(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 114996426 A
(43)申请公布日 2022.09.02

(21)申请号 202210709686.6

(22)申请日 2022.06.22

(71)申请人 京东科技信息技术有限公司
地址 100176 北京市北京经济技术开发区科创十一街18号院2号楼6层601

(72)发明人 王培英 袁韶祖 陈蒙

(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 韩黎捷 张效荣

(51)Int.Cl.
G06F 16/332(2019.01)
G06F 16/33(2019.01)
G06F 40/35(2020.01)
G06K 9/62(2022.01)

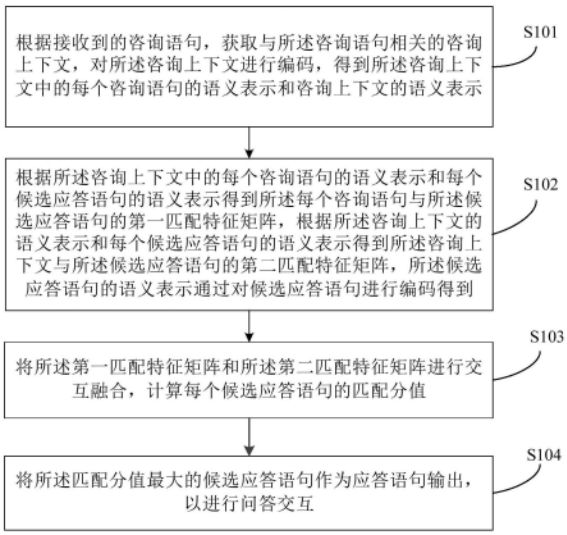
权利要求书2页 说明书13页 附图3页

(54)发明名称

一种问答交互的方法和装置

(57)摘要

本发明公开了一种问答交互的方法和装置，涉及计算机技术领域。该方法的一具体实施方式包括:根据咨询语句,获取咨询上下文,对咨询上下文进行编码,得到咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示;根据咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到第一匹配特征矩阵,根据咨询上下文的语义表示和每个候选应答语句的语义表示得到第二匹配特征矩阵;将第一匹配特征矩阵和第二匹配特征矩阵进行交互融合,计算每个候选应答语句的匹配分值;将匹配分值最大的候选应答语句作为应答语句输出。该实施方式既考虑了咨询语句层面的匹配特征,又兼顾了咨询上下文层面的匹配特征,提高了应答语句的筛选准确率。



CN 114996426 A

1.一种问答交互的方法,其特征在于,包括:

根据接收到的咨询语句,获取与所述咨询语句相关的咨询上下文,对所述咨询上下文进行编码,得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示;

根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵,根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵,所述候选应答语句的语义表示通过对候选应答语句进行编码得到;

将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合,计算每个候选应答语句的匹配分值;

将所述匹配分值最大的候选应答语句作为应答语句输出,以进行问答交互。

2.根据权利要求1所述的方法,其特征在于,对所述咨询上下文进行编码,得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示,包括:

基于所述咨询上下文中的每个咨询语句的词汇,通过词嵌入层,得到所述咨询上下文中的每个咨询语句的词向量表示;

将所述咨询上下文中的每个咨询语句的词向量表示进行拼接,得到所述咨询上下文的词向量表示;

根据所述每个咨询语句的词向量表示和所述咨询上下文的词向量表示,通过语义分析模型,分别得到所述咨询上下文中的每个咨询语句的语义表示和所述咨询上下文的语义表示。

3.根据权利要求1所述的方法,其特征在于,根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵,包括:

根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示,通过计算所述每个咨询语句和候选应答语句的词汇匹配度,得到每个咨询语句和候选应答语句的交互矩阵;

通过对所述交互矩阵进行卷积和池化操作,得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵。

4.根据权利要求1所述的方法,其特征在于,根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵,包括:

根据所述咨询上下文的语义表示和每个候选应答语句的语义表示,采用注意力机制,通过提取所述咨询上下文和每个候选应答语句之间的关键信息匹配度,得到注意力权重;

根据所述注意力权重,对所述咨询上下文进行加权表示,得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵。

5.根据权利要求1所述的方法,其特征在于,将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合,计算每个候选应答语句的匹配分值,包括:

根据所述第一匹配特征矩阵和所述第二匹配特征矩阵,计算交互融合的注意力权重矩阵;

