呼叫的示例过程500。简而言之,过程500可以包括:由用于拨打呼叫并且在呼叫期间在呼叫发起系统的机器人程序和人类被呼叫者之间进行对话的呼叫发起系统的呼叫触发模块来接收指示第一事件的数据(502),通过所述呼叫触发模块并使用指示所述第一事件的数据来确定所述第一事件是触发开始于发起电话呼叫的所述呼叫发起系统的工作流的触发事件(504),基于所确定的触发事件来选择特定工作流(506),并且响应于所述选择而发起对由所述特定工作流指定的被呼叫者的电话呼叫(508)。

[0148] 过程500可以包括由用于拨打呼叫并且在呼叫期间在呼叫发起系统的机器人程序和人类被呼叫者之间进行对话的呼叫发起系统的呼叫触发模块来接收指示第一事件的数据(502)。例如,触发模块110可以从不一致性检测器402接收数据,该数据指示在商店网站上发布的萨丽糖果车(Sally's Saloon of Sweets)的营业时间与存储在与该商家相关的搜索索引中的营业时间之间的差异。

[0149] 过程500可以包括由呼叫触发模块使用指示第一事件的数据来确定所述第一事件是触发事件,该触发事件触发了开始于电话呼叫的呼叫发起系统的工作流(504)。在一些示例中,所确定的触发事件是与第一数据源相关联的值和与第二数据源相关联的对应值的不一致性。例如,触发模块110可以使用来自不一致性检测器402的检测到的不一致性来确定该不一致性是触发事件,其将触发工作流以确定萨丽糖果车的实际营业时间。

[0150] 在一些示例中,指示第一事件的数据由用户提供。例如,用户可以报告萨丽糖果车网站上发布的时间与萨丽糖果车店面发布的时间之间存在差异。

[0151] 在一些示例中,所确定的触发事件是用户请求。例如,用户可以通过用户界面(诸

如用户界面416)向诸如第三方API 404的第三方API提供输入,以请求排程和执行对特定被呼叫者或特定被呼叫者的集合的呼叫。

[0152] 在一些示例中,所确定的触发事件是以下之一的特定类型的事件:天气事件、体育赛事、娱乐事件、或季节性事件。例如,事件识别器408可以确定查尔斯河帆船赛正在马萨诸塞州波士顿发生,并且可以向触发模块110提供事件数据。然后,触发器模块110可以确定该帆船赛是触发事件。

[0153] 在一些示例中,所确定的触发事件是在提交给搜索引擎的搜索请求中检测到的趋势。例如,趋势检测器406可以从搜索引擎418接收搜索引擎数据并确定西班牙小吃餐厅正在趋势化。趋势检测器406可以向触发模块110提供指示该趋势的数据,并且触发模块110可以确定该趋势是触发事件。

[0154] 在一些示例中,所确定的触发事件是流逝预定时间段。例如,不一致性检测器402可以基于来自定时器414的知识数据库412中的数据确定纽约曼哈顿的古巴餐馆的菜单在四个月内还没有被更新。不一致性检测器402可以将定时数据提供给触发器模块110,并且触发器模块110可以确定流逝了四个月而没有更新曼哈顿的古巴餐馆的菜单数据是触发事件。然后,触发器模块110可以向流管理器132提供建议呼叫曼哈顿的古巴餐馆以获得更新的菜单信息的数据。

[0155] 过程500可以包括基于所确定的触发事件来选择特定工作流(506)。触发模块110可以向拨号器106或流管理器132提供触发事件数据,以用于选择特定工作流或工作流的节点。例如,触发模块110可以向流管理器132提供指示萨丽糖果车的所发布的营业时间的不一致性的触发事件数据,流管理器132使用该数据选择特定工作流来呼叫萨丽糖果车以解决该差异。

[0156] 过程500可以包括响应于选择发起到由特定工作流指定的被呼叫者的电话呼叫(508)。流管理器132可以将指示将被联系的特定被呼叫者的指令提供给拨号器106。例如,流管理器132可以将指令提供给拨号器106以呼叫Sally的发廊(Saloon)。

[0157] 在本文所述的系统和方法对工作流的发起——特别地是对呼叫的分配——可以通过触发事件而相对自动化时,可以在系统100中包括保护措施以防止违反局部规则的不想要的呼叫。例如,如果被呼叫者指示他们不再希望接收来自系统的呼叫,则系统可以对此进行记录并且针对到被呼叫者的号码的呼叫建立检查以防止进一步的呼叫。

[0158] 此外,在本文所述的系统和方法收集数据的情况下,在存储或使用数据之前可以以一种或多种方式对所述数据进行处理,使得个人可识别信息被移除或永久性被模糊。例如,被呼叫者的身份可以被永久移除或被处理使得不能确定个人可识别信息,或者可以从其获得位置信息的被呼叫者的地理位置可以被泛化,使得在适当的情况下不能确定用户的特定位置。其中,在呼叫期间接收到个人的、隐私的或敏感的信息的情况下,无论该信息作为工作流的部分被请求、被呼叫者自愿提供、还是非故意地被接收,工作流可以包括用于对所述来自系统的信息进行永久性移除或模糊的步骤。

[0159] 在某些示例中,系统100可以自动地或在用户请求之后为用户提供当前的努力状态以执行对辅助任务的请求。例如,系统100可以通过用户正在使用的设备(诸如计算机、移动设备等)上的通知为用户提供任务被实施的状态。在一些示例中,系统100可以通过其他装置(诸如消息传递应用、通过电话通信等)通知用户正在进行的任务的状态。

[0160] 图6是系统100的任务管理器模块的框图600。任务管理器模块140被连接到通信平台102、触发模块110、任务信息存储150、以及会话存储116。当用户通过通信平台传送任务时,任务信息被存储在任务信息存储150并且任务管理器模块140确定所述任务何时应该被排程。任务管理器可以将任务与触发事件相关联。任务可以具有初始设置为“新”的状态或不存在对请求的处理的一些其他指示符。一旦触发事件发生,触发模块110将开始拨号过程。在一些实施方式中,任务管理器模块140监视会话存储以在任务的状态从初始化改变为进行中、改变为完成时,更新每个任务的状态。

[0161] 通过会话信息,任务管理器模块可以确定每个呼叫的状态和结果。例如,机器人程序可以在电话接通到某个人来做出预约之前尝试多次呼叫餐厅。会话存储保持与机器人程序做出的每个呼叫有关的信息。在一些实施方式中,任务管理器模块可以周期性地轮询会话存储以确定呼叫任务的状态,即,呼叫是否被初始化、在进行中还是已经完成。在其他实施方式中,会话存储可以将呼叫的结果发送到任务管理器模块以便更新任务信息存储中的任务的状态。

[0162] 在一些实施方式中,操作者通过显示呼叫任务和与任务的进行有关的信息的操作者控制板查核呼叫。

[0163] 图7A图示示出与现有呼叫任务的进行有关的信息的操作者控制板。例如,图7示出了理发预约的任务。操作者控制板可以提供与包括以下各项的预约有关信息:预订时间、请求者的姓名、所请求的服务、公司名称、日期、以及预约时间。操作者可以能够通过请求相关联的呼叫查核请求以及相关会话信息,以确定所请求的预约是否已经被正确地预订。

[0164] 图7B图示了用于查核用户所请求的任务中的一个的操作者查核屏幕。屏幕可以为操作者示出任务的当前状态。如图7B中所图示的,由于已经做出预约,任务完成。然而,在一些情况下,任务可能没有完成并且预约可能还没有做出。操作者可以具有以下选项:播放与任务相关联的记录或看来自呼叫的其他所存储的信息(例如,转录、所提取的意图等)、呼叫与任务相关联的商家、或对未来的自动呼叫进行排程。附加地,操作者可以具有将任务的当前状态提供给进行请求的用户的选项。

[0165] 用户还可以通过通信平台102请求任务的状态。附加地或备选地,任务管理器模块140可以基于其他触发事件(诸如任务状态改变或时间)确定何时向用户发送状态更新。

[0166] 图8是示出用于提供任务的状态的过程800的示例的流程图。过程800可以包括由任务管理器模块确定触发事件已经发生以提供用户呼叫请求的当前状态(802)。如上所述,触发事件可以包括:对状态的用户请求、某个时间量的流逝、或者特定任务状态的改变。接着,过程800包括由任务管理器模块确定用户呼叫请求的当前状态(804)。任务管理器模块可以通过检查任务信息存储中的状态来确定当前状态。当任务被添加到任务信息存储150时,任务的状态被初始化。当与任务相关联的呼叫被做出并完成时,任务的状态被更新。接着,任务管理器生成用户呼叫请求的当前状态的表示(806)。表示可以是传达任务的当前状态视觉的或音频表示。过程800为用户提供所生成的用户呼叫请求的当前状态的表示(808)。

[0167] 图9A示出了在预约排程正在进行时图1B的理发预约请求的视觉状态。用户可以能够访问用户界面以检查任务请求的状态或可以被发送给用户设备(诸如智能电话、智能手

表、膝上型计算机、个人家庭助理设备或其他电子设备)的状态。状态可以通过电子邮件、SMS或其他机制被发送。

[0168] 图9B示出了一旦预约已经被成功排程时图1B的理发预约请求的视觉状态。该状态可以被用户请求,或一旦预约已经被成功预订就可以在没有用户提示的情况下被发送至用户。

[0169] 图10A图示了口头状态请求和图1C的餐厅预约请求的更新。如图10A中所示,响应于用户询问餐厅预约是否已经做出,系统可以解释其已经采取来完成的任务的步骤,诸如呼叫餐厅两次。系统还会告诉用户下一次系统被排程以尝试呼叫的时间,并且可以向用户通知呼叫尝试之后的状态。

[0170] 图10B图示了在没有用户提示的情况下系统针对图1C的餐厅预约请求所提供的口头状态更新。一旦系统知道用户的任务已经完成,系统可以为用户提供状态更新。在一些实施方式中,系统立刻为用户提供状态更新。在其他实施方式中,系统确定方便的时间和方法来通知用户。例如,用户可以请求英国伦敦的晚餐预约。然而,用户可以当前被定位在美国加利福尼亚的山景城中。系统可以在用户正在睡觉的时候尝试呼叫餐厅。如果系统确认在伦敦中午12点的预约,系统可以确定在PDT上午4点发送状态更新文本消息可能使用户醒来。接着,系统可以选择备选状态更新方法,即,电子邮件,或保持在更方便的时间对用户进行状态更新。系统可以使用来自用户的日程、时区、习惯的信息或用户的其他个人信息确定适当的方便的时间和方式,来向用户提供状态更新。

