



(12) 发明专利申请

(10) 申请公布号 CN 102369568 A

(43) 申请公布日 2012. 03. 07

(21) 申请号 201080014750. 4

(51) Int. Cl.

(22) 申请日 2010. 01. 19

G10L 15/18(2006. 01)

(30) 优先权数据

G10L 15/22(2006. 01)

12/364, 927 2009. 02. 03 US

H04M 3/493(2006. 01)

(85) PCT申请进入国家阶段日

2011. 09. 29

(86) PCT申请的申请数据

PCT/IB2010/000186 2010. 01. 19

(87) PCT申请的公布数据

W02010/089645 EN 2010. 08. 12

(71) 申请人 索夫特赫斯公司

地址 波兰格丁尼亚

(72) 发明人 E·乌拉斯由克

(74) 专利代理机构 北京嘉和天工知识产权代理

事务所 11269

代理人 严慎

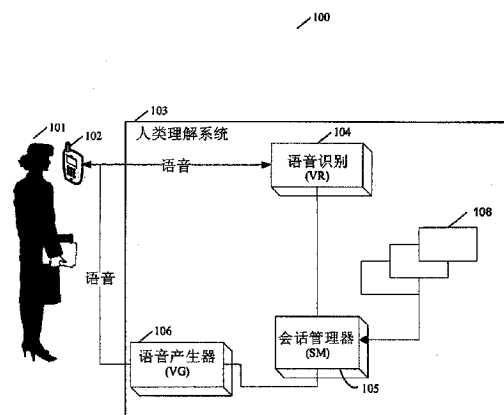
权利要求书 2 页 说明书 8 页 附图 5 页

(54) 发明名称

使用语音通信交互式地访问托管服务的系统和方法

(57) 摘要

一种在包括语音识别模块、会话管理器和语音产生器模块的系统中,用于向使用者提供服务的方法包括:通过所述语音识别模块接收话语;使用依附于本体论的词汇将所述话语转换为一个或更多个结构;使用所述结构辨识所述话语中的概念;如果所述话语包括足够的信息,则基于所述概念选择服务;基于被选服务产生文本消息;以及使用所述语音产生器来将所述文本消息转换为语音消息。



1. 一种在包括语音识别模块、会话管理器和语音产生器模块的系统中,用于向使用者提供服务的方法,所述方法包括:

通过所述语音识别模块接收话语;

使用依附于本体论的词汇将所述话语转换为一个或更多个结构;

使用所述结构辨识所述话语中的概念;

如果所述话语包括足够的信息,则基于所述概念选择服务;

基于被选服务产生文本消息;以及

使用所述语音产生器将所述文本消息转换为语音消息。

2. 如权利要求1所述的方法,还包括:将所述话语转换为文本,并且然后将所述文本转换为所述一个或更多个结构。

3. 如权利要求1所述的方法,还包括:当所述话语不包括足够的信息时,产生被设计来引出另外的信息的文本消息,使用语音产生器将所述文本消息转换为语音消息,并且将所述语音消息传送给所述使用者。

4. 如权利要求1所述的方法,还包括:将所述语音消息传送给所述使用者。

5. 如权利要求1所述的方法,还包括:

响应于传送给所述使用者的所述语音消息,通过所述语音识别模块接收另外的话语;

使用依附于本体论的词汇将所述话语转换为一个或更多个结构;

使用所述结构辨识所述话语中的概念;

基于所辨识的概念获得提供给所述使用者的信息;

基于所获得的信息或者所辨识的概念中的至少一个产生另一文本消息;

使用所述语音产生器将所述文本消息转换为语音消息;以及

将所述文本消息传送给所述使用者。

6. 如权利要求1所述的方法,其中所述使用依附于本体论的词汇将所述话语转换为一个或更多个结构的步骤包括将所述话语与和多个概念链接的一个或更多个单元相关联,所述多个概念又与多个服务相关联。

7. 如权利要求6所述的方法,其中所述使用所述结构辨识所述话语中的概念的步骤包括确定哪些概念与所述单元链接。

8. 如权利要求7所述的方法,还包括:基于所述概念选择服务的步骤包括基于用于所述多个服务中的每个的通过所述单元与所述话语链接的相关联概念的数量来确定哪个服务与所述话语最紧密相关。

9. 如权利要求1所述的方法,还包括:辨识与被选服务相关联的特殊概念,并且基于所述话语中所包括的信息将值分配给所辨识的特殊概念。

10. 如权利要求9所述的方法,还包括:确定与被选服务相关联的所有特殊概念是否存在于所述话语中,并且如果不是,则产生被设计来引出与任何丢失的特殊概念相关的信息的文本消息,并且使用所述语音产生器将所述文本消息转换为语音消息。

11. 如权利要求10所述的方法,还包括:当所有的特殊概念存在并且被分配值时,则执行或完成所述服务。

12. 如权利要求11所述的方法,其中所述执行服务的步骤包括储存被包括在所述话语中并且与所述特殊概念相关联的信息。

13. 一种用于向使用者提供服务的系统,所述系统包括:

语音识别模块,所述语音识别模块被配置来接收话语并且将所述话语转换为文本;使用依附于本体论的词汇转换为一个或更多个结构;

会话管理器,所述会话管理器被配置来与所述语音识别模块耦合,并且被配置来:

使用依附于本体论的词汇产生与所述文本相关联的结构;

使用所述结构辨识所述话语中的概念;

如果所述话语包括足够的信息,则基于所述概念选择服务;以及

基于被选服务产生回复文本消息;以及

语音产生器模块,所述语音产生器模块与所述会话管理器耦合,并且被配置来将所述回复文本消息转换为语音消息。

14. 如权利要求 13 所述的系统,其中所述会话管理器被进一步配置来当所述话语不包括足够的信息时,产生被设计来引出另外的信息的文本消息,并且其中所述语音产生器被进一步配置来将所述文本消息转换为语音消息,并且将所述语音消息传送给所述使用者。

