

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 4 月 16 日 (16.04.2020)



(10) 国际公布号
WO 2020/073532 A1

(51) 国际专利分类号:

G06F 16/332 (2019.01)

(21) 国际申请号: PCT/CN2018/125251

(22) 国际申请日: 2018 年 12 月 29 日 (29.12.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:

201811191523.3 2018 年 10 月 12 日 (12.10.2018) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN]; 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 许开河(XU, Kaihe); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 杨坤(YANG,

Kun); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 王少军(WANG, Shaojun); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 肖京(XIAO, Jing); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市隆天联鼎知识产权代理有限公司(SHENZHEN LUNGTIN LIANDING INTELLECTUAL PROPERTY AGENT LTD.); 中国广东省深圳市福田区南园路上田大厦 4A 刘抗美, Guangdong 518000 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,

(54) **Title:** METHOD AND DEVICE FOR IDENTIFYING CONVERSATION STATE OF CUSTOMER SERVICE ROBOT, ELECTRONIC DEVICE, AND COMPUTER-READABLE STORAGE MEDIUM

(54) 发明名称: 客服机器人对话状态识别方法及装置、电子设备、计算机可读存储介质

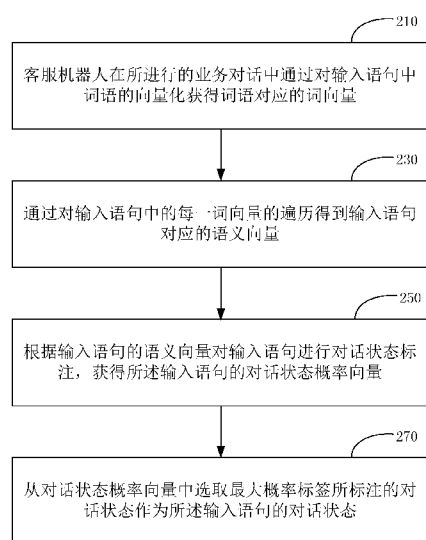


图 2

210 A customer service robot obtaining word vectors corresponding to words by means of performing vectorization on words in an input sentence in a performed service conversation
230 Obtain a semantic vector corresponding to the input sentence by means of traversing each word vector in the input sentence
250 Label a conversation state of the input sentence according to the semantic vector of the input sentence to obtain a conversation state probability vector of the input sentence
270 Select a conversation state labelled with a maximum probability label from the conversation state probability vector as the conversation state of the input sentence

(57) **Abstract:** A method and device for identifying a conversation state of a customer service robot. The method comprises: a customer service robot obtaining word vectors corresponding to words by means of performing vectorization on words in an input sentence in a performed service conversation (210); obtaining a semantic vector corresponding to the input sentence by means of traversing each word vector in the input sentence (230); labeling a conversation state of the input sentence according to the semantic vector of the input sentence to obtain a conversation state probability vector of the input sentence (250); and selecting a conversation state labelled having a maximum probability label from the conversation state probability vector as the conversation state of the input sentence (270). The method above enables the customer service robot to adaptively converse with a customer, and has better initiative.

(57) **摘要:** 一种客服机器人对话状态识别方法及装置, 所述方法包括: 客服机器人在所进行的业务对话中通过对输入语句中词语的向量化获得所述词语对应的词向量(210); 通过对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量(230); 根据所述输入语句的语义向量对所述输入语句进行对话状态标注, 获得所述输入语句的对话状态概率向量(250); 从所述对话状态概率向量中选取最大概率标签所标注的对话状态作为所述输入语句的对话状态(270)。上述方法使得客服机器人可以自适应地与客户进行对话, 主动性更佳。

JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,
LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明，要求每一种可提供的地区
保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, 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)。

本国际公布：

— 包括国际检索报告 (条约第21条(3))。

客服机器人对话状态识别方法及装置、电子设备、计算机可读存储介质

技术领域

[0001] 本申请要求2018年10月12日递交、申请名称为“客服机器人对话状态识别方法及装置、设备、存储介质”的中国专利申请201811191523.3的优先权，在此通过引用将其全部内容合并于此。

[0002] 本申请涉及数据处理技术领域，尤其涉及一种客服机器人对话状态识别方法及装置、电子设备、计算机可读存储介质。

背景技术

[0003] 客服机器人一般的工作流程为:当客服机器人接收到用户提出的问题，将接收的问题与知识库中设定的答案进行匹配，并将匹配到的答案返回给客户，从而产生一系列的对话。

[0004] 发明人意识到，目前市面上的客服机器人无法从整体上理解整个对话，只能根据用户的提问被动地回答问题，不能判断当前对话是否需要寒暄，不能判断当前对话是否需要开始回答问题，导致客服机器人的主动性较弱，在实际应用中导致客户体验感不佳。

发明概述

技术问题

[0005] 为了解决上述技术问题，本申请的一个目的在于提供一种客服机器人会话状态识别方法及装置、电子设备、计算机可读存储介质。

问题的解决方案

技术解决方案

[0006] 其中，本申请所采用的技术方案为：

[0007] 一方面，一种客服机器人会话状态识别方法，包括：客服机器人在所进行的业务对话中，通过对输入语句中词语的向量化获得所述词语对应的词向量；通过

对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量；根据所述输入语句的语义向量对所述输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量；从所述对话状态概率向量中选取最大概率标签所标注的对话状态作为所述输入语句的对话状态。

[0008] 另一方面，一种客服机器人对话状态识别装置，包括：词向量获取模块，用于客服机器人在所进行的业务对话中，通过对输入语句中词语的向量化获得所述词语对应的词向量；语义向量获取模块，用于通过对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量；对话状态标注模块，用于根据所述输入语句的语义向量对所述输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量；对话状态获取模块，用于从所述对话状态概率向量中选取最大概率标签所标注的对话状态作为所述输入语句的对话状态。

[0009] 另一方面，一种电子设备，包括处理器及存储器，所述存储器上存储有计算机可读指令，所述计算机可读指令被所述处理器执行时实现如上所述的客服机器人对话状态识别方法。

[0010] 另一方面，一种计算机可读存储介质，其上存储有计算机程序，所述计算机程序被处理器执行时实现如上所述的客服机器人对话状态识别方法。

发明的有益效果

有益效果

[0011] 在上述技术方案中，输入语句为客户向客服机器人进行的提问，通过对输入语句进行对话状态标注获得输入语句的对话状态。客服机器人可以根据输入语句的对话状态，从知识库中提取与输入语句对话状态相适应的语句与客户进行对话，增加了客服机器人的主动性。

[0012] 应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本申请。

对附图的简要说明

附图说明

[0013] 此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本申请的实施例，并于说明书一起用于解释本申请的原理。