根据所述交互融合的注意力权重矩阵,对所述第一匹配特征矩阵和所述第二匹配特征

矩阵进行交互融合，得到每个候选应答语句对应的融合特征表示；

将所述融合特征表示进行逻辑回归处理，得到每个候选应答语句的匹配分值。

6.根据权利要求5所述的方法，其特征在于，对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合之前，还包括：

使用激活函数分别对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行变换，得到第一变换匹配特征矩阵和第二变换匹配特征矩阵；

所述对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，包括：

对所述第一变换匹配特征矩阵和所述第二变换匹配特征矩阵进行交互融合。

7.根据权利要求5所述的方法，其特征在于，根据所述第一匹配特征矩阵和所述第二匹配特征矩阵，计算交互融合的注意力权重矩阵，包括：

根据所述第一匹配特征矩阵和第二匹配特征矩阵，得到二阶交互特征表示；

对所述二阶交互特征表示中的交互权重矩阵进行降秩拆分，以将所述交互权重矩阵表示为两个二维交互子权重矩阵相乘的形式；

基于所述降秩拆分后的交互权重矩阵，将所述第一匹配特征矩阵和所述第二匹配特征矩阵中的对应元素两两进行二阶交互，并进行池化处理，得到交互融合的注意力权重矩阵。

8.根据权利要求1所述的方法，其特征在于，所述方法基于预先训练的问答交互模型来实现，且所述问答交互模型在进行训练时采用交叉熵作为损失函数。

9.一种问答交互的装置，其特征在于，包括：

语义表示获取模块，用于根据接收到的咨询语句，获取与所述咨询语句相关的咨询上下文，对所述咨询上下文进行编码，得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示；

匹配特征向量获取模块，用于根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵，根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵，所述候选应答语句的语义表示通过对候选应答语句进行编码得到；

交互融合模块，用于将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值；

应答语句获取模块，用于将所述匹配分值最大的候选应答语句作为应答语句输出，以进行问答交互。

10.一种移动电子设备终端，其特征在于，包括：

一个或多个处理器；

存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求1-8中任一所述的方法。

11.一种计算机可读介质，其上存储有计算机程序，其特征在于，所述程序被处理器执行时实现如权利要求1-8中任一所述的方法。

一种问答交互的方法和装置

技术领域

[0001] 本发明涉及计算机技术领域，尤其涉及一种问答交互的方法和装置。

背景技术

[0002] 智能人机问答交互技术是指让机器理解和运用自然语言实现人机通信的技术，最近几年随着人工智能的兴起，智能人机问答交互技术的研究也越来越深入，从一问一答问答类型，到任务完成类型，再到多轮对话问答类型。对于多轮对话问答需要从给定的候选回复中，选取与问题最匹配的回复，现有的多轮对话应答选择方法可以分为两大类：一类通过将多轮对话拼接起来进行编码，然后与候选应答语句进行匹配；另外一类基于多轮对话中的每个语句与候选应答语句之间的交互特征进行建模，再对候选应答语句进行筛选。

[0003] 在实现本发明过程中，发明人发现现有技术中存在如下问题：

[0004] 目前采用的多轮对话应答语句的选择方法不是仅考虑了多轮对话上下文整体与候选应答语句之间的匹配关系，就是仅考虑了多轮对话中的每个语句与候选应答语句之间的交互特征，不能最大程度的提取多轮对话中的交互特征，影响了智能人机多轮对话问答中应答语句的准确率，不能很好的满足实际应用需要。

发明内容

[0005] 有鉴于此，本发明实施例提供一种问答交互的方法和装置，基于咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示，得到表征每个咨询语句与每个候选应答语句的第一匹配特征矩阵和表征咨询上下文与每个候选应答语句的第二匹配特征矩阵，通过将第一匹配特征矩阵和第二匹配特征矩阵进行交互融合，计算并选择候选应答语句中匹配分值最大的候选应答语句作为应答语句，实现了既考虑了咨询语句层面的匹配特征，又兼顾了咨询上下文层面的匹配特征的候选应答语句的筛选，提高了智能人机多轮对话问答中应答语句的准确率，更好的满足了实际应用。

[0006] 为实现所述目的，根据本发明实施例的一个方面，提供了一种问答交互的方法，包括：

[0007] 根据接收到的咨询语句，获取与所述咨询语句相关的咨询上下文，对所述咨询上下文进行编码，得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示；