[0171] 在一些实施方式中,系统可以使用用户信息来确定任务的紧迫性或是否重复努力以完成任务。例如,系统可以正在尝试为用户在加利福尼亚的山景城中的特定餐厅做出预约。用户到山景城的旅途可能到五月十五日结束。如果系统在五月十五日仍然没有成功,那么系统再继续请求五月十六日或者用户的旅行结束之后的预约就是不合理的。然而,为了得到餐厅的某人来做预约,那么与较早的几天相比在五月十四日呼叫两倍的次数是合理的。随着截止日期的接近任务可以变的更加紧急,并且随着截止日期的过去任务可以变的不那么紧急或过时的。

[0172] 在一些实施方式中,图1B的施援模块136确定当呼叫正在进行中时应该针对呼叫引入的干预的类型。施援模块136可以选择实时对机器人程序对话进行人工施援并解释其他人将会接手呼叫。在其他实施方式中,该模块可以允许人类操作者默默地接手呼叫。附加地或备选地,施援模块136可以在没有人工干预的情况下选择礼貌地结束机器人程序和人类之间的电话呼叫。

[0173] 图11示出了用于将电话呼叫从机器人程序转换到人类的示例过程1100。过程1100可以包括在电话呼叫的第一端上的第一人类和电话呼叫的第二端上的机器人程序之间的电话呼叫期间,由初始系统分析第一人类和机器人程序之间的实时对话(1102)。接着,呼叫初始系统可以基于对实时对话的分析来确定电话呼叫是否应该从机器人程序被转换到电话呼叫的第二端上的第二人类(1104)。响应于确定电话呼叫应该被转换到电话呼叫的第二端上的第二人类,由呼叫初始系统将电话呼叫从机器人程序转换到第二人类(1106)。

[0174] 为了确定对于特定的机器人程序电话呼叫最合适的干预的类型,施援模块136可以识别压力事件或寻找呼叫应该被终止或切换到人类操作者的其他指示。

[0175] 在一些实施方式中,施援模块136识别指示在人类上或机器人程序上的压力以适

当地响应于人类的问题。每次施援模块136识别压力事件,其都会增加存储的呼叫的本地压力和全局压力两者的水平。每当对话好像重回轨道的时候,施援模块136重置本地压力水平。例如,当机器人程序已经呼叫餐厅以为6个人的聚会做出预约,人类可以问机器人程序“*How many highchairs will your party need? (你的聚会需要多少高脚椅?)*”。机器人程序可以用“*we all need chairs. (我们都需要椅子)*”做出响应。人类基于机器人程序的响应带有些许恼怒的音调并回复“*Yes, I know you all need chairs, but how many high chairs for babies do you need? (是,我知道你们都需要椅子,但是你需要多少张用于婴儿的高脚椅?)*”。系统可以检测语调模式,即,人类陈述开始处、陈述结束处或贯穿陈述的高音。在一些实施方式中,语调模式预先与压力和恼怒相关联。系统可以将预先关联的模式与实时对话中所检测的模式匹配。在一些实施方式中,语调模式可以检测重复的单词、讲话有意放慢、或者关键字或短语(“*are you even listening to me? (你到底有没有在听我说?)*”“*am I talking to a robot? (我是在跟机器人说话么?)*”)。

[0176] 当系统检测到人类的有些许恼怒的音调时,其增加呼叫的本地压力水平。本地压力是反映了与当前状态相关联的可能的压力的运行分值。如果在实时对话内的人类陈述中出现任何压力指示符,则压力分值升高直到所述分值达到干预阈值。如果没有出现压力指示符,则系统可以根据 workflow 指示呼叫继续进行,并且本地压力分值降低或保持低(或零)。如果机器人程序通过提供人类所期望的响应来适当地对问题做出响应(诸如,“*We have no children in our party, (我们的聚会中没有孩子)*”),则系统可以降低本地压力。如果系统检测到人类做出响应时他或她的话音中不带有恼怒,则施援模块可以确定呼叫重回轨道并且将本地压力重置为默认值或重置为零。

[0177] 电话呼叫的全局压力只会累积升高。本地压力试图评估与人类的当前联络(correspondence)是否带有压力,而全局压力试图评定整个呼叫的总压力。例如,可以将阈值设置为在机器人程序施援转到人类操作员之前有三次误解。如果机器人程序连续三次没有理解人类,则本地压力将会为高并会使得机器人程序被施援。在不同的呼叫中,如果机器人程序连续两次没有理解对方,但确实理解了第三句,则本地压力将会在第三交互重置,并且很可能会继续该对话。全局压力将仍然保持指示机器人程序与人类之间存在过两次误解的信息。如果在该通话期间的稍后时间,机器人程序再次连续两次没有理解人类,则全局压力水平将达到阈值以上并且机器人程序很可能将被施援,尽管本地压力仍然在所设定的三次误解的阈值以下。

[0178] 如上所讨论,如果本地压力水平或全局压力水平中任一达到某阈值,则施援模块136将向系统100指示到了人工干预或礼貌地退出该呼叫。在一些示例中,任何时候当施援模块136需要重复自身、道歉、请求阐明等以及当人类纠正系统100或抱怨该呼叫(例如“*I can't hear you, can you hear me? (我听不到你,你能否听到我?)*”)时会将事件视为压力事件。

[0179] 在一些示例中,如果人类询问机器人程序是否是机器人、嘲弄机器人程序等——即通过询问荒谬的问题、或者以系统没有预期的某些其他方式行动(例如系统在试图作出餐厅预订时被问及体育赛事),则施援模块136会将事件视为压力事件。

[0180] 在一些实施方式中,施援模块136是基于特征的规则集合,其确定系统何时应当施援转到人工干预。一个基于特征的规则可以是陈述当两个连续的未知输入意图出现时系统

应当施援的规则。一个不同的规则会陈述当在通话期间的任何处出现四个未知输入意图时系统施援转到人工操作员。系统跟踪对话中出现的事件并且确定是否已经发生满足规则的判据的事件。

[0181] 在其他实施方式中,施援模块136使用机器学习来预测何时自动地施援转到人类操作员。例如,施援模块136能够从与人类的对话中接收意图作为一个或多个机器学习模型的输入。机器学习模型能够基于所接收的意图和历史意图与结果来确定是否要施援转到人类操作员。系统能够关于来自标示施援何时应当出现或不出现的注释记录的特征来训练机器学习模型。然后,机器学习模型能够在给出输入特征集合的情况下预测施援何时很可能应当出现。

[0182] 施援模块136使用许多因素来确定施援,所述因素包括:人类的举止、人类的语气、人类的所确定的恼怒水平、人类使用的语言、或者人类的词语选择。

[0183] 系统100能够将正由机器人程序施行的对话逐步提升至人类操作员来处理。例如,如果在特定对话中存在阈值量的压力,则施援模块136能够向流管理器132提供反馈数据。流管理器132可以在有或没有向人类被呼叫者可听地警示的情况下指令机器人程序将通话递交至通过操作员控制器134提供输入的人类操作员。例如,机器人程序能够说“Sure, thank you for your time today. Here is my supervisor. (当然,感谢您今日来电,这是我的主管。)”然后,人类操作员能够通过操作员控制器134完成机器人程序试图执行的任务。

[0184] 施援模块136也能够确定置信度水平,其限定了系统在正被完成的当前任务中具有的执行度。例如,机器人程序可能接到任务来为用户作出用餐预订。如果机器人程序呼叫餐厅并且人类询问了该机器人程序不知道答案的多个问题,则系统在正在完成的当前任务中可以具有低置信度。在系统接收到该系统不具有答案的问题之后,系统在完成该任务中的置信度水平可能变得更低。如果系统复原并且系统确定对话正向着完成任务前进,则系统可以升高置信度水平。

[0185] 在一些实施方式中,系统将电话对话切换至监视呼叫的人类操作员。系统可以使用操作员用户界面或一些其他通知机制来向操作员警示需要转接电话呼叫。一旦被通知,操作员在系统确定终止该呼叫之前可以具有有限的时间来转接该电话呼叫。系统可以使用与操作员相同的话音。在这样的情况下,从机器人程序向操作员的转接对于另一侧能够是透明的,因为话音保持相同。

[0186] 在另一实施方式中,系统将电话对话切换至请求该任务的人类用户。系统能够向该用户警示进行中对电话呼叫。当完成该任务存在问题时或当机器人程序被询问该机器人程序不知道答案的问题时,系统能够使该用户知道。机器人程序可以以文本、电子邮件、或以一些其他方式来传递机器人程序需要用户输入的对话的详情。在一些实施方式中,机器人程序将等待阈值时间量——即5秒以供用户在无需用户输入继续该对话之前进行响应。由于该对话正在实时发生的,因此机器人程序无法长时间段来等待用户响应。在一些实施方式中,当系统已经确定电话呼叫需要从机器人程序转接走时,系统可以试图将电话呼叫转接至做出请求的用户。如上所讨论,系统可以等待阈值时间量以供用户进行响应并接手电话呼叫。在一些实施方式中,如果用户没有在阈值时间量中接手电话呼叫,则系统将向操作员转接该电话呼叫。在其他实例中,系统将结束该电话对话。系统也可以使用与人类用户

相同的话音,使得对于对话的另一侧从机器人程序向用户的转接是无缝的。

[0187] 图12示出了计算设备1200的示例和移动计算设备1250的示例,其能够用于实现上述技术。计算设备1200意图表示各种形式的数字计算机,诸如膝上型计算机、桌面计算机、工作站、个人数字助理、服务器、刀片服务器、大型主机、和其他适当的计算机。计算设备1250意图表示各种形式的移动设备,诸如个人数字助理、蜂窝电话、智能电话、和其他类似的计算设备。这里所示的组件、它们的连接和关系以及它们的功能仅仅意味着是示例性的,并不意味着限制本文中描述和/或要求保护的发明的实施方式。