- [0014] 图1是根据一示例性实施例示出的一种客服机器人的硬件框图；
- [0015] 图2是根据一示例性实施例示出一种客服机器人对话状态识别方法的流程图；
- [0016] 图3是根据图2对应实施例示出的对步骤210进行描述的流程图；
- [0017] 图4是根据图2对应实施例示出的对步骤250进行描述的流程图；
- [0018] 图5是根据另一示例性实施例示出一种客服机器人对话状态识别方法的流程图；
- [0019] 图6是根据一示例性实施例示出的一种客服机器人对话状态识别装置的框图。
- [0020] 通过上述附图，已示出本申请明确的实施例，后文中将有更详细的描述，这些附图和文字描述并不是为了通过任何方式限制本申请构思的范围，而是通过参考特定实施例为本领域技术人员说明本申请的概念。

发明实施例

本发明的实施方式

- [0021] 这里将详细地对示例性实施例执行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本申请相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本申请的一些方面相一致的装置和方法的例子。
- [0022] 图1是根据一示例性实施例示出的一种客服机器人的硬件框图。需要说明的是，该客服机器人只是一个适配于本公开的示例，不能认为是提供了对本公开使用范围的任何限制。
- [0023] 如图1所示，客服机器人可以包括以下一个或者多个组件：处理组件101，存储器102，电源组件103，多媒体组件104，音频组件105，传感器组件107以及通信组件108。其中，上述组件并不全是必须的，客服机器人可以根据自身功能需求增加其他组件或减少某些组件，本实施例不作限定。
- [0024] 处理组件101通常控制客服机器人的整体操作，诸如与显示，数据通信，相机操作以及日志数据处理相关联的操作等。处理组件101可以包括一个或多个处理器109来执行指令，以完成上述操作的全部或部分步骤。此外，处理组件101可以包括一个或多个模块，便于处理组件101和其他组件之间的交互。例如，处理

组件101可以包括多媒体模块，以方便多媒体组件104和处理组件101之间的交互。

[0025] 存储器102被配置为存储各种类型的数据以支持在客服机器人的操作。这些数据的示例包括用于在客服机器人上操作的任何应用程序或方法的指令。存储器102可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如SRAM（静态随机存取存储器），EEPROM（电可擦除可编程只读存储器），EPROM（可擦除可编程只读存储器），PROM（可编程只读存储器），ROM（只读存储器），磁存储器，快闪存储器，磁盘或光盘。存储器102中还存储有一个或多个模块，该一个或多个模块被配置成由该一个或多个处理器109执行，以实现以下任一所示方法中的全部或者部分步骤。

[0026] 电源组件103为客服机器人的各种组件提供电力。电源组件103可以包括电源管理系统，一个或多个电源，及其他与为客服机器人生成、管理和分配电力相关联的组件。

[0027] 多媒体组件104包括在所述客服机器人和用户之间的提供一个输出接口的屏幕。在一些实施例中，屏幕可以包括LCD（液晶显示器）和TP（触摸面板）。如果屏幕包括触摸面板，屏幕可以被实现为触摸屏，以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。所述触摸传感器可以不仅感测触摸或滑动动作的边界，而且还检测与所述触摸或滑动操作相关的持续时间和压力。

[0028] 音频组件105被配置为输出和/或输入音频信号。例如，音频组件105包括一个麦克风，当客服机器人处于操作模式，如记录模式和语音识别模式时，麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器102或经由通信组件108发送。音频组件105还包括一个扬声器，用于输出音频信号，以实现客服机器人与客户之间进行会话操作。

[0029] 传感器组件107包括一个或多个传感器，用于为计算机设备提供各个方面的状态评估。例如，传感器组件107可以检测到客服机器人的打开/关闭状态，组件的相对定位，传感器组件107还可以检测客服机器人或客服机器人一个组件的坐标改变以及客服机器人的温度变化。在一些实施例中，该传感器组件107还可以包

括磁传感器，压力传感器或温度传感器。

[0030] 通信组件108被配置为便于客服机器人和其他设备之间有线或无线方式的通信。客服机器人可以接入基于通信标准的无线网络，如WiFi，2G或3G，或它们的组合。在一个示例性实施例中，通信组件108经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。

[0031] 在示例性实施例中，客服机器人可以被一个或多个ASIC（应用专用集成电路）、DSP（数字信号处理器）、PLD（可编程逻辑器件）、FPGA（现场可编程门阵列）、控制器、微控制器、微处理器或其他电子元件实现，用于执行上述方法。

[0032] 在本实施例中，客服机器人是用于自动执行对话工作的机器装置，具体可以是一种智能手机、平板电脑、笔记本电脑等终端设备，或者是其他具备特定外形和功能的机器设备。

[0033] 本实施例中的客服机器人的处理器执行操作的具体方式将在有关客服机器人对话状态识别方法的实施例中进行详细描述，此处将不做详细阐述说明。

[0034] 图2是根据一示例性实施例示出的一种客服机器人对话状态识别方法的流程图，该方法适用于图1所示的客服机器人。如图2所示，该方法可以包括以下步骤：

[0035] 在步骤210中，客服机器人在所进行的业务对话中通过对输入语句中词语的向量化获得词语对应的词向量。

[0036] 其中，输入语句是客服机器人在与客户进行的业务对话中，用户向客服机器人提问的对话消息。

[0037] 输入语句可以是客服机器人通过识别客户输入的语音信号获得。例如，客服机器人通过自身所配置的麦克风获取客户输入的提问语音，并对获取的提问语音进行语音识别获得输入语句。

[0038] 输入语句也可以通过客服机器人所配置的触摸屏获取。例如，客户通过在触摸屏上输入想要提问的问题，此时，客服机器人直接获取触摸屏上输入的文本信息为输入语句。

[0039] 客服机器人获取输入语句后，通过对输入语句中的词语进行向量化，获得这些

词语所对应的词向量。

[0040] 在一实施例中，如图3所示，客服机器人对输入语句中的词语进行向量化可以具体包括以下步骤：

[0041] 在步骤211中，通过对输入语句进行分词处理获得输入语句中的目标词语。

[0042] 其中，对输入语句进行分词处理，是将输入语句分割成由若干个目标词语顺序排列的文本序列。例如，输入语句为“您好，我有几个问题想请问您？”，进行分词处理得到的文本序列为“您好^我^有^几个^问题^想^请问^您”。

[0043] 对输入语句进行分词处理可以是采用分词算法进行的，例如，可采用基于字符串匹配的分词方法、基于理解的分词方法或者基于统计的分词方法等，本处不进行限定。

[0044] 在步骤213中，根据所对应上下文信息将每一目标词语映射为低维向量，获得每一目标词语对应的词向量。

[0045] 其中，可采用one-hot（独热码）向量编码方式或者word2vec（word embeddings，词向量）向量编码方式将文本序列中的每一目标词语映射为低维向量，获得每一目标词语对应的词向量，以通过每一词向量建立不同目标词语之间的语义关联。词向量的维数具体为文本序列中目标词语的数量。

[0046] 需要说明的是，采用one-hot向量编码方式以及word2vec向量编码方式获得输入语句中每一目标词语对应的词向量是本领域的现有技术，其具体实现过程本处不进行赘述。