[0008] 根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵，根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵，所述候选应答语句的语义表示通过对候选应答语句进行编码得到；

[0009] 将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值；

[0010] 将所述匹配分值最大的候选应答语句作为应答语句输出，以进行问答交互。

[0011] 可选地，对所述咨询上下文进行编码，得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示，包括：基于所述咨询上下文中的每个咨询语句的词汇，通过词嵌入层，得到所述咨询上下文中的每个咨询语句的词向量表示；将所述咨询上下文中的每个咨询语句的词向量表示进行拼接，得到所述咨询上下文的词向量表示；根据所述每个咨询语句的词向量表示和所述咨询上下文的词向量表示，通过语义分析模型，分别得到所述咨询上下文中的每个咨询语句的语义表示和所述咨询上下文的语义表示。

[0012] 可选地，根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵，包括：根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示，通过计算所述每个咨询语句和候选应答语句的词汇匹配度，得到每个咨询语句和候选应答语句的交互矩阵；通过对所述交互矩阵进行卷积和池化操作，得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵。

[0013] 可选地，根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵，包括：根据所述咨询上下文的语义表示和每个候选应答语句的语义表示，采用注意力机制，通过提取所述咨询上下文和每个候选应答语句之间的关键信息匹配度，得到注意力权重；根据所述注意力权重，对所述咨询上下文进行加权表示，得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵。

[0014] 可选地，将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值，包括：根据所述第一匹配特征矩阵和所述第二匹配特征矩阵，计算交互融合的注意力权重矩阵；根据所述交互融合的注意力权重矩阵，对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，得到每个候选应答语句对应的融合特征表示；将所述融合特征表示进行逻辑回归处理，得到每个候选应答语句的匹配分值。

[0015] 可选地，对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合之前，还包括：使用激活函数分别对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行变换，得到第一变换匹配特征矩阵和第二变换匹配特征矩阵；所述对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，包括：对所述第一变换匹配特征矩阵和所述第二变换匹配特征矩阵进行交互融合。

[0016] 可选地，根据所述第一匹配特征矩阵和所述第二匹配特征矩阵，计算交互融合的注意力权重矩阵，包括：根据所述第一匹配特征矩阵和第二匹配特征矩阵，得到二阶交互特征表示；对所述二阶交互特征表示中的交互权重矩阵进行降秩拆分，以将所述交互权重矩阵表示为两个二维交互子权重矩阵相乘的形式；基于所述降秩拆分后的交互权重矩阵，将所述第一匹配特征矩阵和所述第二匹配特征矩阵中的对应元素两两进行二阶交互，并进行池化处理，得到交互融合的注意力权重矩阵。

[0017] 可选地，所述方法基于预先训练的问答交互模型来实现，且所述问答交互模型在进行训练时采用交叉熵作为损失函数。

[0018] 根据本发明实施例的第二方面，提供一种问答交互的装置，包括：

[0019] 语义表示获取模块，用于根据接收到的咨询语句，获取与所述咨询语句相关的咨询上下文，对所述咨询上下文进行编码，得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示；

[0020] 匹配特征向量获取模块，用于根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵，根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵，所述候选应答语句的语义表示通过对候选应答语句进行编码得到；

[0021] 交互融合模块，用于将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值；

[0022] 应答语句获取模块，用于将所述匹配分值最大的候选应答语句作为应答语句输出，以进行问答交互。

[0023] 根据本发明实施例的第三方面，提供一种问答交互的电子设备，其特征在于，包括：

[0024] 一个或多个处理器；

[0025] 存储装置，用于存储一个或多个程序，

[0026] 当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现本发明实施例第一方面提供的方法。

[0027] 根据本发明实施例的第四方面，提供一种计算机可读介质，其上存储有计算机程序，所述程序被处理器执行时实现本发明实施例第一方面提供的方法。

[0028] 所述发明中的一个实施例具有如下优点或有益效果：通过根据接收到的咨询语句，获取与咨询语句相关的咨询上下文，对咨询上下文进行编码，得到咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示；根据咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到每个咨询语句与候选应答语句的第一匹配特征矩阵，根据咨询上下文的语义表示和每个候选应答语句的语义表示得到咨询上下文与候选应答语句的第二匹配特征矩阵，候选应答语句的语义表示通过对候选应答语句进行编码得到；将第一匹配特征矩阵和第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值；将匹配分值最大的候选应答语句作为应答语句输出，以进行问答交互的技术方案，实现了基于咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示，得到表征每个咨询语句与每个候选应答语句的第一匹配特征矩阵和表征咨询上下文与每个候选应答语句的第二匹配特征矩阵，通过将第一匹配特征矩阵和第二匹配特征矩阵进行交互融合，计算并选择候选应答语句中匹配分值最大的候选应答语句作为应答语句，实现了既考虑了咨询语句层面的匹配特征，又兼顾了咨询上下文层面的匹配特征的候选应答语句的筛选，提高了智能人机多轮对话问答中应答语句的准确率，更好的满足了实际应用。