15. 如权利要求 13 所述的系统,其中所述会话管理器被进一步配置来通过将所述话语与和多个概念链接的一个或更多个单元相关联以产生与所述文本相关联的结构,所述多个概念又与多个服务相关联。

16. 如权利要求 15 所述的系统,其中所述会话管理器被进一步配置来通过确定哪些概念与所述单元链接以使用所述结构辨识所述话语中的概念。

17. 如权利要求 16 所述的系统,其中所述会话管理器被进一步配置来通过下述方式基于所述概念选择服务,即,基于用于所述多个服务中的每个的通过所述单元与所述话语链接的相关联概念的数量来确定哪个服务与所述话语最紧密相关。

18. 如权利要求 13 所述的系统,其中所述会话管理器被进一步配置来辨识与被选服务相关联的特殊概念并且基于所述话语中所包括的信息将值分配给所辨识的特殊概念。

19. 如权利要求 18 所述的系统,其中所述会话管理器被进一步配置来确定与被选服务相关联的所有特殊概念是否存在于所述话语中,并且如果不是,则产生被设计来引出与任何丢失的特殊概念相关的信息的文本消息,并且其中所述语音产生器模块被进一步配置来使用所述语音产生器将所述文本消息转换为语音消息。

20. 如权利要求 19 所述的系统,其中所述会话管理器被进一步配置来当所有的特殊概念存在并被分配值时,执行或完成所述服务。

21. 如权利要求 20 所述的系统,其中所述会话管理器被进一步配置来通过储存被包括在所述话语中并且与所述特殊概念相关联的信息来执行所述服务。

## 使用语音通信交互式地访问托管服务的系统和方法

### 技术领域

[0001] 本申请中所公开的实施方案总地涉及交互式语音应答系统 (voice response system), 并且更具体地, 涉及能够通过电话实现对托管服务 (比如购物、汽车租赁、汽车旅馆等) 的语音通信访问的系统和方法。

[0002] 相关技术

[0003] 由于成本或方便的原因, 现今的公司惯常通过互联网和电话提供客户服务。目前, 使用者可以从提供视觉呈现能力的接入设备获得这样的互联网服务, 所述接入设备例如具有互联网 web 浏览器的个人电脑 (PC), 互联网 web 浏览器请求并接收由 Web 服务器产生的超文本标记语言 (HTML) 文档。对于电子商务应用, Web 服务器具有处理使用者输入的服务逻辑和事务服务器接口或者提供对这些接口的访问。使用任意数量的流行的 Web 编程工具对服务逻辑进行编程。

[0004] 使用者通过具有音频交互能力的接入设备获得电话服务, 所述接入设备例如调用具有音频输入、输出和电话功能的交互式语音应答 (IVR) 平台及其自己的服务逻辑和事务服务器接口的电话或者网络电话 (VOIP) 设备。使 IVR 系统自动地允许电话使用者通过口头命令访问系统上所链接的服务。通常以使用平台的应用编程接口 (API) 的通用软件语言或者平台特定脚本语言来对服务逻辑进行编程。

[0005] 传统的交互式 IVR 系统包括菜单、定向对话和可由话语识别技术的改进产生的混合主动式对话。菜单式交互通常使用预先录制的语音提示, 这些语音提示要求使用者按下电话键区上的数字或者讲简单的答复 (比如, “是”、“不是”) 或者简单的数字, 以从一组选项中选择项目。在定向对话中, 所述系统通过询问要求分别 (discrete) 答复的分别的问题来引导使用者通过数据集。例如, 为了找出操作者居住的地方, 一分别对话系统将首先要求操作者说出他所住的国家, 其后要求说出他所住的城市。混合主动式对话系统让使用者进入单个话语中的多个数据片段并且提供部分信息。

[0006] 尽管有这些进展, 但是常规的 IVR 往往仍是缓慢的、非个人的, 并且提供用于帮助系统与使用者之间的交互的笨重平台。通过电话上的纷繁复杂的菜单选项和选择进行操纵往往会非常耗时, 并且定向和混合主动式对话系统的语音命令识别 / 理解特征没有被设计来有效地处理不对脚本编写的问题进行应答的语音命令。简而言之, 现存的 IVR 均不允许由使用者对服务进行真正的交互式处理。

### 发明内容

[0007] 本文公开使用语音通信交互式地访问托管服务的方法和系统。

[0008] 在一个方面, 在包括语音识别模块、会话管理器和语音产生器模块的系统中, 用于向使用者提供服务的方法包括: 通过所述语音识别模块接收话语; 使用依附于本体论的词汇将所述话语转换为一个或更多个结构; 使用所述结构辨识所述话语中的概念; 如果所述话语包括足够的信息, 则基于所述概念选择服务; 基于被选服务产生文本消息; 以及使用所述语音产生器将所述文本消息转换为语音消息。

[0009] 在另一方面,用于向使用者提供服务的系统包括:语音识别模块,所述语音识别模块被配置来接收话语,并且使用依附于本体论的词汇将该话语转换为文本;使用依附于本体论的词汇转换为一种或更多种结构;会话管理器,所述会话管理器被配置来与所述语音识别模块耦合,并且被配置来使用依附于本体论的词汇产生与所述文本相关联的结构,使用所述结构辨识所述话语中的概念;如果所述话语包括足够的信息,则基于所述概念选择服务,并且基于被选服务产生回复文本消息;以及语音产生器模块,所述语音产生器模块与所述会话管理器耦合,并且被配置来将所述回复文本消息转换为语音消息。

[0010] 以下在标题为“具体实施方式”的部分中对本发明的这些和其他特征、方面和实施方案进行描述。

### 附图说明

[0011] 为了更完整地理解本文公开的原理及其优点,现在对结合附图进行的以下描述进行论述,在附图中:

[0012] 图 1 是图示说明根据一个实施方案的使得使用者使用语音通信交互式地访问托管服务的交互式语音应答 (IVR) 系统的功能元件的示图;