[0047] 由于采用one-hot向量编码方式所得的one-hot向量没有存储输入语句中各目标词语之间的关联性，还需对每一目标词语所得的one-hot向量加上权重信息。每一目标词语所加的权重大小与该目标词语对输入语句语义的关联程度有关，例如在输入语句“您好，我有几个问题想请问您？”中，目标词语“问题”和目标词语“请问”最大程度表达了客户的意图，对输入语句的语义关联性较大，所对应的权重应当较高。将所得的one-hot向量加上权重，则获得每一目标词语对应的词向量。

[0048] 而采用word2vec向量编码方式对输入语句中每一目标词语映射为低维向量，是将每一目标词语相对于输入语句中其他周边目标词语依次进行条件映射获得的

，所得的每一词向量能够与输入语句的语义相关联。

[0049] 因此，通过本实施例提供的方法对输入语句进行向量化，所得的每一词向量能够与输入语句的语义建立关联。

[0050] 在步骤230中，通过对输入语句中的每一词向量的遍历得到输入语句对应的语义向量。

[0051] 其中，对输入语句中每一目标词语对应的词向量进行遍历是通过神经网络模型实现的，本实施例可采用LSTM（Long Short-Term Memory，长短期神经网络）模型。

[0052] 具体地，将输入语句中的每一词向量顺序输入LSTM模型中，并按照时间先后顺序对输入的每一词向量进行遍历，将遍历所得的隐状态向量作为输入语句对应的语义向量。

[0053] 在遍历的过程中，会将上一词向量所对应的隐藏层状态输出给下一词向量对应的隐藏层，使得下一词向量的隐藏层根据上一词向量的隐藏层状态进行状态更新，直至对最后一词向量对应的隐藏层进行状态更新后，完成遍历操作。将每一词向量所对应隐藏层更新得到的状态向量称为对该词向量进行遍历所得的隐状态向量。

[0054] 在一实施例中，由于在对输入语句中最后一词向量的隐藏层进行状态更新时，考虑了该词向量之前全部词向量的隐藏层状态，使得对最后一词向量进行遍历所得的隐状态向量能够对输入语句建立全局上的语义表达。因此，可将对最后一词向量进行遍历所得的隐状态向量作为输入语句的语义向量。

[0055] 在另一实施例中，可以对每一词向量进行遍历所得的隐状态向量进行拼接，将拼接所得的向量作为输入语句的语义向量。

[0056] 应当说明的是，对输入语句中每一词向量进行遍历还可以是通过Bi-LSTM（Bi-Long Short-Term Memory，双向长短期神经网络）模型进行的，或者是其他模型，本处并不对此进行限定。

[0057] 在步骤250中，根据输入语句的语义向量对输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量。

[0058] 其中，对输入语句进行对话状态标注的目的在于，获取输入语句可能被标注为

对话状态标注空间中每一对话状态的概率，并对所得的概率按照对话状态标注空间中每一对话状态的顺序进行排列，将所得的概率序列称为输入语句的对话状态概率向量。

[0059] 对话状态标注空间为预置的样本空间，用于为输入语句提供所标注的对话状态。在一实施例中，对话状态标注空间中预置有9类对话状态，依次为：客户寒暄、坐席寒暄、客户提问、坐席回答、客户追问、坐席追问、客户回答、客户结束、坐席结束。

[0060] 假设输入语句“您好，我有几个问题想请问您？”分别被标注为上述9种对话状态的概率依次为0.85、0.11、0.43、0.03、0.18、0.21、0.27、0.16、0.15，则获取的对话状态概率向量为（0.85，0.11，0.43，0.03，0.18，0.21，0.27，0.16，0.15）。

[0061] 如图4所示，在一示例性实施例中，步骤250具体可以包括以下步骤：

[0062] 在步骤251中，对输入语句的语义向量进行非线性映射，获得输入语句在对话状态标注空间中的状态向量。

[0063] 其中，对输入语句的语义向量进行非线性映射的目的在于，将语义向量映射至对话状态标注空间，获得输入语句在对话状态标注空间中的状态向量，从而与对话状态标注空间建立关联。

[0064] 便于理解的，若将输入语句中每一词向量表示为 $(x_1, x_2, \dots, x_n)$ ，将所述输入语句遍历所得隐状态向量表示为 $H = (h_1, h_2, \dots, h_n)$ ，其中 $H \in \mathbb{R}^{(n \times m)}$ ，其维数为 m ，并将输入语句在对话状态标注空间中的状态向量可表示为 $P = (p_1, p_2, \dots, p_n)$ ，其中 $P \in \mathbb{R}^{(n \times k)}$ ，其维数为 k 维，则将输入语句的语义向量映射为状态标注空间空的状态向量的实质为，将语义向量从 m 维映射为 k 维。其中， m 是输入语句中最后一词向量所对应隐藏层的状态向量的元素数量， k 是状态标注空间中预置对话状态的数量。

[0065] 在步骤253中，根据目标状态转移矩阵对输入语句在对话状态标注空间中的状态向量进行对话状态标注打分。

[0066] 在一实施例中，对输入语句在对话状态标注空间中的状态向量进行对话状态标注打分，是通过CRF（Conditional Random Fields，条件随机场）模型进行的，以

通过进行对话状态标注打分，获得对将输入语句分别标注为对话状态标注空间中每一对话状态的打分情况。得分越高，将输入语句标注为该种对话状态的可能性越大。

[0067] 假设对话状态标注空间中预置的一个长度等于输入语句长度的对话状态序列为 $y = (y_1, y_2, \dots, y_n)$ ，并假设输入语句表示为 X ，则使用CRF模型将输入语句标记为对话状态序列 y 的打分为：

$$s(X, y) = \sum_{i=0}^n P_{i, y_i} + \sum_{i=0}^n A_{y_i, y_{i+1}}$$

[0068] 其中，“ P_{i, y_i} ”表示输入语句的语义向量被标注为第 i 个对话状态的得分， A 为CRF模型的目标状态转移矩阵，“ $A_{y_i, y_{i+1}}$ ”表示从第 i 个对话状态转移到第 $i+1$ 个对话状态的转移得分。

[0069] 可以看出，对输入语句在对话状态标注空间中的状态向量进行对话状态预测打分由两部分得到，其中一部分打分由输入语句在对话状态标注空间中的状态向量决定，另一部分打分由CRF模型的目标状态转移矩阵决定。

[0070] 在步骤255中，对输入语句的对话状态预测打分进行概率归一化，获得输入语句的对话状态概率向量。

[0071] 对输入语句的对话状态预测打分进行概率归一化采用Softmax多分类预测函数进行，本实施例定义的Softmax多分类预测函数为：

$$\sigma(y|X) = \frac{e^{s(X, y)}}{\sum_y e^{s(X, y)}}$$

[0072] 其中，“ y^* ”表示输入语句的真实对话状态，“ y ”表示输入语句所被标注的对话状态。通过Softmax多分类预测函数将输入语句在对话状态标注空间中的状态向量映射成由 $(0,1)$ 之间的概率值序列组合形成的概率向量，以获得输入语句所对应的对话状态概率向量。