附图说明

[0029] 附图用于更好地理解本发明，不构成对本发明的不当限定。其中：

[0030] 图1是根据本发明实施例的问答交互方法的主要流程的示意图；

[0031] 图2是本发明实施例问答交互模型的整体框架示意图；

[0032] 图3是根据本发明实施例的问答交互装置的主要模块示意图；

[0033] 图4是本发明实施例可以应用于其中的示例性系统架构图；

[0034] 图5是适于用来实现本发明实施例的终端设备或服务器的计算机系统的结构示意

图。

具体实施方式

[0035] 以下结合附图对本发明的示范性实施例做出说明，其中包括本发明实施例的各种细节以助于理解，应当将它们认为仅仅是示范性的。因此，本领域普通技术人员应当认识到，可以对这里描述的实施例做出各种改变和修改，而不会背离本发明的范围和精神。同样，为了清楚和简明，以下的描述中省略了对公知功能和结构的描述。

[0036] 目前采用的多轮对话应答语句的选择方法不是仅考虑了多轮对话中的每个语句与候选应答语句之间的交互特征，就是仅考虑了多轮对话上下文整体与候选应答语句之间的匹配关系，不能最大程度的提取多轮对话中的交互特征，影响了智能人机多轮对话问答中应答语句的准确率，不能很好的满足实际应用需要。

[0037] 为了解决现有技术中存在的上述问题，本发明提出一种问答交互方法，基于咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示，得到表征每个咨询语句与每个候选应答语句的第一匹配特征矩阵和表征咨询上下文与每个候选应答语句的第二匹配特征矩阵，通过将第一匹配特征矩阵和第二匹配特征矩阵进行交互融合，计算并选择候选应答语句中匹配分值最大的候选应答语句作为应答语句，实现了既考虑了咨询语句层面的匹配特征，又兼顾了咨询上下文层面的匹配特征的候选应答语句的筛选，提高了智能人机多轮对话问答中应答语句的准确率，更好的满足了实际应用。

[0038] 在本发明的实施例介绍中，所涉及的名词及其含义如下：

[0039] Transformer：是一种NLP经典模型，使用了Self-Attention机制，不采用RNN的顺序结构，使得模型可以并行化训练，而且能够拥有全局信息；

[0040] max-pooling：是CNN当中的最大值池化操作；

[0041] Adam优化器：是一种可以替代传统随机梯度下降过程的一阶优化算法，它能基于训练数据迭代地更新神经网络权重；

[0042] softmax函数：一种归一化的指数函数，用于多分类过程中，它将多个神经元的输出，映射到(0,1)区间内；

[0043] softmax层：最常用的一种通过神经网络解决多分类问题；

[0044] ES：Elasticsearch，一个基于Lucene的搜索服务器。它提供了一个分布式多用户能力的全文搜索引擎，是一种流行的企业级搜索引擎。

[0045] 图1是根据本发明实施例的问答交互方法的主要流程的示意图，如图1所示，本发明实施例的问答交互方法包括如下的步骤S101至步骤S104。

[0046] 步骤S101、根据接收到的咨询语句，获取与所述咨询语句相关的咨询上下文，对所述咨询上下文进行编码，得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示。

[0047] 具体地，在智能人机多轮问答交互中，由于咨询语句可能是片段化的，也有可能随着对话的进行发生变化和修改，所以需要根据接收到的咨询语句，获取与咨询语句相关的咨询上下文。系统根据当前接收到的咨询语句，将本次多轮问答交互中在咨询语句之前的接收到的所有咨询语句和当前接收到的咨询语句组成咨询上下文。