[0013] 图 2 是根据一个实施方案的可以被包括在图 1 所描绘的系统中的会话管理器的内部部件和这些部件如何与人类理解系统 (Human Understanding System) 中的其余模块交互的详细图示;

[0014] 图 3 是图示说明根据一个实施方案的图 2 和 3 中所图示说明的系统的操作的流程图;

[0015] 图 4 是更详细地图示说明根据一个实施方案的图 2 和 3 中所图示说明的系统的操作的流程图;以及

[0016] 图 5 是图示说明根据一个实施方案的结构的产生的示图,所述结构例如图 1 和 2 的会话管理器中可以发生的话语。

### 具体实施方式

[0017] 以下对使用语音通信交互式地访问托管服务的方法和系统进行描述。然而,将理解的是,可以在没有这些特定细节中的一些或全部的情况下实施本文所描述的系统和方法。在其他实例中,为了简洁,不对公知的处理操作进行详细描述。

[0018] 如本文所使用的,电话是提供远距离语音通信的通用设备,比如,陆线电话、移动电话、互联网通信设备等。电话包括通过模拟式信号(即,在时间和幅值上连续的信号)传输语音通信的传统模拟电话系统和通过数字式信号(即,离散二进制码)传输语音通信的最近的数字电话系统。网络电话 (VOIP) 是现代形式的基于数字的电话,其使用传输控制协议 / 互联网协议 (TCP/IP) 和用于通过互联网传输数字化的语音数据的其他网络传输格式。

[0019] 互联网或万维网 (WWW) 是由许多服务器构成的广域网 (WAN),这些服务器链接在一起,以使得可以使用诸如 TCP/IP、可靠使用者数据报协议 (RUDP) 或者它们的等同形式的网络数据传输协议将数据从一个服务器传输到另一个服务器。通常,互联网将位于广阔的地理区域中的众多服务器链接在一起。相反,局域网 (LAN) 是小型服务器网络,诸如覆盖小的局部区域(诸如家庭、办公室或学院)的那些服务器网络。

[0020] 鉴于前述,应该意识到 IVR 系统可以从本文所描述的系统和方法获益,本文所描述的系统和方法交互式地使用语音通信来确定哪些服务被客户请求,并且将这些服务传递给他们,而不使用菜单驱动的或者预先用脚本编写的对话。

[0021] 图 1 是图示说明根据一个实施方案的交互式语音应答 (IVR) 系统的功能元件的示图,该交互式语音应答 (IVR) 系统使得使用者可以使用语音通信交互式地访问托管服务。如本文所描绘的,所述系统接待 (handle) 操作电话设备 102 的使用者 101,电话设备 102 被配置来与人类理解系统 103 通信,人类理解系统 103 包含各种托管服务 108。为了通过人类理解系统 103 访问每个服务 106,必须首先在该系统中定义每个服务 106。

[0022] 因此,系统 100 可以被配置来处理多个服务 108,并且通过人类理解系统 103 将它们提供给使用者 101。在示例性注册处理期间,应该提供用于每个服务的以下元素:可使用该服务的地理区域;该服务用自然语言的详细描述;来自使用者 101 的以便该服务被提供给使用者 101 的所需数据的列表;以及通话脚本。

[0023] 在一个实施方案中,每个注册的托管服务与一组或更多组服务提供商(比如,购物、汽车租赁、汽车旅馆)相关。应该理解,人类生活的许多方面可以由合适的托管服务来服务,只要这些服务可以由使用者 101 通过电话设备 102 订购即可。

[0024] 在一个实施方案中,电话设备 102 可以通过被配置来使用模拟信号传输语音数据的模拟陆线(比如模拟物理有线连接)与人类理解系统 103 以通信方式链接。在另一个实施方案中,电话设备 102 可以通过被配置来使用离散的数字二进制信号传输语音数据的数字陆线(比如数字光纤连接)与人类理解系统 103 以通信方式链接。

[0025] 在另一个实施方案中,比如移动电话、卫星电话的电话设备 102 可以通过无线通信链路与人类理解系统 103 以通信方式链接,所述无线通信链路被配置来使用射频 (RF) 或微波信号将语音数据传输到人类理解系统 103。传输格式可以是模拟或数字的,并且无线通信链路可以是与人类理解系统 103 直接链接的链路或者可以通过下述基本单元 (base unit) 与人类理解系统 103 链接的链路,该基本单元通过陆线或另一无线连接与人类理解系统 103 连接。在又另一实施方案中,电话设备 102(即,互联网通信设备)可以以网络连接的方式通过陆线、无线连接或者这二者与人类理解系统 103 以通信方式链接,所述网络连接被配置来使用网络电话 (VOIP) 或等效协议来传输语音数据。网络连接可以被分布为局域化网络(即,局域网)或者广域网(即,互联网)。

[0026] 在一个实施方案中,所述系统可以被配置来通过操作移动电话(即,电话设备 102)的使用者 101 操作来呼入人类理解系统 103,以访问由人类理解系统 103 提供的服务。移动电话 102 通过 RF 链路的方式与移动电话提供商(即,蜂窝网络提供商)通信,移动电话提供商自身与公共交换电话网 (PSTN)(即,陆线)链接,公共交换电话网 (PSTN) 与人类理解系统 103 通信。人类理解系统 103 又可以通过互联网或 LAN 与多域系统以通信方式链接。在另一种情况下,使用者 101 可以操作启用 VOIP 的计算机(即,电话设备 102),来以 VOIP 呼入与互联网链接的人类理解系统 103。启用 VOIP 的计算机通过宽带互联网连接进行通信,所述宽带互联网连接通过网络连接(比如,互联网、LAN 等)与人类理解系统 103 以通信方式链接。

[0027] 应该意识到,包括以上提供的情况只是为了说明的目的,而并非意图以任何方式限制系统 100 可使用的通信配置。存在多种可想到的建立使用者 101 与人类理解系统 103

之间的通信的方法;仅受所得系统 (resulting system) 的这样的能力的限制,即以足够的清晰度和特异性 (specificity) 将语音数据传输到人类理解系统 103 以允许人类理解系统 103 处理和理解语音数据的能力。