[0073] 在步骤270中，从对话状态概率向量中选取最大概率标签所标注的对话状态作为所述输入语句的对话状态。

[0074] 其中，如前所述，对话状态概率向量中每一元素为分别将输入语句标注为对话

状态标注空间中每一对话状态的概率，输入语句则最有可能被标注为最大概率所标注的对话状态。

[0075] 例如在上述举例中，输入语句“您好，我有几个问题想请问您？”被标注被“客户寒暄”的概率最大（具体为0.85），则获得该输入语句的对话状态为“客户寒暄”。

[0076] 在一种应用场景中，客服机器人获取输入语句的对话状态后，可从自身知识库中提取对话状态为“坐席寒暄”的语句返回给客户。例如，可提取语句“您好，我是阳光，很高兴为您服务，请问有什么可以帮您？”与客户进行寒暄。同理，如果客服机器人获取输入的对话状态为“客户提问”，客服机器人则从自身知识库中提取对话状态为“坐席回答”的语句，从而回答客户提出的问题。

[0077] 需要说明的是，客服机器人自身所配置的知识库中预先存储有大量语句，这些语句预先通过上述实施例所描述的方法进行对话状态标注。因此，客服机器人从知识库中获取与输入语句语义相匹配的语句后，可相应获取该语句的对话状态。

[0078] 因此，在本实施例中，客服机器人可以根据所获取输入语句的对话状态自适应地与客户进行业务对话，增加了客服机器人的主动性。同时，也提高了用户使用客服机器人进行业务办理的体验感。

[0079] 还应当说明的是，在本实施例中，客服机器人从知识库中提取相应语句后，可通过所配置的扬声器进行语音播放，或者通过所配置的LCD屏幕对语句进行文本显示，从而与客户进行业务对话。

[0080] 并且，本实施例对输入语句进行对话状态标注的过程可以是由客服机器人所配置的处理器执行的，也可以由与客服机器人预先建立有线或者无线网络连接的服务器执行，本实施例并不对比进行限定。

[0081] 可选地，上述本申请提供的客服机器人对话状态识别方法可分别用作离线训练阶段和在线预测阶段。其中，对客服机器人进行离线训练的目的在于优化步骤253中的目标状态转移矩阵，以获得最优的目标状态转移矩阵。在线预测阶段则直接使用离线训练阶段获得的最优目标状态转移矩阵对实时输入的输入语句进行对话状态标注，输出该输入语句的最优对话状态。

[0082] 图5是另一示例性实施例示出的一种客服机器人对话状态识别方法，该方法应用于客服机器人的离线训练阶段。该方法在上述实施例的基础上还包括以下步骤：

[0083] 在步骤310中，计算输入语句的对话状态和真实对话状态之间的语句识别误差。

[0084] 在一实施例中，语句识别误差是根据交叉熵损失函数计算得出的。本实施例定义的交叉熵损失函数为：

$$[-\log(\sigma(y|X)) = s(X, y) - \log(\sum_y e^{s(X, y)})]$$

[0085] 如果计算得到的语句识别误差“ $-\log(\sigma(y|X))$ ”小于预设阈值，则表示当前对输入语句进行对话状态标注所使用的目标状态转移矩阵不是最优的。

[0086] 在步骤330中，通过最小化语句识别误差对目标状态转移矩阵进行更新。

[0087] 如前所述，如果计算所得的语句识别误差小于预设阈值，需对输入语句识别误差进行最小化。

[0088] 在一实施例中，采用梯度下降算法对语句识别误差进行最小化。具体地，对上述交叉熵损失函数进行求导运算，获得该交叉熵损失函数相对当前目标状态转移矩阵的偏导数。得到的偏导数也称为梯度值。然后将当前的目标状态转移矩阵与求得的偏导数进行减法运算，获取新的状态转移矩阵，并根据新的状态转移参数对目标参数矩阵进行更新。

[0089] 获取更新的状态转移矩阵后，仍按照步骤210至步骤270所描述的方法对输入语句进行下一次离线训练，此时采用的目标状态转移矩阵为更新的状态转移矩阵。获得输入语句的对话状态后，根据上述交叉熵损失函数计算语句识别误差。

[0090] 若获得的语句识别误差仍小于预设阈值，则重复步骤310和步骤330所描述的方法对此次离线训练采用的目标状态转移矩阵更新，并按照步骤210至步骤270所描述的方法对输入语句进行下一次离线训练，直至获得的语句识别误差大于预设阈值，则表示对输入语句进行离线训练的目标状态转移矩阵最优，此时离线训练阶段完成。

[0091] 离线训练所得的语句识别误差会穿过CRF模型到达LSTM模型中，因此，在CR

F模型的目标状态转移矩阵进行更新的过程中，LSTM模型中的训练参数也会进行更新，具体可采用BP（error Back Propagation，误差反向传播）算法进行。

[0092] 在线预测阶段则直接使用此次离线训练所使用的目标状态转移矩阵进行输入语句的在线预测，在线预测阶段输出的输入语句的对话状态可视为对输入语句进行准确标注的对话状态。

[0093] 在另一示例性实施例中，上述客服机器人对话状态识别方法还可以包括以下步骤：

[0094] 在客服机器人完成业务对话后，获取该业务对话中每一对话的对话状态；

[0095] 对该业务对话进行指定对话抽取，所抽取的指定对话对应于同一对话状态。

[0096] 其中，如前所述，由于客服机器人在进行业务对话的过程中，客服机器人对客户向其输入的语句以及从自身知识库中提取的语句均进行了对话状态标注，在业务对话完成后，客服机器人可从存储器中获取每一对话对应的对话状态。

[0097] 因此，可从客服机器人所存储的海量业务对话中提取某一种或多种对话状态所对应的对话。例如，可从客服机器人在目标时间段内产生的业务对话中提取对话状态为“客户提问”的全部对话，以分析目标时间段内客户频繁询问的问题是什么，或者抽取目标时间段内对话状态为“客户评价”的全部对话并进行分析，以根据分析结果优化客服机器人的功能。

[0098] 如图6所示，在一示例性实施例中，本申请还提供一种客服机器人对话状态识别装置，该装置包括词向量获取模块410、语义向量获取模块430、对话状态标注模块450和对话状态获取模块470。

[0099] 词向量获取模块410用于客服机器人在所进行的业务对话中通过对输入语句中词语的向量化获得所述词语对应的词向量。

[0100] 语义向量获取模块430用于通过对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量。

[0101] 对话状态标注模块450用于根据所述输入语句的语义向量对所述输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量。