[0188] 计算设备1200包括处理器1202、存储器1204、存储设备1206、连接到存储器1204和高速扩展端口1210的高速接口1208、以及连接到低速总线1214和存储设备1206的低速接口1212。处理器1202、存储器1204、存储设备1206、高速接口1208、高速扩展端口1210、和低速接口1212中的每一个使用各种总线互连,并且可以安装在公共主板上或视情况以其他方式安装。

[0189] 处理器1202可以处理在计算设备1200内执行的指令,包括存储在存储器1204中或存储设备1206上的指令,以在诸如耦合到高速接口1208的显示器1216的外部输入/输出设备上显示GUI的图形信息。在其他实施方式中,可以视情况连同多个存储器和多种类型的存储器一起使用多个处理器和/或多条总线。此外,可以连接多个计算设备1200,其中每个设备提供必要操作的一部分(例如,作为服务器组、刀片服务器的群组、或多处理器系统)。

[0190] 存储器1204在计算设备1200内存储信息。在一个实施方式中,存储器1204是一个或多个易失性存储器单元。在另一实施方式中,存储器1204是一个或多个非易失性存储器单元。存储器1204还可以是另一种形式的计算机可读介质,诸如磁盘或光盘。

[0191] 存储设备1206能够为计算设备1200提供大容量存储。在一个实施方式中,存储设备1206可以是或包含计算机可读介质,诸如软盘设备、硬盘设备、光盘设备、或磁带设备;闪存或其它类似的固态存储设备;或包括存储区域网络或其他配置中的设备的设备阵列。计算机程序产品可以有形地体现在信息载体中。计算机程序产品还可以包含指令,该指令在被执行时执行一个或多个方法,诸如上述的方法。计算机程序产品也能够有形地体现在计算机可读介质或机器可读介质中,诸如存储器1204、存储设备1206、或处理器1202上的存储器。

[0192] 高速接口1208管理计算设备1200的带宽密集型操作,而低速接口1212管理较低带宽密集型操作。功能的这样的分配仅是示例性的。在一些实施方式中,高速接口1208耦合到存储器1204、显示器1216(例如,通过图形处理器或加速器)以及可接受各种扩展卡(未示出)的高速扩展端口1210。在该实施方式中,低速接口1212耦合到存储设备1206和低速扩展端口1214。可以包括各种通信端口(例如,USB、蓝牙、以太网、无线以太网)的低速扩展端口1214可以耦合至一个或多个输入/输出设备——诸如键盘、指示设备、扫描仪、或例如通过网络适配器而耦合到诸如交换机或路由器的联网设备。

[0193] 计算设备1200可以以多种不同的形式来实现,如图所示。例如,它可以被实现为标准服务器,或者在这样的服务器的群组中多次实现。此外,其可以在诸如膝上型计算机1222的个人计算机中实现。它也可以被实现为机架式服务器系统1224的一部分。替选地,来自计算设备1200的组件可以与诸如移动计算设备1250的计算设备中的其他组件(未示出)组合。这样的设备中的每一个可以包含计算设备1200、移动计算设备1250中的一个或多个,并且

整个系统可以由彼此通信的多个计算设备组成。

[0194] 除了其他组件之外,移动计算设备1250包括处理器1252、存储器1264,诸如显示器1254、通信接口1266、和收发器1268的输入/输出设备。设备1250还可以设置有诸如微驱动器或其他设备的存储设备1250,以提供额外的存储。处理器1252、存储器1264、显示器1254、通信接口1266、和收发器1268中的每一个使用各种总线互连,并且若干组件可以安装在公共主板上或视情况以其他方式安装。

[0195] 处理器1252可以执行移动计算设备1250内的指令,包括存储在存储器1264中的指令。处理器可以被实现为包括分离的和多个模拟和数字处理器的芯片的芯片集。处理器可以例如提供移动计算设备1250的其他组件的协调,诸如对用户界面、由设备1250运行的应用、以及由设备1250进行的无线通信的控制。

[0196] 处理器1252可以通过耦合到显示器1254的控制接口1258和显示接口1256与用户通信。显示器1254可以是例如TFT(薄膜晶体管液晶显示器)显示器或OLED(有机发光二极管)显示器或其他适当的显示技术。显示接口1256可以包括用于驱动显示器1254以向用户呈现图形和其它信息的适当电路。控制接口1258可以接收来自用户的命令并将其变换以用于提交给处理器1252。此外,外部接口1262可以提供与处理器1252的通信,以便使得移动计算设备1250能够与其他设备进行近区域通信。外部接口1262可以例如在一些实施方式中提供有线通信,或在其他实施方式中提供无线通信,并且还可以使用多个接口。

[0197] 存储器1264将信息存储在移动计算设备1250内。存储器1264可以被实现为计算机可读介质或多个介质、易失性存储器单元、或非易失性存储器单元或多个单元中的一个或多个。还可以提供扩展存储器1274并通过扩展接口1272来将其连接到设备1250,扩展接口1272可以包括例如SIMM(单列直插存储器模块)卡接口。扩展存储器1274可以为移动计算设备1250提供额外的存储空间,或者还可以存储移动计算设备1250的应用或其他信息。具体地,扩展存储器1274可以包括用于执行或补充上述过程的指令,并且还可以包括安全信息。因此,例如,扩展存储器1274可以被提供作为移动计算设备1250的安全模块,并且可以用允许安全使用移动计算设备1250的指令进行编程。另外,安全应用可以经由SIMM卡连同附加信息一起被提供,诸如以不可非法侵入方式将识别信息放置在SIMM卡上。

[0198] 存储器可以包括例如闪存和/或NVRAM存储器(非易失性随机存取存储器),如下所述。在一些实施方式中,计算机程序产品被有形地体现在信息载体中。计算机程序产品包含指令,其当被执行时,执行一个或多个方法——例如上述的那些方法。计算机程序产品可以是计算机可读介质或机器可读介质,诸如存储器1264、扩展存储器1274、或处理器1252上的存储器。在一些实施方式中,计算机程序产品能够以传播信号来接收,例如通过收发器1268或外部接口1262来接收。

[0199] 设备1250可以通过通信接口1266进行无线通信,通信接口1266可以在必要时包括数字信号处理电路。通信接口1266可以提供在诸如GSM语音呼叫(全球移动通信系统)、SMS(短消息服务)、EMS(增强消息传递服务)、或MMS消息传递(多媒体消息传递服务)、CDMA(码分多址)、TDMA(时分多址)、PDC(个人数字蜂窝)、WCDMA(宽带码分多址)、CDMA2000、或GPRS(通用分组无线电服务)等的各种模式或协议下的通信。这样的通信可以例如通过使用射频的收发器1268发生。另外,可能会诸如使用蓝牙、WiFi、或其他这样的收发器(未示出)发生短距离通信。此外,GPS(全球定位系统)接收器模块2570可以向移动计算设备1250提供附加

的导航和位置相关的无线数据,其视情况可以由在移动计算设备1250上运行的应用来使用。

[0200] 移动计算设备1250还可以使用音频编解码器1260可听地通信,音频编解码器1260可以从用户接收说出的信息并将其变换为可用的数字信息。音频编解码器1260可以同样例如在移动计算设备1250的送受话器中诸如通过扬声器为用户产生可听见的声音。这样的声音可以包括来自语音电话呼叫的声音,可以包括记录的声音(例如,语音消息、音乐文件等等),并且还可以包括在移动计算设备1250上运行的应用产生的声音。

[0201] 如图所示,移动计算设备1250可以以多种不同的形式来实现。例如,它可以被实现为蜂窝电话1280。它也可以被实现为智能电话1282、个人数字助理、平板计算机、可穿戴计算机、或其他类似移动设备的一部分。

[0202] 已经描述了多种实施方式。然而,将理解,在不背离本公开的宗旨和范围的情况下可以进行各种修改。例如,上文示出的各种形式的流程可以以步骤重新排序、添加或者移除来使用。

[0203] 在本说明书中描述的所有功能操作可以在数字电子电路或者计算机软件、固件或者硬件中实施,包括在本说明书中公开的结构以及它们的在结构上的等同物,或者以它们的一个或多个的组合来实施。公开的技术可以被实施为一个或多个计算机程序产品,即,在计算机可读介质上编码的一个或多个计算机程序指令的模块,用于通过数据处理装置执行,或者控制数据处理装置的操作。计算机可读机制可以是机器可读存储设备、机器可读存储基板、存储设备、影响机器可读传播信号的事项的组合,或者它们中的一个或多个的组合。计算机可读介质可以是非暂时性计算机可读介质。术语“数据处理装置”包括所有用于处理数据的装置、设备和机器,包括,例如,可编程处理器、计算机,或者多个处理器或计算机。除了硬件之外,装置还可以包括代码,该代码创建用于讨论中的计算机程序的执行环境,例如,构成处理架构、协议栈、数据库管理系统、操作系统、或者它们的一个或多个的组合的代码。传播信号是人工生成的信号,例如,机器生成的电、光或者电磁信号,被生成以编码信息用于发送到适当的接收器装置

[0204] 计算机程序(也可被称为程序、软件、软件应用、模块、脚本或代码)可以以任何形式的编程语言编写,包括编译或解释语言或声明性或过程性语言,并且其可以以任何形式部署,包括作为独立程序或作为适于在计算环境中使用的模块、组件、子例程或其他单元。计算机程序可以但不需要对应于文件系统中的文件。程序可以存储在保存其他程序或数据的文件的一部分中(例如存储在标记语言文档中的一个或多个脚本)、存储在专用于所述程序的单个文件中,或者存储在多个协同文件(例如存储一个或多个模块、子程序或代码部分的文件)。计算机程序可以被部署为在一个计算机上或在位于一个地点或分布在多个地点并通过通信网络互连的多个计算机上执行。

[0205] 本说明书中描述的过程和逻辑流程可以由执行一个或多个计算机程序的一个或多个可编程计算机执行,以通过对输入数据进行操作并生成输出来执行功能。过程和逻辑流程也可以由例如FPGA(现场可编程门阵列)或ASIC(专用集成电路)的专用逻辑电路来执行,并且装置也可以实现为该专用逻辑电路。