[0028] 继续图 1,人类理解系统 103 包括以下:语音识别模块 104、语音产生器模块 106 和会话管理器模块 105。语音识别模块 104 可以被配置来通过电话设备 102 接收来自使用者 101 的话语,电话设备 102 比如使用上述电话通信配置中的任何一种与人类理解系统 103 以通信方式链接。在某些实施方案中,话语包括关于使用者 101 的信息,比如,辨识信息、认证信息以及关于使用者 101 正请求访问的被提供的服务的信息。语音识别模块 104 可以被配置来将从使用者 101 接收的语音数据转化为文本数据并且将该数据传送到会话管理器模块 105。

[0029] 将理解的是,人类理解系统 103 可以包括执行本文所描述的功能所需的部件,包括硬件和软件这二者。还将理解的是,人类理解系统 103 可以包括其他部件和功能,并且某些功能可以由相同或不同的部件来实现。因此,图 1 不应该被看作将本文所描述的系统和方法限于某种架构或配置。相反,图 1 仅以实施例的方式呈现。

[0030] 在一个实施方案中,语音识别模块 104 被配置来识别最常用世界语言。语音识别模块可以识别的语言的一些示例是:英语、汉语、印地语、西班牙语、孟加拉语、葡萄牙语、俄语、德语、日语和法语。在另一个实施方案中,语音识别模块 104 被配置来仅识别由人类理解系统 103 所处理的服务所指定的语言。然而,应该理解,语音识别模块 104 可以被系统管理员配置来识别任何语言,只要该语言的语言特性有益于将通过计算机处理被转换的语言即可。语音识别模块 104 还可以被配置来将通过设备 102 提供的使用者 101 的语音转换为文本。

[0031] 会话管理器模块 105 可以被配置来以两个状态操作:1) “选择服务状态”,在该状态下,会话管理模块 105 被配置来选择服务 106;以及 2) “通话脚本状态”,在该状态下,会话管理模块 105 被配置来执行适合于被选服务 106 的通话脚本 203。

[0032] 会话管理器模块 105 配备有用于与本体论联系的特定自然语言的词汇集。这使得它能够将自然讲话的文本转换为包含语法和本体信息的特殊结构。该本体信息可以被用于确定哪个服务 106 是所需的,并且还可以被用在“通话脚本状态”中,比如,以查询某些概念是否事实上被使用者 101 传送。

[0033] 当会话管理器 105 正在“选择服务状态”下工作时,它可以被配置来执行以下处理:选择已经定义了与客户话语中出现的本体概念最接近的本体概念的服务 108,并且同时排除为该服务定义的本体条件与该客户话语冲突的服务 108。当信息不足以让会话管理器 105 确定被请求的服务 106 的身份 (identity) 时,会话管理器 105 可以被配置来产生向使用者 101 询问进行该确定操作所需的必要信息的合适文本。会话管理器 105 然后将该文本转发给语音产生器模块 106,语音产生器模块 106 被配置来将该文本转换为话语,语音产生器模块 106 将该话语传送给电话装置 102 以供使用者 101 收听。

[0034] 在一个实施方案中,每当需要时,会话管理器 105 可以重复进行该处理,直到会话管理器 105 已接收到足够信息以确定由使用者 101 提供的语音数据中所请求的服务 108 的身份为止。在另一个实施方案中,系统管理员可以利用合适的通话脚本 203 启用或实现特殊服务,比如,服务选择 (service-to-choose) 服务,该通话脚本 203 可以要求客户提供更

多的信息和选择特定服务 108。在某些实施方案中, 通话脚本 203 可以调用其他服务。例如, 在租车应用中, 可以调用被配置来基于国家、城市、汽车类型等选择合适的租车公司的特殊服务, 以便完成该服务。

[0035] 应该意识到, 上面讨论的各种实施方案被配置来高度有效地实现使用者 101 与人类理解系统 103 之间的交互式对话。意图是尽可能接近地模仿使用者 101 与处理的服务 108 之间的通信环境, 包括试图确定哪个服务 108 或其他信息正被使用者 101 请求。与一系列菜单选项必须被选择和 / 或一系列信息请求被应答的常规系统相反, 其中, 选择和对应答的请求跟在通过预定决策树 (decision tree) 的相对固定的线性进程之后。这些常规交互完全不像两个人之间的自然通话。

[0036] 如在上面所描述的语音识别模块 104 的情况下, 在一个实施方案中, 语音产生器模块 106 可以被配置来仅能够进行 30 种最常用世界语言的通话。在另一个实施方案中, 语音产生器模块 106 可以被配置来仅识别人类理解系统 103 所处理的服务所指定的语言。然而, 应该意识到, 语音产生器模块 106 可以被系统管理员配置来识别任何语言, 只要该语言的语音特性有益于将通过计算机处理被转换的语言即可。

[0037] 一旦会话管理器 105 已成功地辨识使用者请求的服务 108, 会话管理器 105 就开始执行为该服务定义的通话脚本。

[0038] 图 2 是根据一个实施方案的会话管理器 105 的内部部件和这些部件如何与人类理解系统 103 中的其余模块交互的详细图示。如该实施方案所示, 会话管理器 105 包括脚本引擎 202、为所有服务 108 定义的通话脚本 203、服务数据库 204 和服务数据管理部件 205。

[0039] 会话管理器还包括文本到结构模块 212, 文本到结构模块 212 可以使用合适的词汇 210 将人的话语的文本转换为特殊结构。这些词汇 210 与本体论 211 链接。这种关联允许文本到结构模块 212 将以任何限定的人类语言的文本转换为与人类语言无关的包含语法和本体信息的结构。