[0102] 对话状态获取模块470用于从所述对话状态概率向量中选取最大概率标签所标注的对话状态作为所述输入语句的对话状态。

- [0103] 在另一示例性实施例中，词向量获取模块410可以具体包括分词处理单元和词向量获取单元。
- [0104] 分词处理单元用于通过对输入语句进行分词处理获得所述输入语句中的目标词语。
- [0105] 词向量获取单元用于根据所对应上下文信息将每一所述目标词语映射为低维向量，获得每一所述目标词语对应的词向量。
- [0106] 在另一示例性实施例中，语义向量获取模块430可进一步用于按照时间先后顺序对所述输入语句中的每一词向量进行遍历，将遍历所得的隐状态向量作为所述输入语句对应的语义向量。
- [0107] 在另一示例性实施例中，对话状态标注模块450可以具体包括向量映射单元、标注打分单元和概率归一单元。
- [0108] 向量映射单元用于对所述输入语句的语义向量进行非线性映射，获得所述输入语句在对话状态标注空间中的状态向量；
- [0109] 标注打分单元用于根据目标状态转移矩阵对所述输入语句在对话状态标注空间中的状态向量进行对话状态标注打分；
- [0110] 概率归一单元用于对所述输入语句的对话状态预测打分进行概率归一化，获得所述输入语句的对话状态概率向量。
- [0111] 在另一示例性实施例中，上述客服机器人对话状态识别装置还可以包括语句识别误差计算模块和参数更新模块。
- [0112] 识别误差计算模块用于计算所述输入语句的对话状态和真实对话状态之间的语句识别误差。
- [0113] 参数更新模块用于通过最小化所述语句识别误差对所述目标状态转移矩阵进行更新。
- [0114] 在另一示例性实施例中，上述客服机器人对话状态识别装置还可以包括业务对话获取模块和业务对话抽取模块。
- [0115] 业务对话获取模块用于在所述客服机器人完成所述业务对话后，获取所述业务对话中每一对话的对话状态。
- [0116] 业务对话抽取模块用于对所述业务会话进行指定对话抽取，所抽取的指定对话

对应于同一所述对话状态。

[0117] 需要说明的是，上述实施例所提供的装置与上述实施例所提供的方法属于同一构思，其中各个模块执行操作的具体方式已经在方法实施例中进行了详细描述，此处不再赘述。

[0118] 在一示例性实施例中，本申请还提供一种电子设备，该电子设备包括：

[0119] 处理器；

[0120] 存储器，该存储器上存储有计算机可读指令，该计算机可读指令被处理器执行时，实现如前所示的客服机器人对话状态识别方法。

[0121] 在一示例性实施例中，本申请还提供一种计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时，实现如前所示的客服机器人对话状态识别方法。

[0122] 上述内容，仅为本申请的较佳示例性实施例，并非用于限制本申请的实施方案，本领域普通技术人员根据本申请的主要构思和精神，可以十分方便地进行相应的变通或修改，故本申请的保护范围应以权利要求书所要求的保护范围为准。

权利要求书

- [权利要求 1] 一种客服机器人对话状态识别方法，包括：
客服机器人在所进行的业务对话中，通过对输入语句中词语的向量化获得所述词语对应的词向量；
通过对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量；
根据所述输入语句的语义向量对所述输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量；
从所述对话状态概率向量中选取最大概率标签所标注的对话状态作为所述输入语句的对话状态。
- [权利要求 2] 如权利要求1所述的方法，其中，所述客服机器人在所进行的业务对话中通过对输入语句中词语的向量化获得所述词语对应的词向量，包括：
通过对输入语句进行分词处理获得所述输入语句中的目标词语；
根据所对应上下文信息将每一所述目标词语映射为低维向量，获得每一所述目标词语对应的词向量。
- [权利要求 3] 如权利要求1或2所述的方法，其中，所述通过对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量，包括：
按照时间先后顺序对所述输入语句中的每一词向量进行遍历，将遍历所得的隐状态向量作为所述输入语句对应的语义向量。
- [权利要求 4] 如权利要求1至3任一项所述的方法，所述根据所述输入语句的语义向量对所述输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量，包括：
对所述输入语句的语义向量进行非线性映射，获得所述输入语句在对话状态标注空间中的状态向量；
根据目标状态转移矩阵，对所述输入语句在对话状态标注空间中的状态向量进行对话状态标注打分；
对所述输入语句的对话状态预测打分进行概率归一化，获得所述输入

语句的对话状态概率向量。

[权利要求 5]

如权利要求4所述的方法，其中，所述方法还包括：

计算所述输入语句的对话状态和真实对话状态之间的语句识别误差；
通过最小化所述语句识别误差对所述目标状态转移矩阵进行更新。

[权利要求 6]

如权利要求1至5任一项所述的方法，其中，所述方法还包括：

在所述客服机器人完成所述业务对话后，获取所述业务对话中每一对话的对话状态；

对所述业务对话进行指定对话抽取，所抽取的指定对话对应于同一所述对话状态。

[权利要求 7]

一种客服机器人对话状态识别装置，包括：

词向量获取模块，配置为客服机器人在所进行的业务对话中，通过对输入语句中词语的向量化获得所述词语对应的词向量；

语义向量获取模块，配置为通过对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量；

对话状态标注模块，配置为根据所述输入语句的语义向量对所述输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量；

对话状态获取模块，配置为从所述对话状态概率向量中选取最大概率标签所标注的对话状态作为所述输入语句的对话状态。

[权利要求 8]

如权利要求7所述的装置，其中，所述词向量获取模块包括：

分词处理单元，配置为通过对输入语句进行分词处理获得所述输入语句中的目标词语；

词向量获取单元，配置为根据所对应上下文信息将每一所述目标词语映射为低维向量，获得每一所述目标词语对应的词向量。

[权利要求 9]

如权利要求7或8所述的装置，其中，所述语义向量获取模块配置为按照时间先后顺序对所述输入语句中的每一词向量进行遍历，将遍历所得的隐状态向量作为所述输入语句对应的语义向量。

[权利要求 10]

如权利要求7至9所述的装置，其中，所述对话状态标注模块包括：

向量映射单元，配置为对所述输入语句的语义向量进行非线性映射，

获得所述输入语句在对话状态标注空间中的状态向量；

标注打分单元，配置为根据目标状态转移矩阵，对所述输入语句在对话状态标注空间中的状态向量进行对话状态标注打分；

概率归一单元，配置为对所述输入语句的对话状态预测打分进行概率归一化，获得所述输入语句的对话状态概率向量。

[权利要求 11]

如权利要求10所述的装置，其中，所述装置还包括：

识别误差计算模块，配置为计算所述输入语句的对话状态和真实对话状态之间的语句识别误差；

参数更新模块，配置为通过最小化所述语句识别误差对所述目标状态转移矩阵进行更新。

[权利要求 12]

如权利要求7至11任一项所述的装置，其中，所述装置还包括：

业务对话获取模块，配置为在所述客服机器人完成所述业务对话后，获取所述业务对话中每一对话的对话状态；

业务对话抽取模块，配置为对所述业务会话进行指定对话抽取，所抽取的指定对话对应于同一所述对话状态。

[权利要求 13]