[0048] 根据本发明的一个实施例，对所述咨询上下文进行编码，得到所述咨询上下文中

的每个咨询语句的语义表示和咨询上下文的语义表示，包括：基于所述咨询上下文中的每个咨询语句的词汇，通过词嵌入层，得到所述咨询上下文中的每个咨询语句的词向量表示；将所述咨询上下文中的每个咨询语句的词向量表示进行拼接，得到所述咨询上下文的词向量表示；根据所述每个咨询语句的词向量表示和所述咨询上下文的词向量表示，通过语义分析模型，分别得到所述咨询上下文中的每个咨询语句的语义表示和所述咨询上下文的语义表示。

[0049] 具体地，设定上述的咨询上下文用$s=\{s_k\}_{k=1}^{m}$来表示，而$s_k=\{w_1,w_2,\cdots,w_n\}$，其中$s_k$表示咨询上下文中的第k个咨询语句，$w_n$表示在$s_k$咨询语句中的第n个词汇。对咨询上下文s中的每个咨询语句，通过词嵌入层，得到每个咨询语句的词向量表示，设定其中一个咨询语句的词向量表示为$E_S=[e_1,e_2\cdots e_n]$；根据每个咨询语句的词向量表示，将咨询上下文中的每个咨询语句的词向量表示进行拼接，设定咨询上下文中共计L个词汇，得到咨询上下文的词向量表示$E_C=[e_1,e_2\cdots e_L]$；根据每个咨询语句的词向量表示和咨询上下文的词向量表示，通过基于自注意力机制的Transformer语义分析模型，得到每个咨询语句的语义表示和咨询上下文的语义表示，以上述的一个咨询语句的词向量表示E_S为例，其语义表示S为：

[0050] $S=\text{Transformer}_{\text{Encoder}}(E_S)$，

[0051] 相应的，咨询上下文的语义表示C为：

[0052] $C=\text{Transformer}_{\text{Encoder}}(E_C)$。

[0053] 步骤S102、根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵，根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵，所述候选应答语句的语义表示通过对候选应答语句进行编码得到。

[0054] 具体地，根据上述咨询上下文，可以通过ES检索出多个候选应答语句，将每个候选应答语句按照上述的编码方法进行编码，得到候选应答语句的语义表示，以其中的一个候选应答语句r为例，$r=\{w_1,w_2,\cdots,w_q\}$，w_p表示在候选应答语句r中的第q个词汇，在获取到此候选应答语句的词向量表示$E_R=[e_1,e_2\cdots e_q]$后，通过基于自注意力机制的Transformer语义分析模型，得到此候选应答语句的语义表示R为：

[0055] $R=\text{Transformer}_{\text{Encoder}}(E_R)$。

[0056] 根据本发明的一个实施例，根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵，包括：根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示，通过计算所述每个咨询语句和候选应答语句的词汇匹配度，得到每个咨询语句和候选应答语句的交互矩阵；通过对所述交互矩阵进行卷积和池化操作，得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵。

[0057] 具体地，根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示，以候选应答语句中的一个候选应答语句为匹配对象，在计算所述咨询上下文中每个咨询语句和匹配对象的词汇匹配度后，按照上述每个咨询语句和匹配对象的词汇匹配度计算的方法，遍历候选应答语句中的其它候选应答语句，计算所述每个咨询语句和候选应答语句的词汇匹配度，进而得到每个咨询语句和候选应答语句的交互矩阵；在此基础

上，通过对所述交互矩阵进行卷积和池化操作，将提取到的匹配特征进行拼接和投影，得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵。

[0058] 示例性地，以候选应答语句中的一个候选应答语句为匹配对象，根据上述的一个咨询语句的语义表示S和一个候选应答语句的语义表示R，计算两个句子相应位置上的词汇匹配度为：

[0059] $M_{ij}=S_{i\cdot}^{T}R_{j\cdot}$，

[0060] 其中$S_{i\cdot}$表示此咨询语句中的第i个词的语义表示，$R_{j\cdot}$表示此候选应答语句的第j个词的语义表示，由此得到此咨询语句和此候选应答语句之间的交互矩阵M，按照此方法，可以得到每个咨询语句和此候选应答语句的交互矩阵，进而得到每个咨询语句和所有候选应答语句的交互矩阵；对于上述交互矩阵M，通过在这个交互矩阵上进行卷积以及池化操作，提取此咨询语句和词候选应答语句之间的关键匹配信息，将池化操作中max-pooling的匹配特征进行拼接，并投影到低维匹配特征空间，得到第i个咨询语句与此候选应答语句之间的第一匹配特征向量h_t^i，可以用卷积神经网络表示为：