[0040] 会话管理器 105 还可以包括特殊概念勘探器 (prospector) 230。为了理解特殊概念勘探器 230, 有必要说明这些特殊概念是什么。它们是不必被理解的概念, 比如, 名称、颜色、地址。相反, 感兴趣的是可以与概念相关联的值或等级。这样的概念具有使用所有自然语言的一些模式, 这些模式允许人们可以谈论它们。因此, 特殊概念勘探器 230 分析从话语生成的文本和结构, 以提取特殊概念值。另外, 可以以合适的语言定义问题集, 这些问题集被设计来生成来自使用者 101 的可以使与所述值相关联的概念更清晰的信息。

[0041] 当脚本引擎 202 遇到对通话脚本 203 中的特殊概念的请求时, 它调用特殊概念勘探器 230 来得到与该概念相关联的值。在缺乏该信息的情况下, 特殊概念勘探器 230 将通过语音产生器模块 106 转发给客户 101 的合适问题返回给脚本引擎 202。在接收到答复之后, 如果有必要的话, 则可以重复进行所述处理。如果客户拒绝提供答复, 则脚本引擎 202 将记下该事实, 以使得脚本可以做出反应。以下对特殊概念和它们的使用进行更详细的描述。

[0042] 特殊指令可以被嵌入通话脚本 203, 以查询合适的数据库。其他指令可以被嵌入与服务 108 相关联的通话脚本 203, 来允许向客户呈现结果, 以便使得他能够选择特定值。另外的指令可以被用于以合适的数据库更新数据库 204。

[0043] 通话脚本 203 中的条件指令使得可以对预期的人类反应作出反应。这些条件指令



可以包括指令,诸如可以使得系统 103(或者更具体地,通话脚本 203)能够处理各种情况(诸如如果使用者 101 拒绝提供答复或特定信息,则该怎么做)的 if、switch、while 等指令。

[0044] 通话脚本 203 可以被配置来对消息数字进行操作,这些消息数字可以根据客户所使用的语言被转换为自然语言的消息。

[0045] 图 3 是更详细地图示说明根据一个实施方案的示例性处理的流程图,通过该处理,系统 100 被用于选择服务 108 和将服务 108 提供给使用者 101(“选择服务状态”)。在步骤 302 中,语音识别引擎 104 可以通过设备 102 和相关联的通信架构从使用者 101 接收语音话语。如所指出的,话语可以包括信息,诸如:辨识信息,即,名称、编号、地址等;认证信息,即,社保号、确认号等;关于期望服务的信息,即,销售、客户服务、预订等;仅仅列出了可以被包括在从使用者 101 接收的话语中的信息类型的几个示例。

[0046] 在步骤 304 中,语音识别模块可以被配置来将话语转化或转换为可以在步骤 306 中被转发给会话管理器 105 的文本。

[0047] 会话管理器 105 然后可以在步骤 308 中使用与本体论链接的词汇将文本转换为结构。在步骤 310 中,会话管理器 105 可以使用所述结构来辨识包括在话语中的概念,即,包括在话语中的本体概念。然后可以在步骤 312 中使用所述概念来辨识正被请求的服务 108。以下对概念及其在系统 103 内的辨识进行更详细的描述。

[0048] 当在话语中不存在用于充分辨识服务 108 的足够信息时,会话管理器则可以被配置来在步骤 314 中产生文本消息,该文本消息被设计来引出另外的信息以允许服务 108 的辨识。所述文本消息可以使用通话脚本 203 来产生,并且可以在步骤 316 中被发送到语音产生器 106,语音产生器 106 可以被配置来将文本消息转换为将通过设备 102 被传送给使用者 101 的讲话。

[0049] 然后将重复进行所述处理,直到在步骤 318 中出现足够的信息来选择服务。一旦服务被选择,合适的通话脚本 203 就可以在步骤 320 中被激活,以便产生可以在步骤 322 中被转发给语音产生器 106 的文本消息,从而与使用者 101 通信。

[0050] 图 4 是更详细地图示说明根据一个实施方案的选择由会话管理器 105 承担的服务的过程的流程图。在步骤 402 中,脚本引擎 202 从语音识别模块 104 接收文本信息,并且将文本发送到文本到结构引擎 212,在步骤 404 中文本到结构引擎 212 可以被配置来使用合适的词汇 210 和相关联的本体论 211 将文本转换为结构。然后在步骤 406 中服务选择器 220 可以使用所述结构来从服务数据库 204 选择服务。

[0051] 在步骤 408 中,被选服务然后可以使脚本引擎 202 使用通话脚本 203 产生提供合适信息所需的合适文本消息或者向使用者 101 索求信息。换句话讲,与通话脚本 203 配合的脚本引擎 202 允许系统 103 以自然的方式如由被选服务决定的那样与使用者 101 进行通信。

[0052] 服务数据管理器 205 可以被配置来将与各种服务相关联的数据储存在服务数据库 204 中。另外,服务数据库 205 可以被配置来储存从使用者 101 以及比如特殊概念处理器 230 接收的信息。因此,当在步骤 408 中产生消息时,在步骤 410 中响应于该消息来接收新的信息。在步骤 412 中,可以将与在步骤 410 中接收的信息相关联的信息或值储存在服务数据库 412 中。

[0053] 图 5 是图示说明根据一个实施方案的用于形成结构并将它们与概念链接的实施例的示意图。图 5 图示说明用于以下话语的处理：“I would like to make an appointment for my wife with her doctor”和等同的波兰语“**Chciałbym umówić moją żonę z jej lekarzem**”。每个话语 502(英语)和 514(波兰语)包括如以用于该语言的相关联的词汇 210 定义的多个词素。一般而言,每种语言的词汇 210 中的每个词位直接与一个合适的概念或多个合适的概念 508 联系,或者通过具有变形的语言中的词根与一个合适的概念或多个合适的概念 508 联系。因为在自然语言交流中,表达相同概念有许多同义的不同方式,所以每个概念 508 与每个词汇 210 中的多个词素 (lexeme) 联系。