一种电子设备，包括：

处理器；及存储器，所述存储器上存储有计算机可读指令，所述计算机可读指令被所述处理器执行时，所述处理器配置为实现以下步骤：

客服机器人在所进行的业务对话中，通过对输入语句中词语的向量化获得所述词语对应的词向量；

通过对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量；

根据所述输入语句的语义向量对所述输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量；

从所述对话状态概率向量中选取最大概率标签所标注的对话状态作为所述输入语句的对话状态。

[权利要求 14]

如权利要求13所述的电子设备，其中，所述客服机器人在所进行的业务对话中通过对输入语句中词语的向量化获得所述词语对应的词向量

，所述处理器配置为实现以下步骤：

通过对输入语句进行分词处理获得所述输入语句中的目标词语；

根据所对应上下文信息将每一所述目标词语映射为低维向量，获得每一所述目标词语对应的词向量。

[权利要求 15] 如权利要求13或14所述的电子设备，其中，所述通过对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量，所述处理器配置为实现以下步骤：

按照时间先后顺序对所述输入语句中的每一词向量进行遍历，将遍历所得的隐状态向量作为所述输入语句对应的语义向量。

[权利要求 16] 如权利要求13至15所述的电子设备，其中，所述根据所述输入语句的语义向量对所述输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量，所述处理器配置为实现以下步骤：

对所述输入语句的语义向量进行非线性映射，获得所述输入语句在对话状态标注空间中的状态向量；

根据目标状态转移矩阵，对所述输入语句在对话状态标注空间中的状态向量进行对话状态标注打分；

对所述输入语句的对话状态预测打分进行概率归一化，获得所述输入语句的对话状态概率向量。

[权利要求 17] 如权利要求16所述的电子设备，其中，所述处理器还配置为实现以下步骤：

计算所述输入语句的对话状态和真实对话状态之间的语句识别误差；

通过最小化所述语句识别误差对所述目标状态转移矩阵进行更新。

[权利要求 18] 如权利要求13至17任一项所述的电子设备，其中，所述处理器还配置为实现以下步骤：

在所述客服机器人完成所述业务对话后，获取所述业务对话中每一对话的对话状态；

对所述业务会话进行指定对话抽取，所抽取的指定对话对应于同一所述对话状态。

- [权利要求 19] 一种计算机可读存储介质，其上存储有计算机程序，其中，所述计算机程序被处理器执行时，所述处理器配置为实现以下步骤：
客服机器人在所进行的业务对话中，通过对输入语句中词语的向量化获得所述词语对应的词向量；
通过对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量；
根据所述输入语句的语义向量对所述输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量；
从所述对话状态概率向量中选取最大概率标签所标注的对话状态作为所述输入语句的对话状态。
- [权利要求 20] 如权利要求19所述的计算机可读存储介质，其中，所述客服机器人在所进行的业务对话中通过对输入语句中词语的向量化获得所述词语对应的词向量，所述处理器配置为实现以下步骤：
通过对输入语句进行分词处理获得所述输入语句中的目标词语；
根据所对应上下文信息将每一所述目标词语映射为低维向量，获得每一所述目标词语对应的词向量。
- [权利要求 21] 如权利要求19或20所述的计算机可读存储介质，其中，所述通过对所述输入语句中的每一词向量的遍历得到所述输入语句对应的语义向量，所述处理器配置为实现以下步骤：
按照时间先后顺序对所述输入语句中的每一词向量进行遍历，将遍历所得的隐状态向量作为所述输入语句对应的语义向量。
- [权利要求 22] 如权利要求19至21所述的计算机可读存储介质，其中，所述根据所述输入语句的语义向量对所述输入语句进行对话状态标注，获得所述输入语句的对话状态概率向量，所述处理器配置为实现以下步骤：
对所述输入语句的语义向量进行非线性映射，获得所述输入语句在对话状态标注空间中的状态向量；
根据目标状态转移矩阵对所述输入语句在对话状态标注空间中的状态向量进行对话状态标注打分；