[0206] 适合于执行计算机程序的计算机包括,例如可以基于通用或专用微处理器或两者或任何其它类型的中央处理单元。通常,中央处理单元将从只读存储器或随机存取存储器

或两者接收指令和数据。计算机的基本元件是用于执行或实施指令的中央处理单元和用于存储指令和数据的一个或多个存储器设备。通常,计算机还将包括用于存储数据的一个或多个大容量存储设备,例如磁盘、磁光盘或光盘,或者可操作地耦合以从其接收数据或向其传送数据或两者。然而,计算机不需要具有这样的设备。此外,计算机可以嵌入在另一设备中,例如移动电话、个人数字助理(PDA)、移动音频或视频播放器、游戏控制台、全球定位系统(GPS)接收器或便携式存储设备,例如,通用串行总线(USB)闪存驱动器,这里仅举几个例子。适于存储计算机程序指令和数据的计算机可读介质包括所有形式的非易失性存储器、介质和存储器设备,包括例如半导体存储器设备,例如EPROM、EEPROM和闪存设备;磁盘,例如内部硬盘或可移动盘;磁光盘;和CD-ROM和DVD-ROM盘。处理器和存储器可以由专用逻辑电路补充或并入专用逻辑电路中。

[0207] 为了提供与用户的交互,本说明书中描述的主题的实施例可以实现在计算机上,该计算机具有:用于向用户显示信息的显示设备(例如,CRT(阴极射线管)或LCD(液晶显示器)监视器)和键盘;以及,指示设备,例如鼠标或轨迹球,用户可以通过该指示设备向计算机提供输入。其他类型的设备也可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的感觉反馈,例如视觉反馈、听觉反馈或触觉反馈;并且可以以任何形式接收来自用户的输入,该任何形式包括声音、语音或触觉输入。

[0208] 在本说明书中描述的主题的实施例可以实现在计算系统中,该计算系统包括诸如作为数据服务器的后端组件,或者包括诸如应用服务器的中间件组件,或者包括诸如具有图形用户界面或Web浏览器的客户端计算机的前端组件,或者包括一个或多个这样的后端、中间件或前端组件的任何组合,用户可以通过该Web浏览器与本说明书中描述的主题的实现交互。系统的组件可以通过任何形式或介质的数字数据通信(例如,通信网络)互连。通信网络的示例包括局域网(“LAN”)和广域网(“WAN”),例如互联网。

[0209] 计算系统可以包括客户端和服务器。客户端和服务器通常彼此远离并且通常通过通信网络交互。客户端和服务器的关系借助于在相应计算机上运行并且彼此具有客户端-服务器关系的计算机程序而产生。

[0210] 虽然本说明书包含许多具体实施细节,但是这些不应被解释为对任何发明的范围或对所要求保护内容的范围的限制,而是作为可以对特定发明的特定实施例特定的特征的描述。在本说明书中在单独实施例的场境中描述的某些特征也可以在单个实施例中组合实现。相反,在单个实施例的场境中描述的各种特征也可以在多个实施例中单独地或以任何合适的子组合来实现。此外,虽然特征可以在上面描述为在某些组合中起作用并且甚至最初如此要求保护,但是来自所要求保护的组合的一个或多个特征在一些情况下可以从组合中删除,并且所要求保护的组合可以涉及子组合或子组合的变体。

[0211] 类似地,虽然在附图中以特定顺序描绘了操作,但是这不应被理解为要求这些操作以所示的特定顺序或以顺序次序执行,或者所有所示的操作被执行,以实现期望的结果。在某些情况下,多任务和并行处理可能是有利的。此外,上述实施例中的各种系统模块和组件的分离不应被理解为在所有实施例中都需要这样的分离,并且应当理解,所描述的程序组件和系统通常可以一起集成在单个软件产品中,或者封装成多个软件产品。