[0061] $$h_t^i = CNN(M)。$$

[0062] 按照上述相似的方式，依次对咨询上下文中剩余的m-1个咨询语句计算第一匹配特征向量，得到每个咨询语句与此候选应答语句之间的第一匹配特征矩阵$H_t = \{h_t^1, \ldots, h_t^m\}]$。对于所有的候选应答语句，与此相似，通过对交互矩阵进行卷积和池化操作，计算每个咨询语句与所述候选应答语句的第一匹配特征矩阵，进而得到每个咨询语句与所有候选应答语句的第一匹配特征矩阵的集合。

[0063] 根据本发明的另一个实施例，根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵，包括：根据所述咨询上下文的语义表示和每个候选应答语句的语义表示，采用注意力机制，通过提取所述咨询上下文和每个候选应答语句之间的关键信息匹配度，得到注意力权重；根据所述注意力权重，对所述咨询上下文进行加权表示，得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵。

[0064] 具体地，由于咨询上下文的规模一般相对大，对于咨询上下文和候选应答语句的匹配度提取优选采用注意力机制，提取关键信息的匹配度，根据咨询上下文的语义表示和每个候选应答语句的语义表示，采用注意力机制，提取咨询上下文和每个候选应答语句之间的关键信息匹配度，得到注意力权重，例如，对于一个候选应答语句response，咨询上下文context和response之间的关键信息匹配度$d_{i,j}$为：

[0065] $d_{i,j}=\tanh(C_{j\cdot}^{T}WR_{i\cdot}+b)$，

[0066] 其中$C_{j\cdot}$为咨询上下文context中第j个词的语义表示，$R_{i\cdot}$为候选应答语句response中第i个词的语义表示，W为权重矩阵，b为偏置项，相应地，其注意力权重$\alpha_{i,j}$为：

[0067] $$\alpha_{i,j} = \frac{\exp(d_{i,j})}{\sum_j \exp(d_{i,j})}。$$

[0068] 类似的，按照此方法，遍历所有的候选应答语句，提取咨询上下文和每个候选应答语句之间的关键信息匹配度，得到注意力权重；根据上述注意力权重，对咨询上下文context进行加权表示，得到咨询上下文context与所述候选应答语句response的第二匹配

特征矩阵，进而得到咨询上下文context与所有候选应答语句的第二匹配特征矩阵集合。例如上述的一个候选应答语句response，咨询上下文context可以被表示为各个词的加权表示$\sum_j \alpha_{i,j} C_j$.，得到咨询上下文context与候选应答语句response的第i个词的第二匹配特征向量h_r^i：$h_r^i = \left(\sum_j \alpha_{i,j} C_j\right) \circ R_i$.，其中，∘表示对应元素相乘，最后得到咨询上下文context与候选应答语句response中每个词的第二匹配特征矩阵$H_r = \{h_r^1, \ldots, h_r^n\}]$，n表示response中词汇的总个数。类似的，按照这种方法，遍历所有的候选应答语句，得到咨询上下文context与所有候选应答语句的第二匹配特征矩阵集合。

[0069] 通过上述第一匹配特征矩阵和第二匹配特征矩阵的计算，既得到了咨询上下文和候选应答语句之间语句层面的匹配特征，又得到了咨询上下文和候选应答语句之间更细粒度的词汇匹配特征，实现了咨询语句和候选应答语句多层面的匹配特征提取。

[0070] 步骤S103、将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值。

[0071] 具体地，候选应答语句中的每个候选应答语句均对应着第一匹配特征矩阵和第二匹配特征矩阵，将每个候选应答语句的第一匹配特征矩阵和第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值。

[0072] 根据本发明的一个实施例，将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值，包括：根据所述第一匹配特征矩阵和所述第二匹配特征矩阵，计算交互融合的注意力权重矩阵；根据所述交互融合的注意力权重矩阵，对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，得到每个候选应答语句对应的融合特征表示；将所述融合特征表示进行逻辑回归处理，得到每个候选应答语句的匹配分值。