对所述输入语句的对话状态预测打分进行概率归一化，获得所述输入语句的对话状态概率向量。

[权利要求 23] 如权利要求22所述的计算机可读存储介质，其中，所述处理器还配置为实现以下步骤：

计算所述输入语句的对话状态和真实对话状态之间的语句识别误差；
通过最小化所述语句识别误差对所述目标状态转移矩阵进行更新。

[权利要求 24] 如权利要求19至23任一项所述的计算机可读存储介质，其中，所述处理器还配置为实现以下步骤：

在所述客服机器人完成所述业务对话后，获取所述业务对话中每一对话的对话状态；
对所述业务会话进行指定对话抽取，所抽取的指定对话对应于同一所述对话状态。

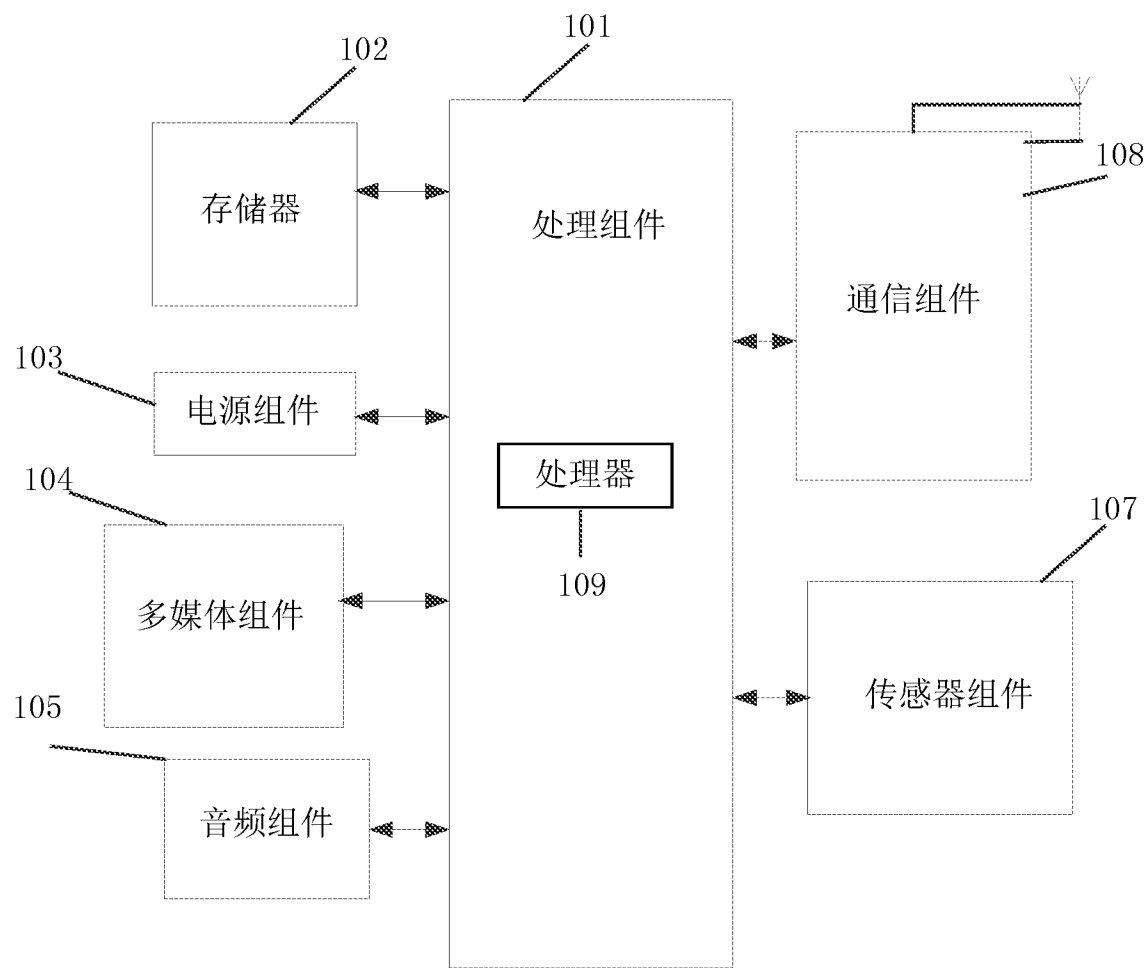


图 1

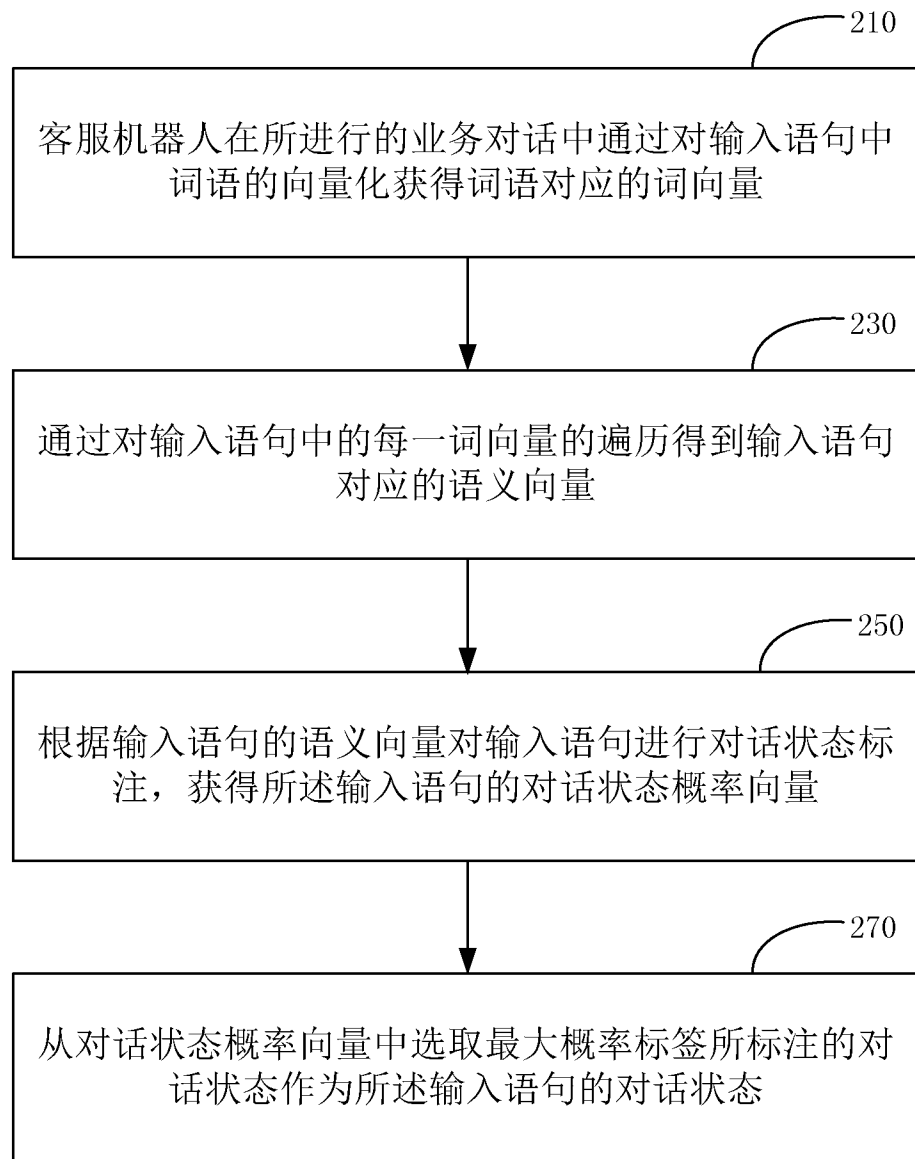


图 2

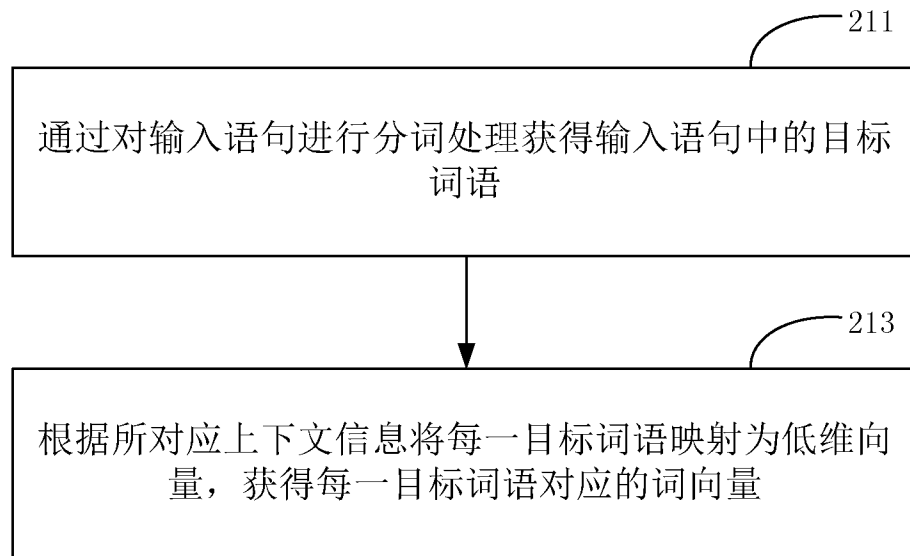


图 3

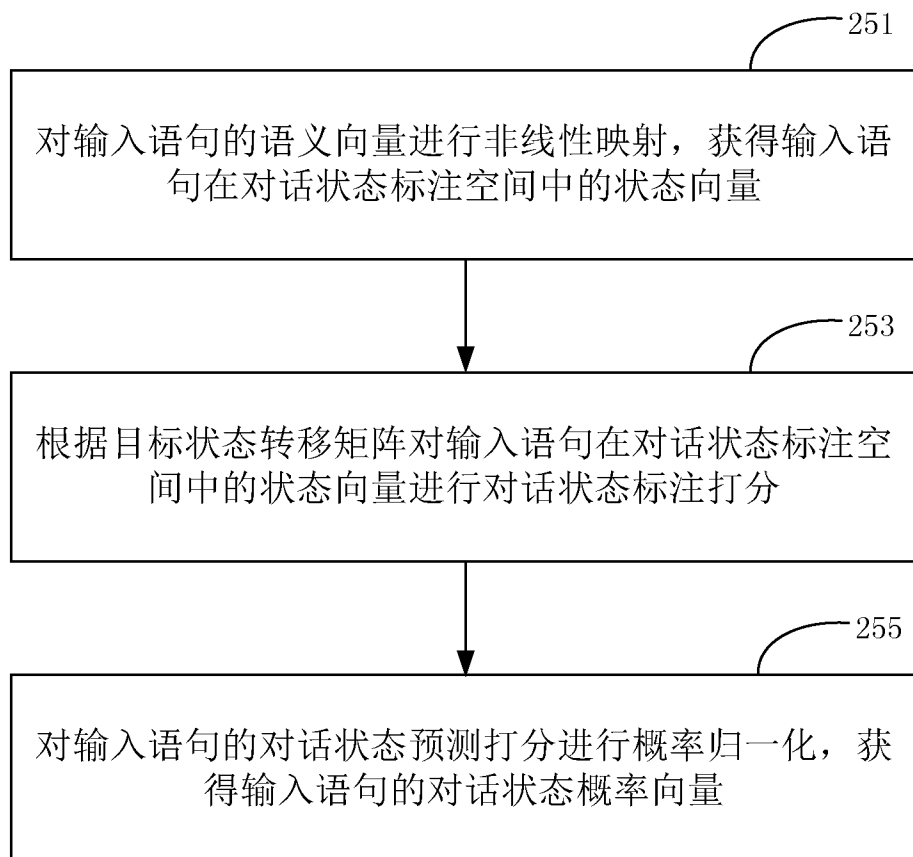


图 4

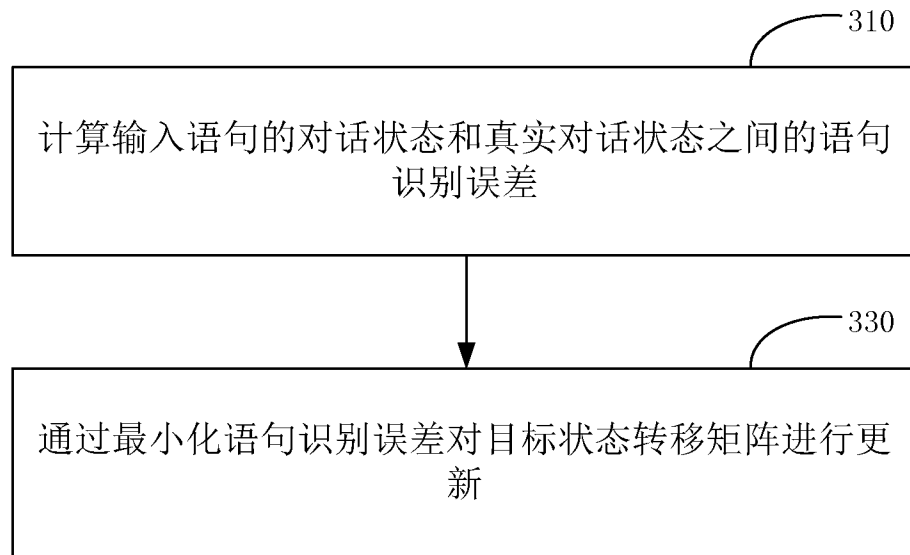


图 5

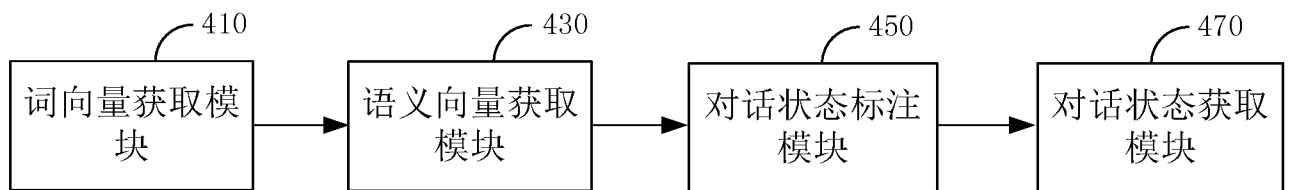


图 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/125251

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/332(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT; CNKI; WPI; EPODOC: 客服, 用户, 对话, 状态, 类型, 类别, 标签, 标记, 向量, 分词, 语义, 标注, 概率, 角色, 意图, customer service, user, conversation, state, type, category, label, tag, vector, participle, semantic, probability, role, intention

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 106503236 A (BEIJING BAIDU NETCOM SCIENCE AND TECHNOLOGY CO., LTD.) 15 March 2017 (2017-03-15) description, paragraphs [0026]-[0073]	1-24
A	CN 107679234 A (SHANGHAI CTRIP INTERNATIONAL TRAVEL AGENCY CO., LTD.) 09 February 2018 (2018-02-09) entire document	1-24
A	CN 108491433 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 04 September 2018 (2018-09-04) entire document	1-24
A	CN 108363690 A (BEIJING SHISAN TECHNOLOGY CO., LTD.) 03 August 2018 (2018-08-03) entire document	1-24

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