[0212] 由此,已经描述了主题的特定实施方式。其他实施方式在所附权利要求的范围内。例如,权利要求中所述的动作可以以不同的顺序执行并且仍然实现期望的结果。

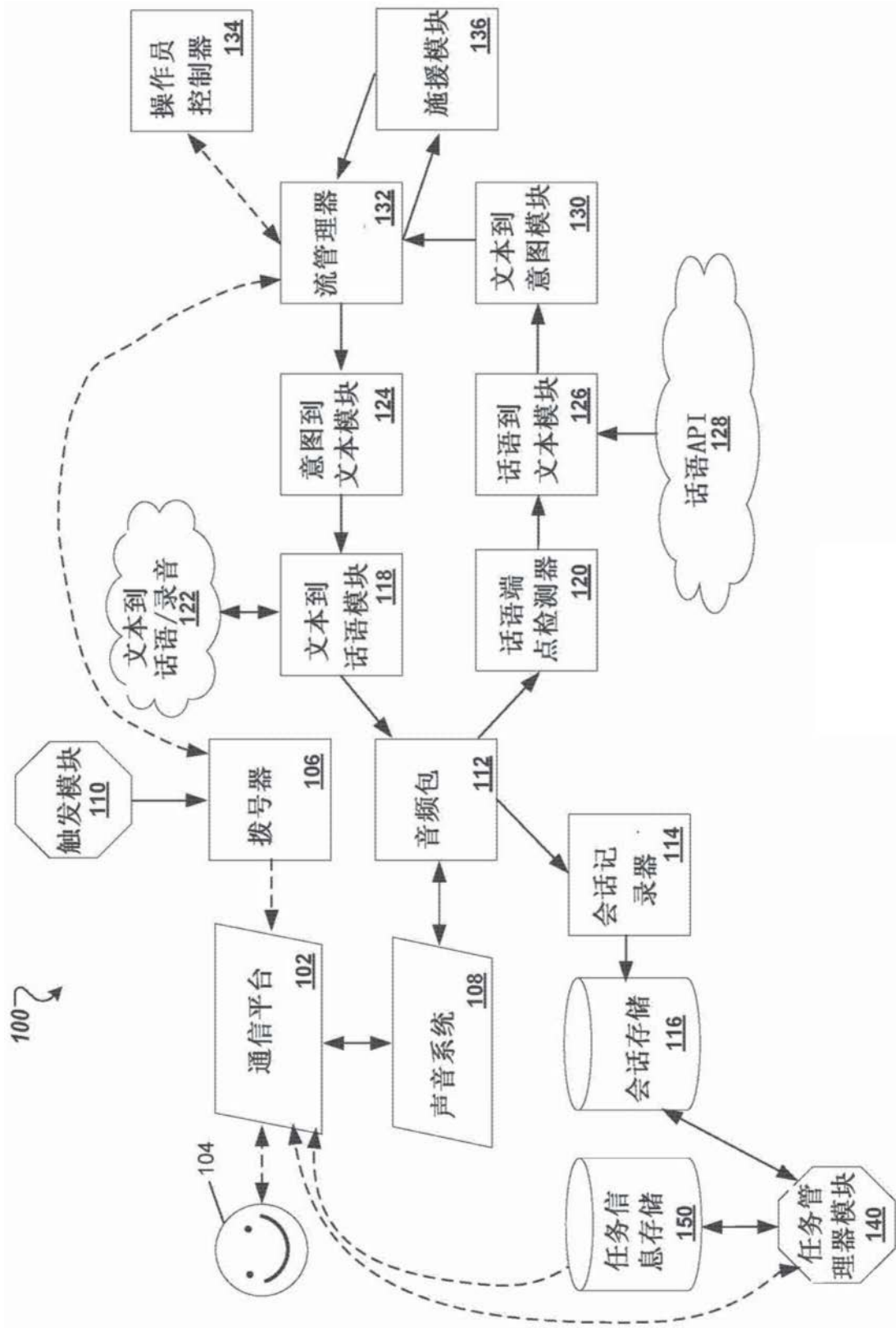


图1A

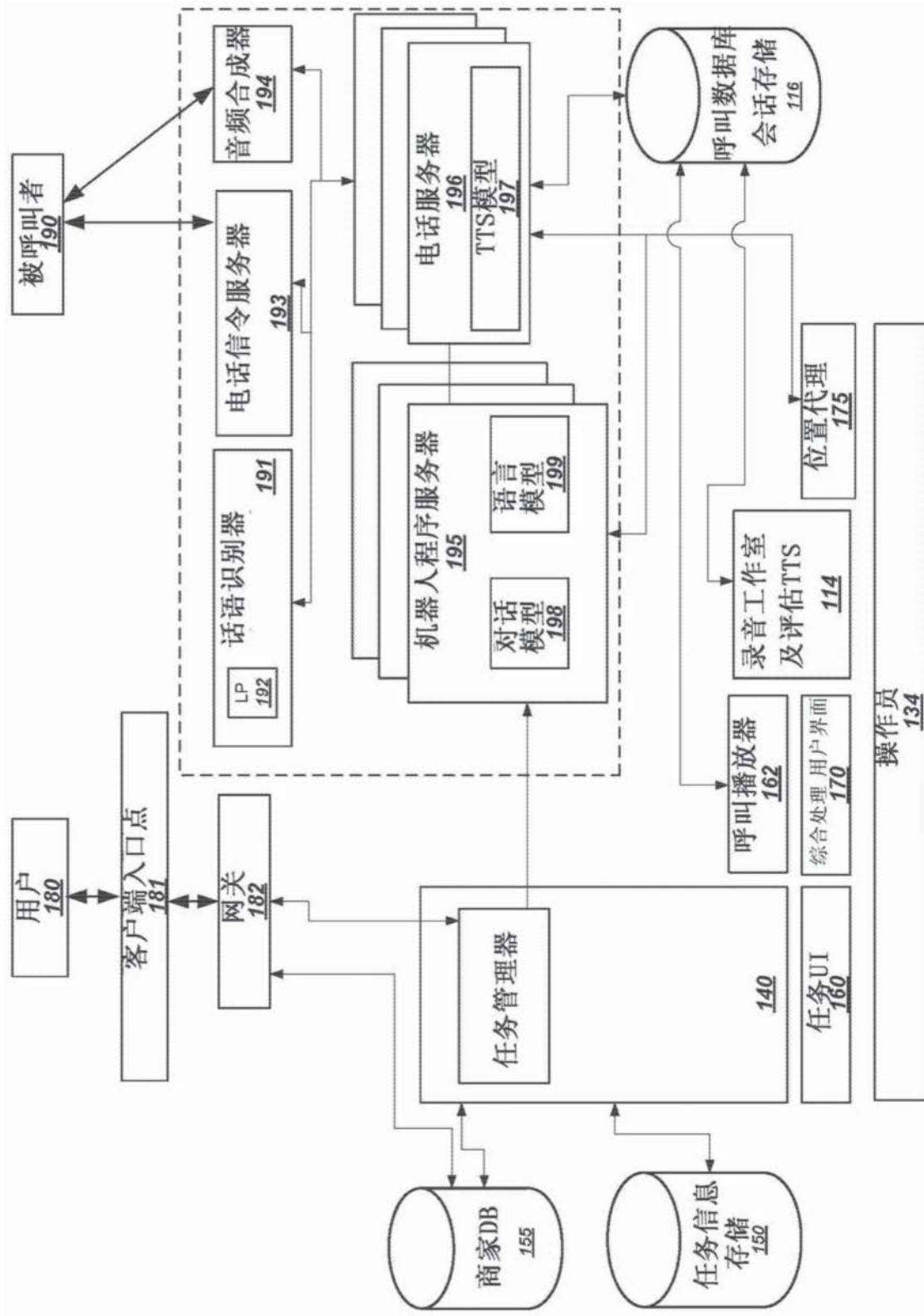


图1B

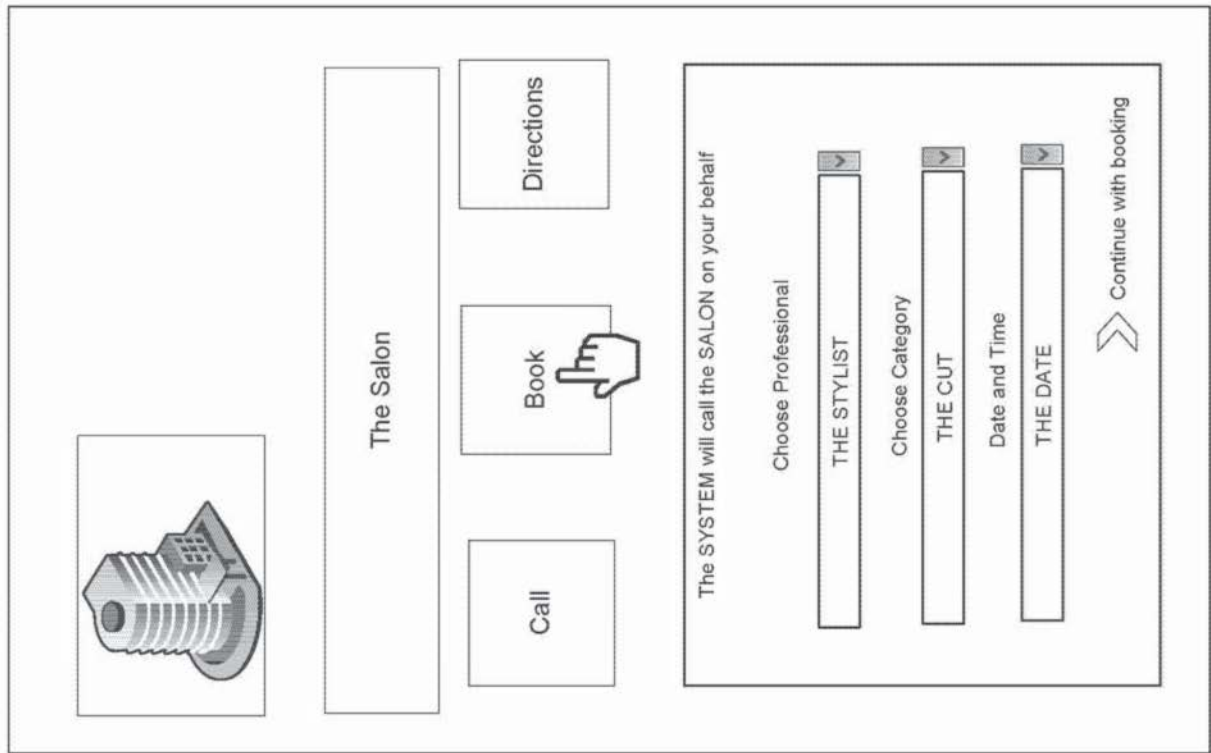


图1C