[0054] 仍参照图 5,话语 502 和 514 可以被转换为单元 504 或 512。每个单元 504 和 512 可以对应于话语中的一个或更多个词素。例如,单元 504 可以对应于单个词素,或者可以具有多个子单元,以使得它可以对应于多个词素,诸如图 5 的实施例中的话语 502 中的“would like to”的分组。每个单元 504 或子单元然后可以基于相关联的本体论 211 通过链接 (links) 506 和 510 被链接到一个或更多个概念 508。

[0055] 单元 504 和 512 与链接 506 和 510 的组合分别形成与各个概念 508 相关联的结构 505 和 515。应该指出,无论输入话语中所包括的指引或导致概念被辨识的词素如何,每个概念 508 都将具有唯一的结构。换句话讲,如果话语 502 被稍微不同地陈述,即,“I would like to make a doctor appointment for my wife”或者“I would like to see if my wife’s doctor is available”,则应该产生非常类似的、将会导致辨识相同概念 508 的结构 505 和 515。

[0056] 因此,系统 103 中所定义的服务 108 可以具有与它相关联的一组概念 508。在许多实施方案中,仅包括重要的概念,以避免不必要的关联。此外,在其他实施方案中,省略仅表达语法的词素。可选地,每个服务还可以具有排除条件和必需条件。例如,如果某个服务 108 将被选择,则必须为将被选择的该某个服务 108 提供某些信息,而不能包括其他信息。

[0057] 概念 508 然后可以与服务 108 相关联。换句话讲,每个服务将具有与它相关联的某些概念。会话管理器 105 然后可以确定与话语(比如,话语 502)相关联的结构(比如,结构 505),并且基于结构 505 确定什么服务 108 与结构 505 最紧密关联,即,具有与话语 502 相关联的大多数概念 508。在某些实施方案中,在确定哪个服务 108 与话语 502 最紧密关联时,会话管理器 105 还可以被配置来确定任何必需条件和排除条件是否已经满足每个服务 108。

[0058] 一旦服务 108 被选择,则下一个步骤就可以是使用通话脚本 203 收集用于该服务的数据。这可以涉及如上所描述的特殊概念。再次,特殊概念是实际上不试图理解的概念系统 103,但是对于该系统 103,试图产生相关联的值或等级。在图 5 的实施例中,这样的概念可以是:人,即,谁要去看医生;时间,即,何时应该进行预约;以及医生,需要哪个医生。

[0059] 首先,可以对话语(比如,话语 502)进行评估,以确定它是否包括必要数据。值或等级是所提供的数据有多强烈地对应于该信息的指示。确定数据是否被包括并且提供值或等级的任务可以落到特殊概念处理器 230。如果表现为话语 502 不包括所需数据,则脚本引擎 202 可以在被选服务 108 的控制下选择被设计来引出所述信息的通话脚本。在以上实施例中,客户将被询问预约的日期(在谈话脚本中),这是因为当预订该服务时,仅该信息尚

未被指定。

[0060] 一旦产生的值指示提供了所有的所需信息,则数据可以被储存,并且可以被用于进行预约。通话脚本 204 则可以被用于产生可以被传送给使用者的确认预约的消息。

[0061] 再次,通话脚本 203 使用自然语言以自然的方式产生接收的话语并对接收的话语做出响应。因此,当使用者 101 被提示提供所需信息时,他们应该觉得他们好像是正在进行正常的通话似的。

[0062] 本文所描述的系统和方法的全部或部分还可以被实施为计算机可读介质上的计算机可读代码。计算机可读介质是可以储存其后可以被计算机系统读取的数据的任意数据储存设备。计算机可读介质的示例包括硬盘驱动器、联网储存器 (NAS)、只读存储器、随机存取存储器、CD-ROM、CD-R、CD-RW、磁带和其他光学和非光学数据储存设备。计算机可读介质还可以被分布在联网的计算机系统上,以使得计算机可读代码被以分布式方式储存和执行。

[0063] 虽然本文已经对几个实施方案进行了详细描述,但是本领域技术人员应该理解,可以以许多其他特定形式实施本文所描述的系统和方法。因此,认为提供的实施例和实施方案是示例性的,而不是限制性的,并且本文所描述的系统和方法不限于本文提供的细节,而是可以在所附的权利要求的范围内被修改和实施。