28 June 2019

Date of mailing of the international search report

18 July 2019

Name and mailing address of the ISA/CN

State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/125251

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106503236	A	15 March 2017	US	2018121801	A1	03 May 2018
CN	107679234	A	09 February 2018	None			
CN	108491433	A	04 September 2018	None			
CN	108363690	A	03 August 2018	None			

国际检索报告

国际申请号

PCT/CN2018/125251

A. 主题的分类 G06F 16/332(2019.01) i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNPAT;CNKI;WPI;EPDOC: 客服, 用户, 对话, 状态, 类型, 类别, 标签, 标记, 向量, 分词, 语义, 标注, 概率, 角色, 意图, customer service, user, conversation, state, type, category, label, tag, vector, participle, semantic, probability, role, intention		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 106503236 A (北京百度网讯科技有限公司) 2017年 3月 15日 (2017 - 03 - 15) 说明书第[0026]-[0073]段	1-24
A	CN 107679234 A (上海携程国际旅行社有限公司) 2018年 2月 9日 (2018 - 02 - 09) 全文	1-24
A	CN 108491433 A (平安科技深圳有限公司) 2018年 9月 4日 (2018 - 09 - 04) 全文	1-24
A	CN 108363690 A (北京十三科技有限公司) 2018年 8月 3日 (2018 - 08 - 03) 全文	1-24
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期 2019年 6月 28日		国际检索报告邮寄日期 2019年 7月 18日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 王璐 电话号码 86-(10)-53961303

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/125251

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	106503236	A	2017年 3月 15日	US 2018121801 A1	2018年 5月 3日
CN	107679234	A	2018年 2月 9日	无	
CN	108491433	A	2018年 9月 4日	无	
CN	108363690	A	2018年 8月 3日	无	