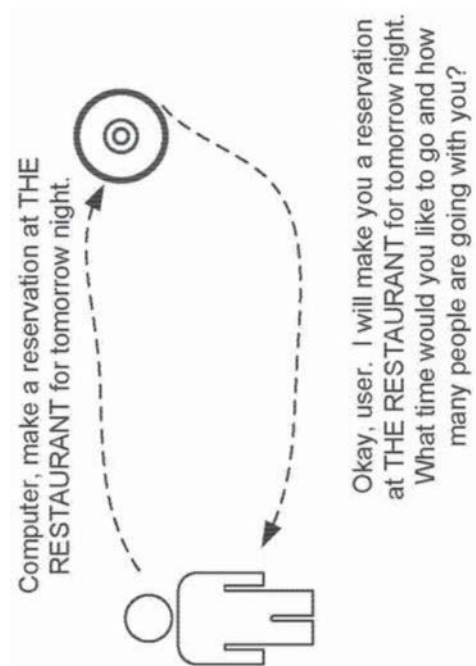


图1D

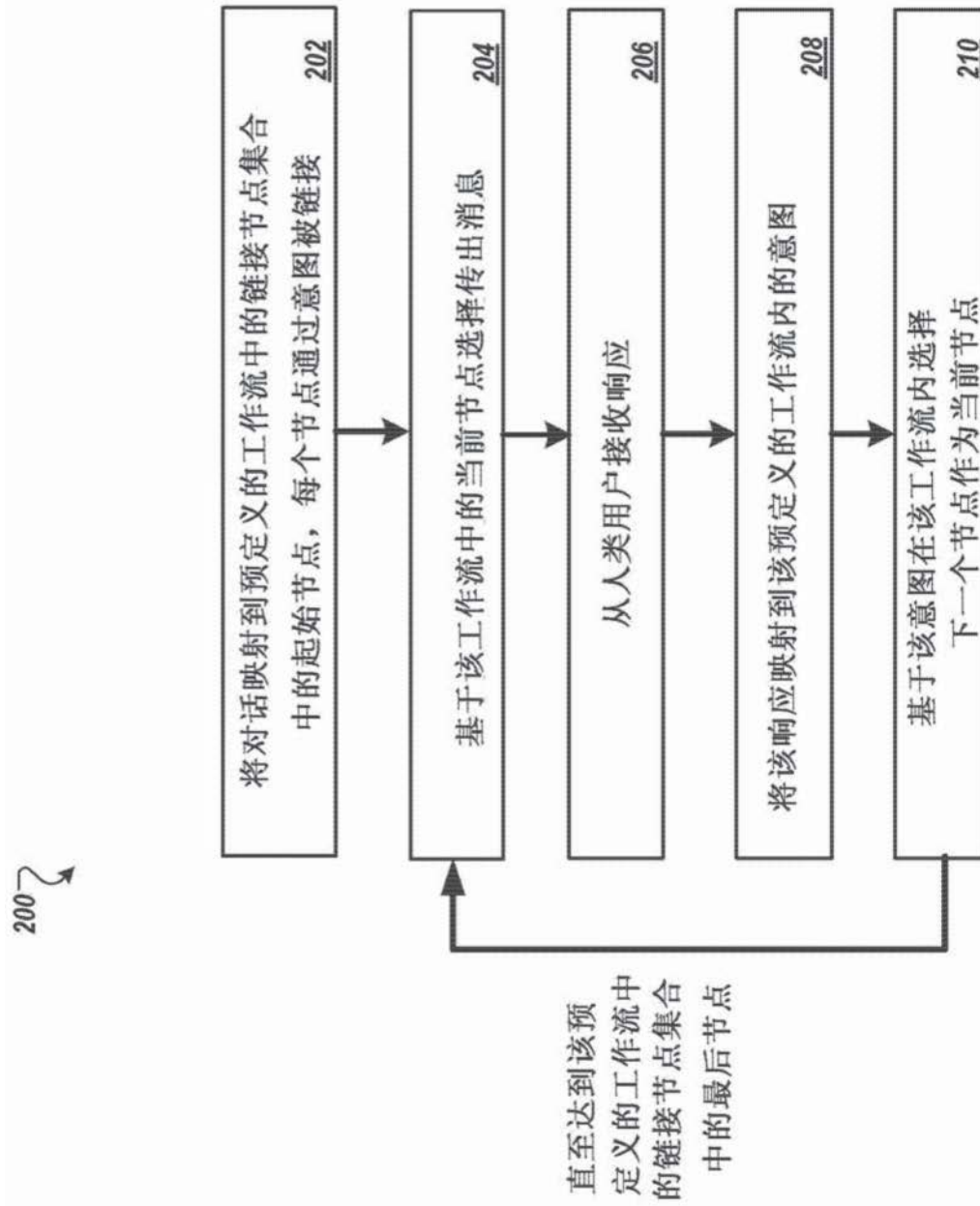


图2A

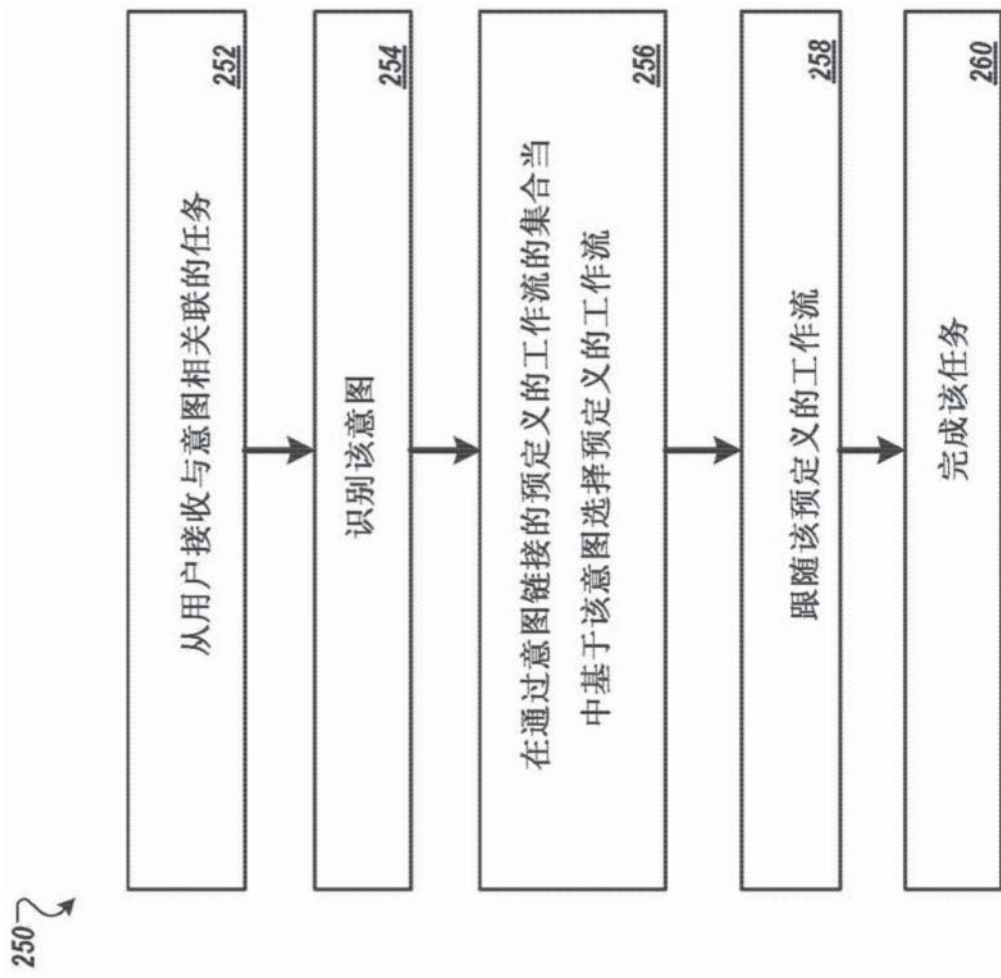


图2B

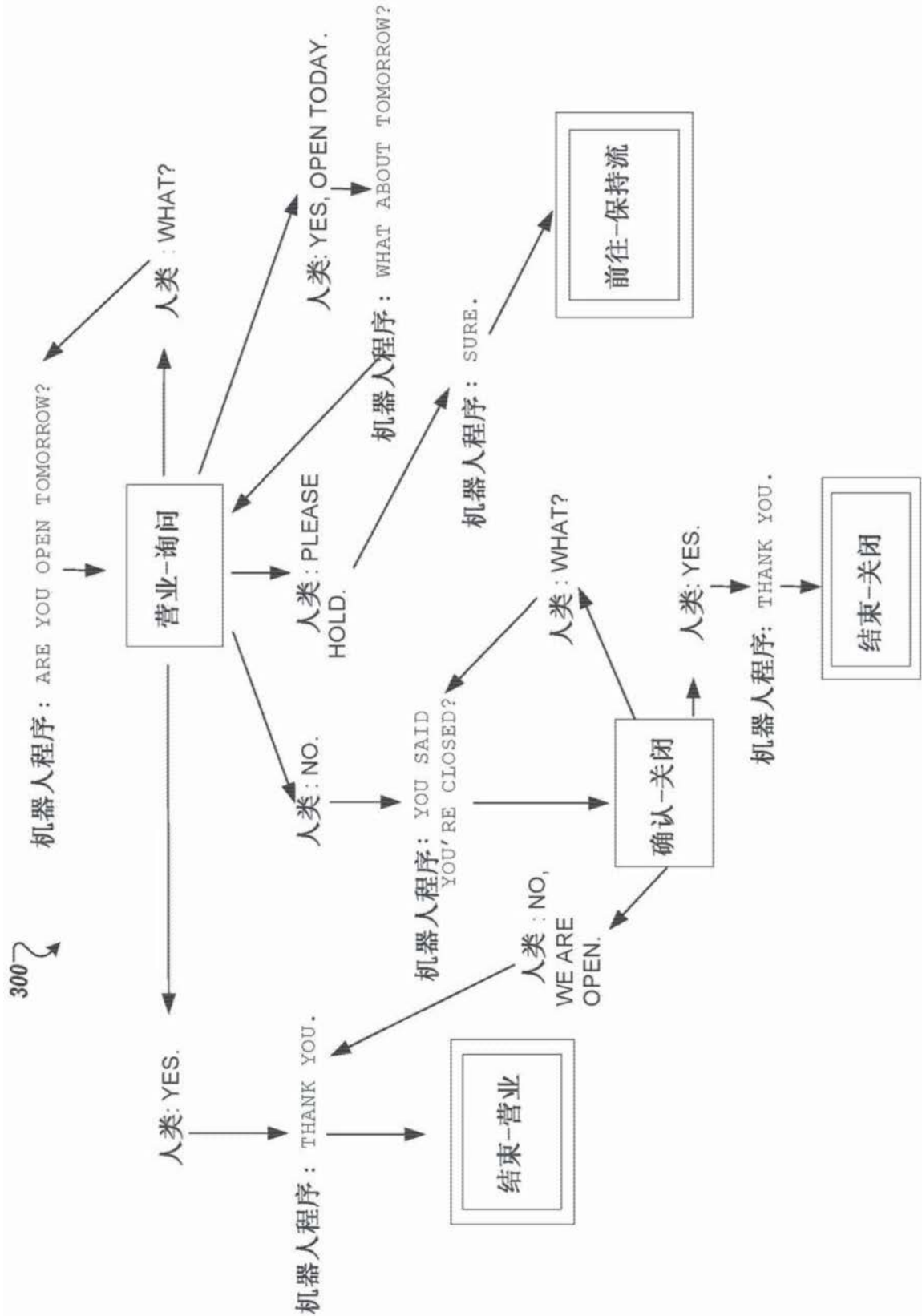


图3

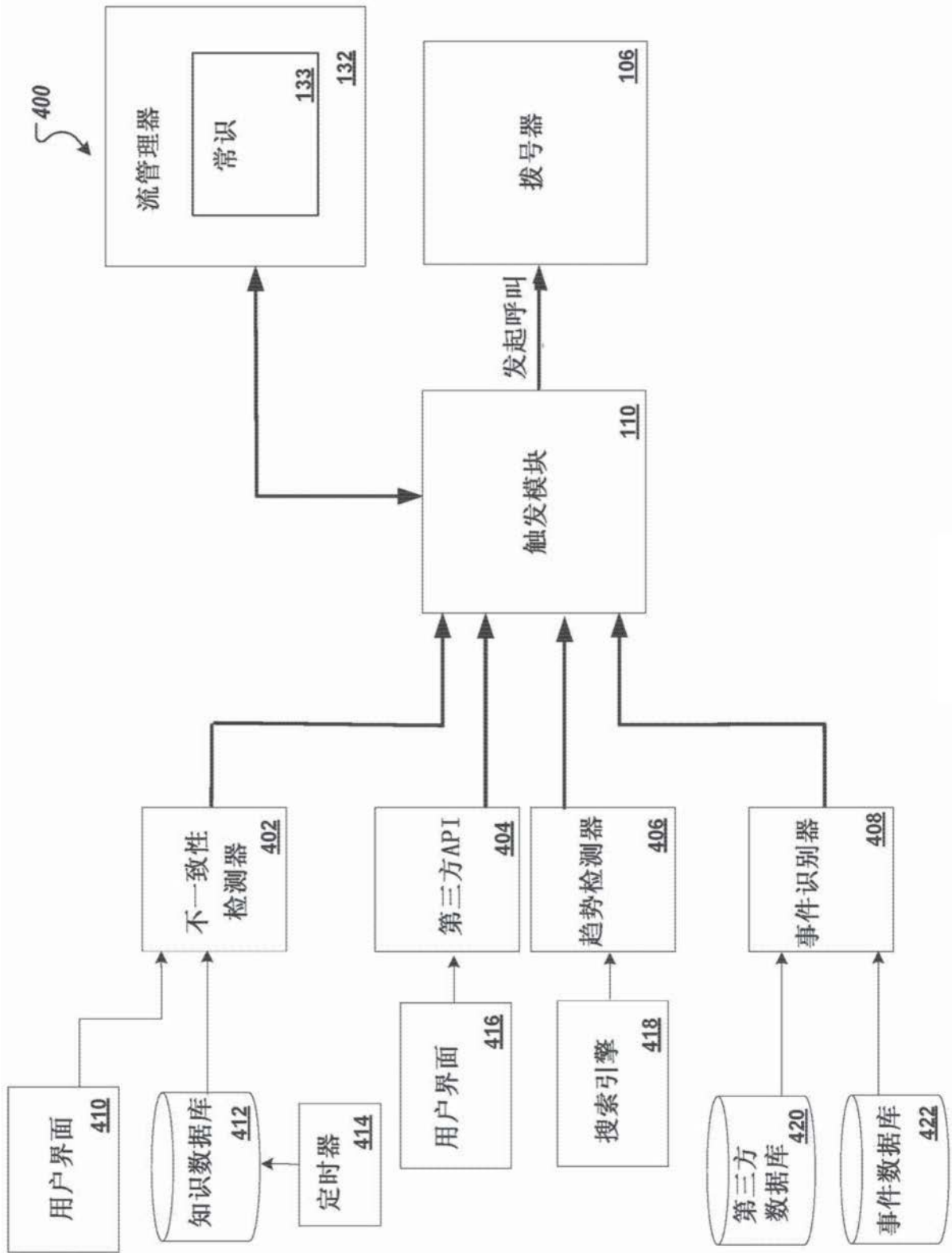