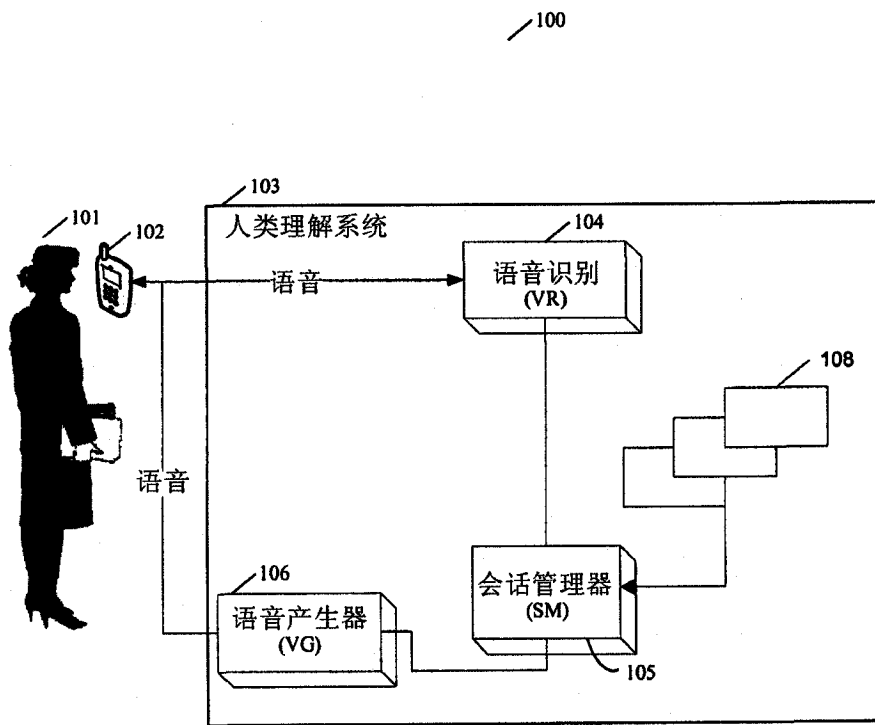


图 1

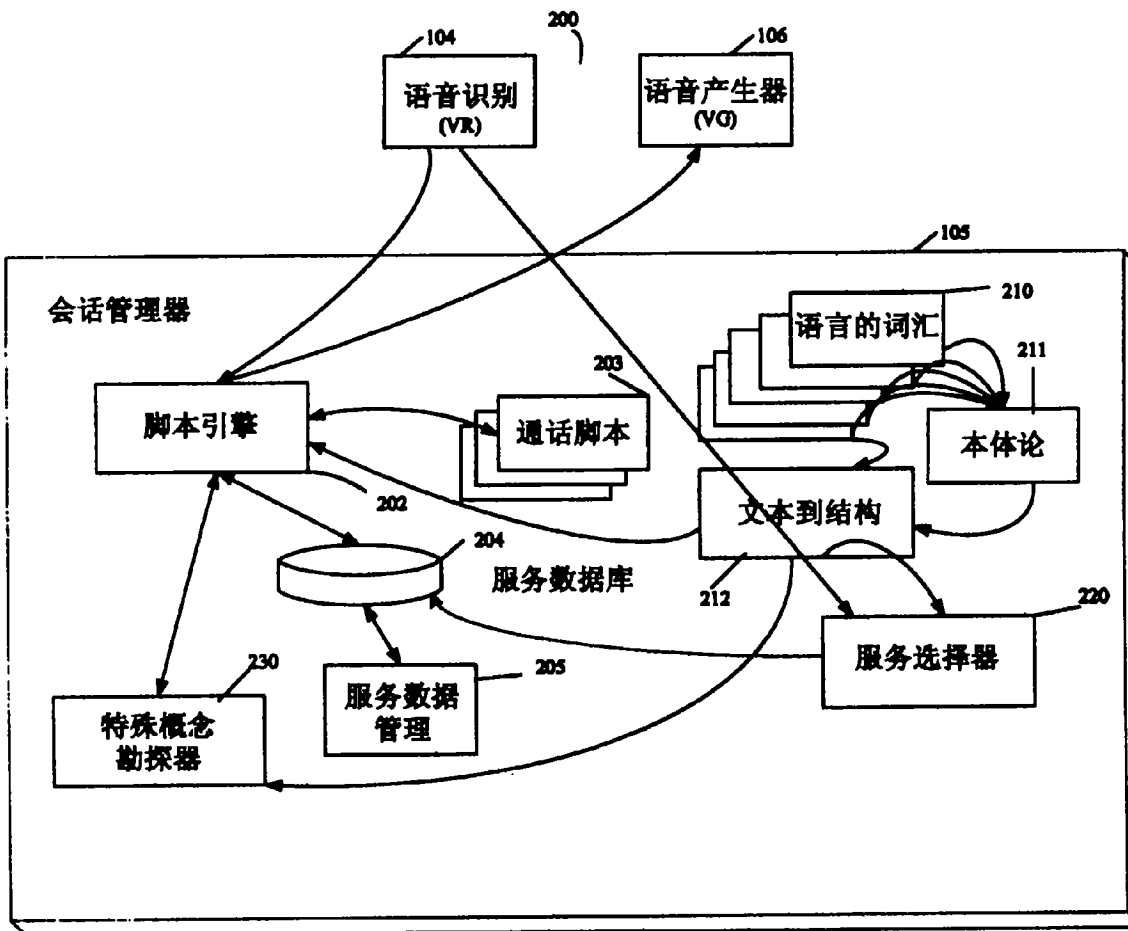


图 2

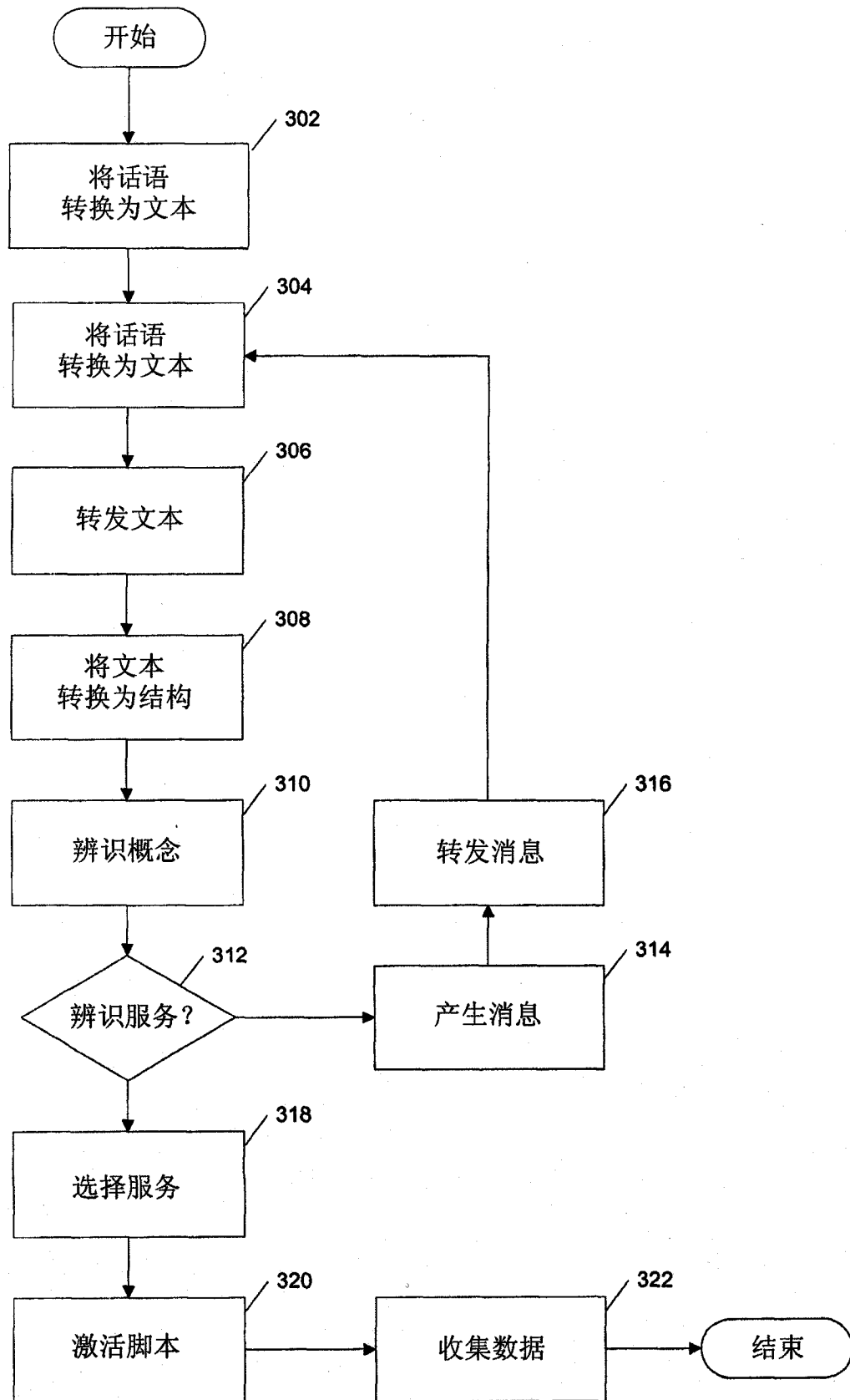


图 3

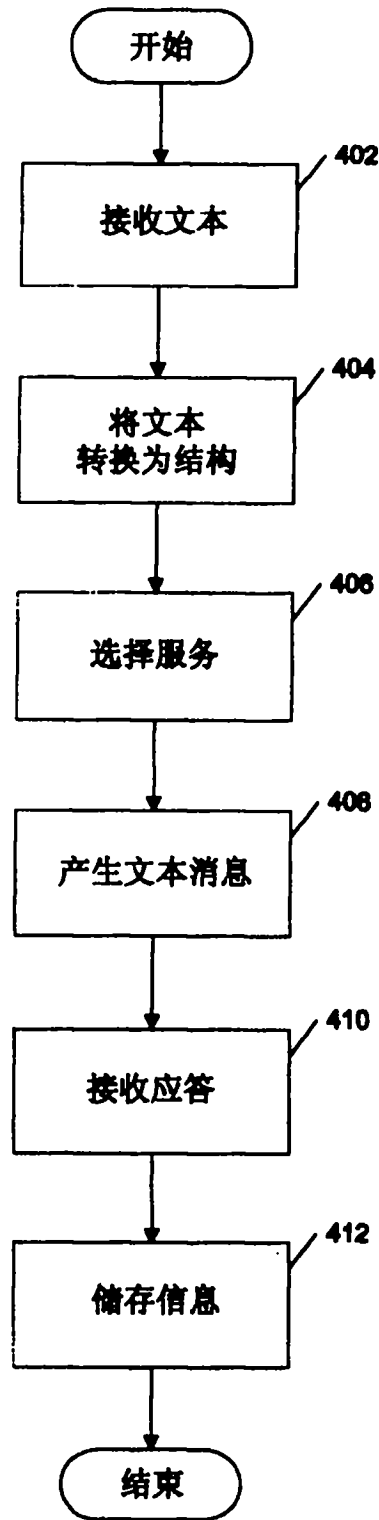


图 4

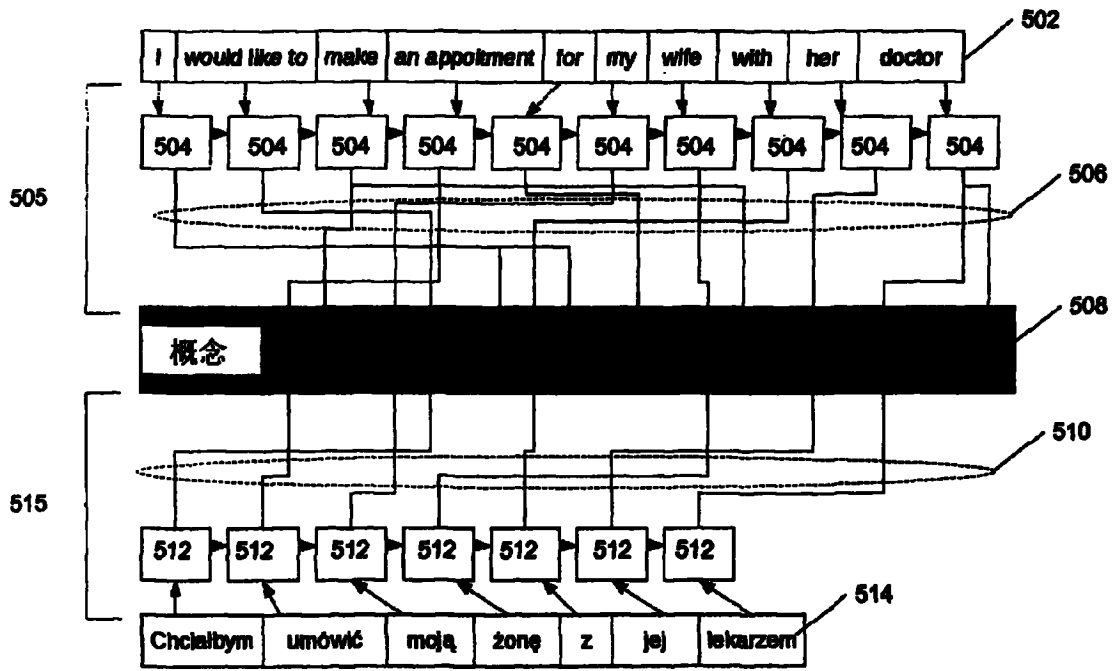


图 5