图4

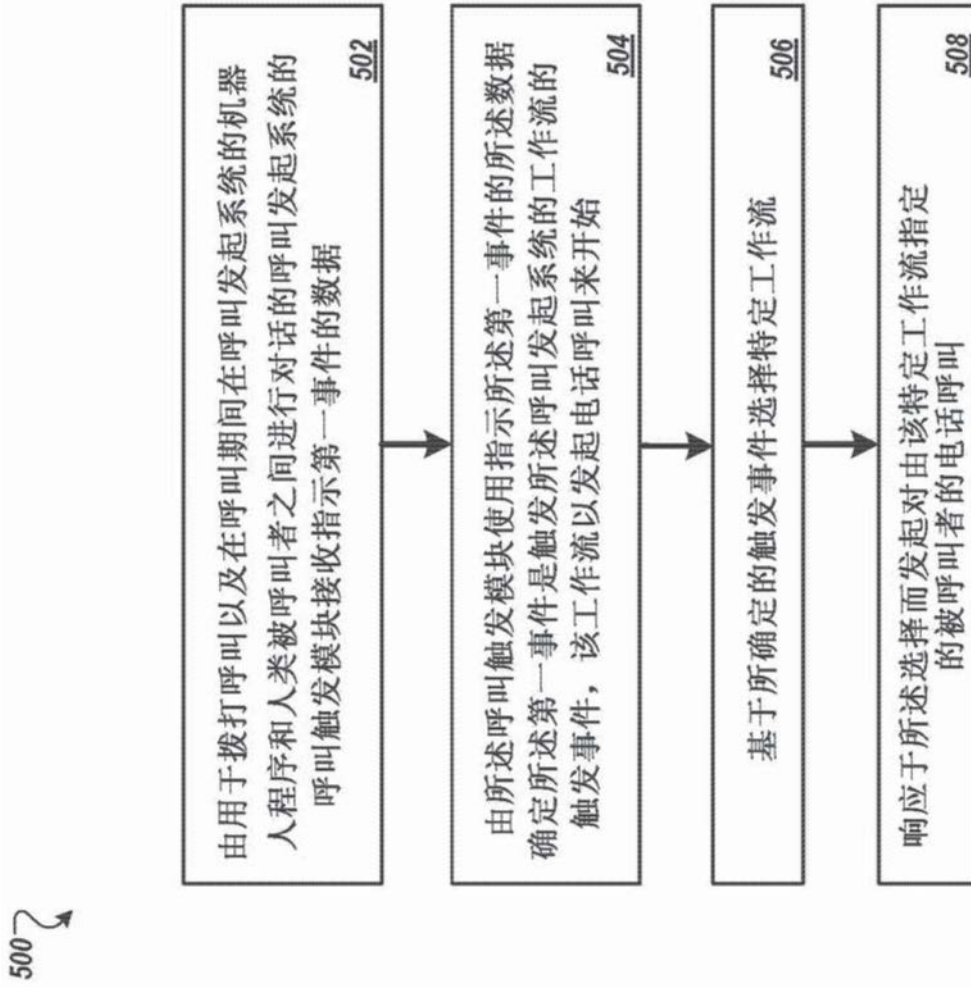


图5

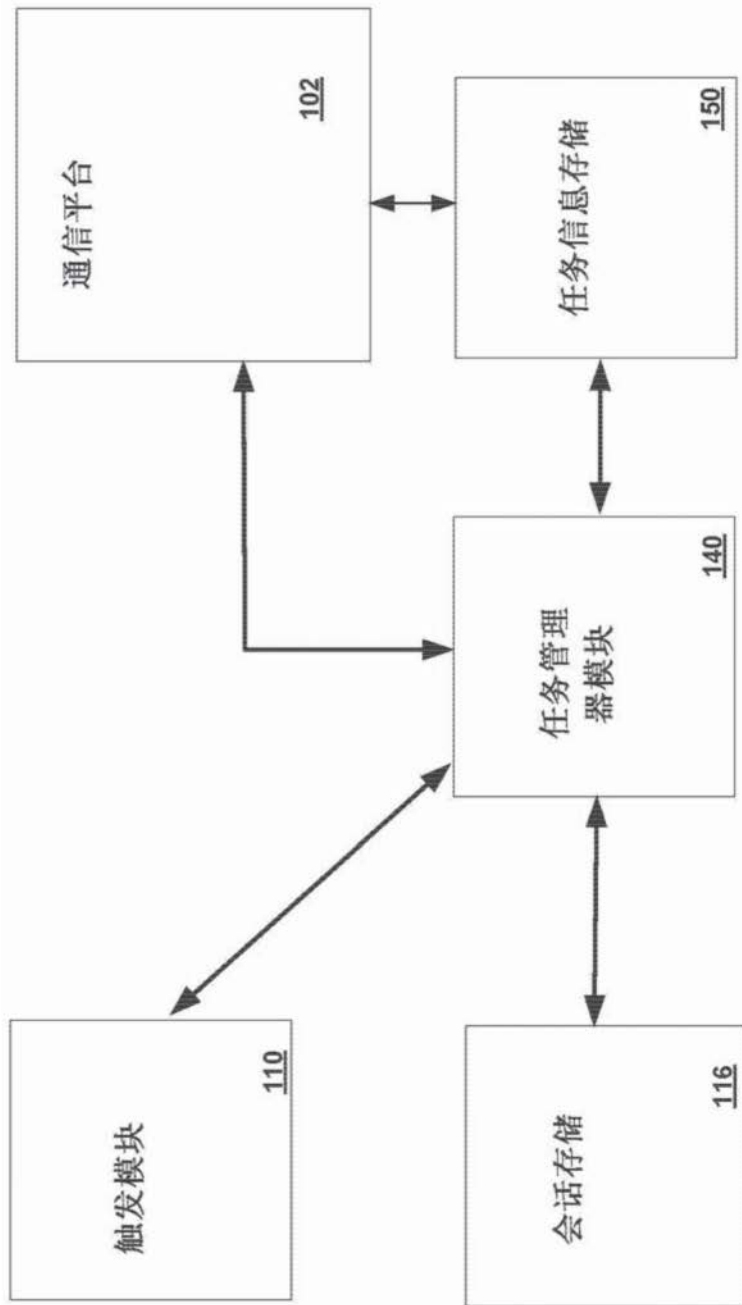


图6

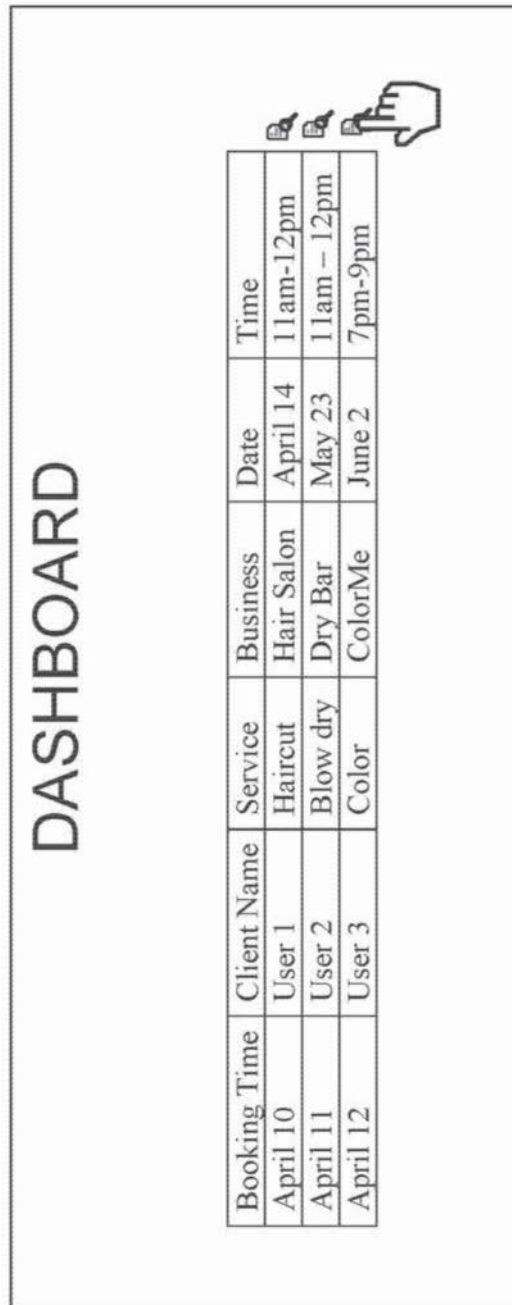


图7A



图7B

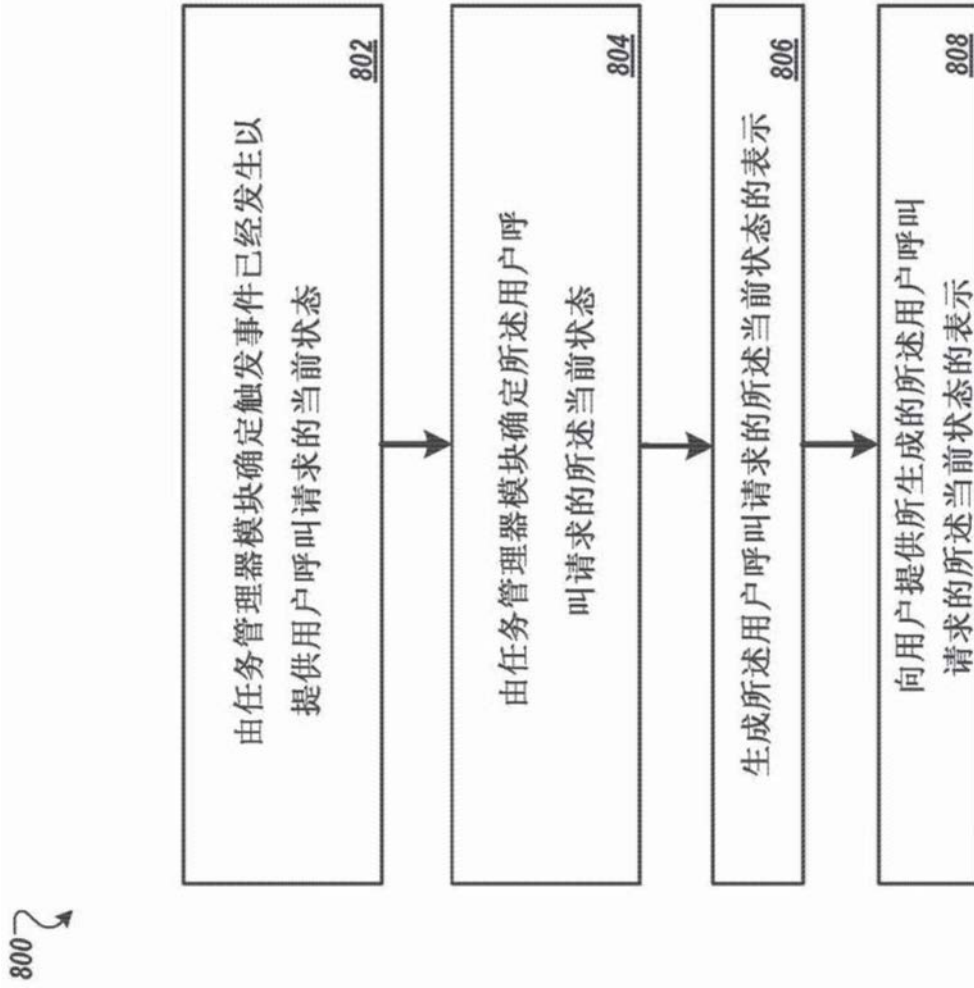


图8

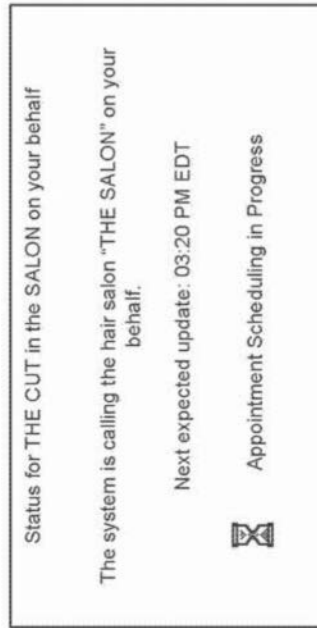


图9A



图9B

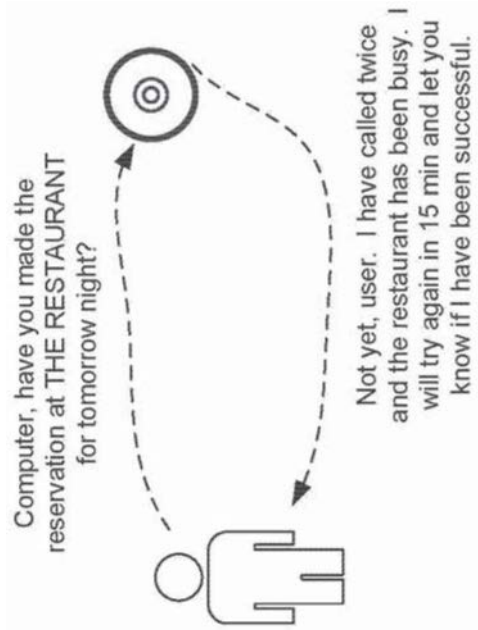


图10A

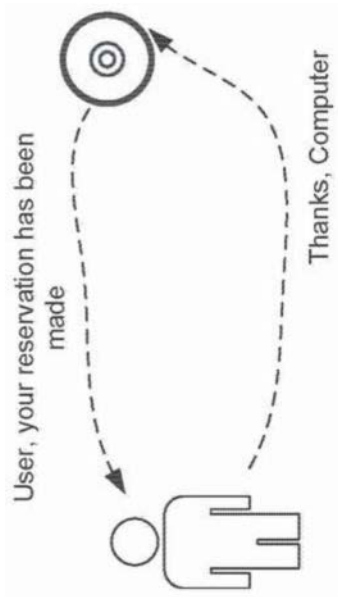


图10B

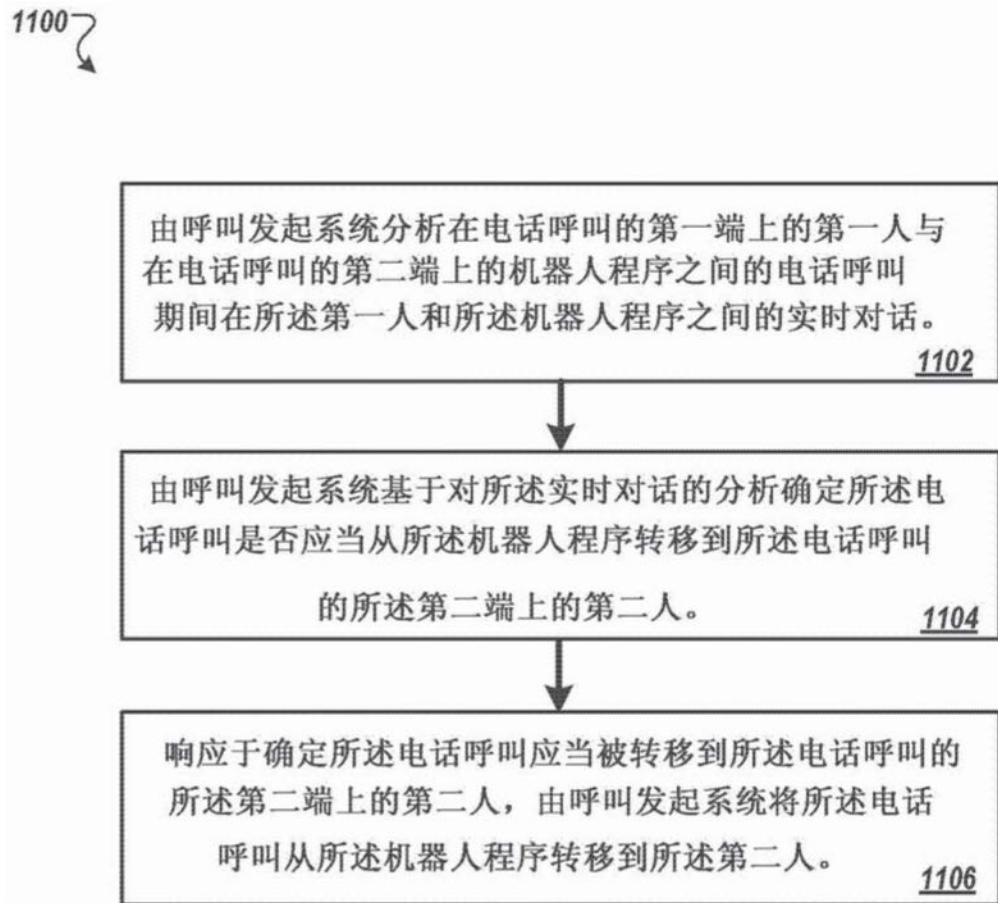


图11

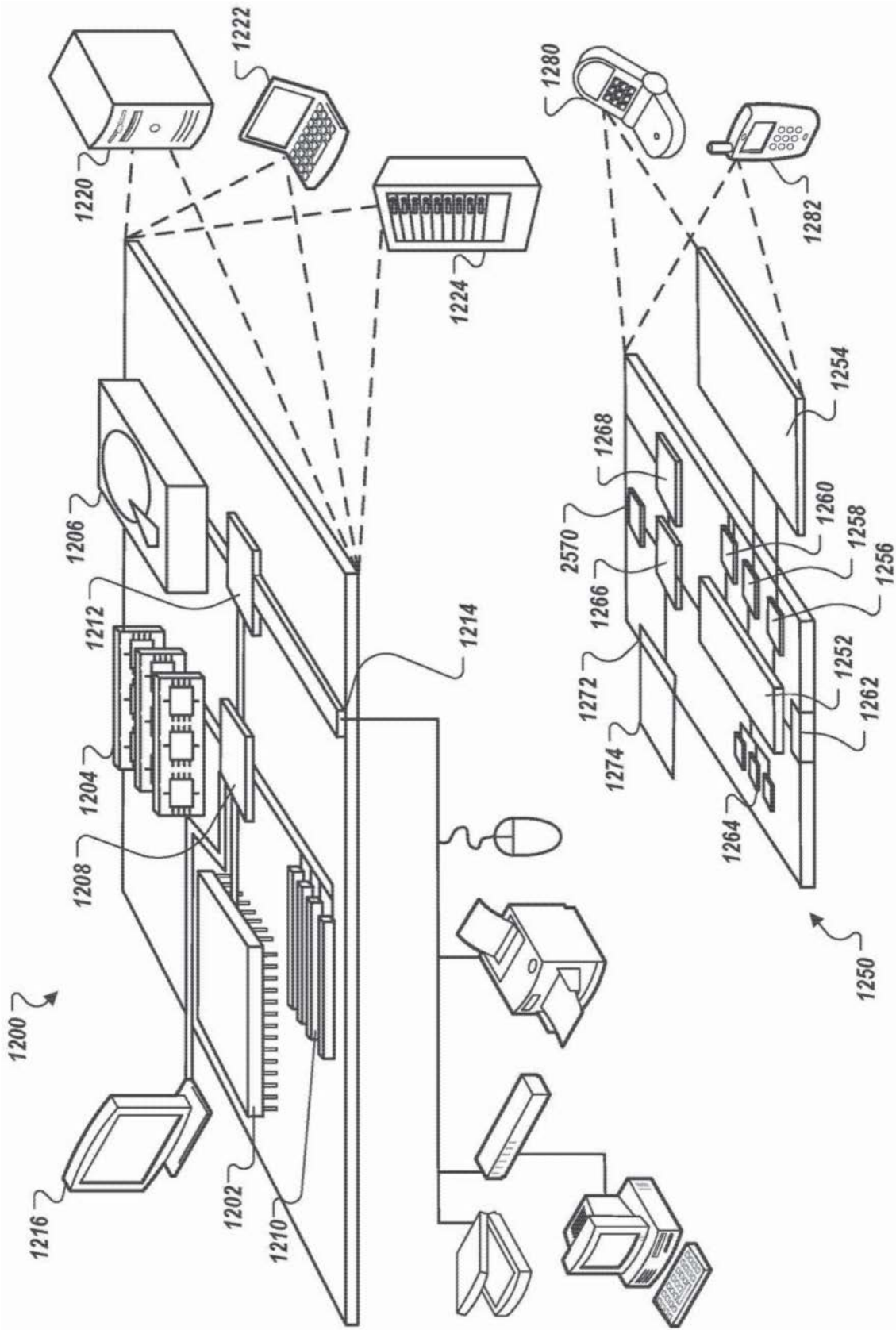


图12