[0073] 根据本发明的另一个实施例，根据所述第一匹配特征矩阵和所述第二匹配特征矩阵，计算交互融合的注意力权重矩阵，包括：根据所述第一匹配特征矩阵和第二匹配特征矩阵，得到二阶交互特征表示；对所述二阶交互特征表示中的交互权重矩阵进行降秩拆分，以将所述交互权重矩阵表示为两个二维交互子权重矩阵相乘的形式；基于所述降秩拆分后的交互权重矩阵，将所述第一匹配特征矩阵和所述第二匹配特征矩阵中的对应元素两两进行二阶交互，并进行池化处理，得到交互融合的注意力权重矩阵。

[0074] 示例性地，对上述第二匹配特征矩阵中的第i个词汇的第二匹配特征向量h_r^i以及第一匹配特征矩阵中的第j个咨询语句的第一匹配特征向量h_t^j，其二阶交互特征表示为：

[0075] $$f_i = {h_r^i}^T V_i h_t^j;$$

[0076] 其中$\mathrm{h_r^i} \in \mathbb{R}^{\mathrm{d_w}}$，$\mathrm{h_t^j} \in \mathbb{R}^{\mathrm{d_c}}$，$\mathrm{V_i}$是交互权重矩阵$\mathrm{V} \in \mathbb{R}^{\mathrm{K} \times \mathrm{d_w} \times \mathrm{d_c}}$的第i个矩阵；为了减少参数量，对交互权重矩阵V进行降秩拆分，使用两个

[0077] 低秩的二维矩阵$\mathrm{U} \in \mathbb{R}^{d_w \times s}$、$\mathrm{V} \in \mathbb{R}^{d_c \times s}$的相乘来代替，其中$\mathrm{s} \leqslant \min(\mathrm{d_w}, \mathrm{d_c})$；基于降秩拆分后的交互权重矩阵，将第二匹配特征矩阵中的第i个词汇的第二匹配特征向量h_r^i以及第一匹配特征矩阵中的第j个咨询语句的第一匹配特征向量h_t^j中的对应元素两两相

乘，使用池化矩阵$\mathrm{P}\in\mathbb{R}^{o\times s}$去达到同样产生二阶交互特征向量$\mathbf{f}\in\mathbb{R}^{o}$的效果，得到：

[0078] $$\mathbf{f}=P^{T}\left(U^{T}h_{r}^{i}\circ V^{T}h_{t}^{j}\right)。$$

[0079] 根据上述二阶交互特征向量的表示，可以得到交互融合的注意力权重矩阵A中的每个元素为：

[0080] $$\mathrm{A_{ij}}=\mathrm{p^{T}}\left(\mathrm{U^{T}h_{r}^{i}}\circ\mathrm{V^{T}h_{t}^{j}}\right);$$

[0081] 其中$\mathrm{p}\in\mathbb{R}^{s}$，得到标量$\mathrm{A_{ij}}$，为了便于后续的运算效率，根据元素$\mathrm{A_{ij}}$，可以得到交互融合的注意力权重矩阵A的计算方法：

[0082] $$\mathrm{H_{r}'}=(1\cdot\mathrm{p^{T}})\circ\mathrm{u}(\mathrm{H_{r}};\mathrm{U},\mathrm{b_{U}})$$

[0083] $\mathrm{H'_{t}}=\mathrm{u}(\mathrm{H_{t}};\mathrm{V},\mathrm{b_{v}})$

[0084] $$\mathrm{A}=\mathrm{g}(\mathrm{H_{r}'H_{t}'^{T}})$$

[0085] 其中，$\mathbf{1}\in\mathbb{R}^{n}$，g函数表示逐元素的softmax，u函数为：$\mathrm{u}(\mathrm{X};\theta,\mathrm{b_{0}})=\sigma(\mathrm{X}\theta+\mathrm{b_{0}})$，σ表示激活函数，$\mathrm{b_{0}}$表示偏置项。

[0086] 基于上述计算得到的交互融合的注意力权重矩阵A，对第一匹配特征矩阵和第二匹配特征矩阵进行交互融合，得到每个候选应答语句对应的融合特征表示。

[0087] 根据本发明的再一个实施例，对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合之前，还包括：使用激活函数分别对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行变换，得到第一变换匹配特征矩阵和第二变换匹配特征矩阵；所述对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，包括：对所述第一变换匹配特征矩阵和所述第二变换匹配特征矩阵进行交互融合。

[0088] 具体地，在对第一匹配特征矩阵和第二匹配特征矩阵进行交互融合之前，使用激活函数u函数，对第一匹配特征矩阵$\mathrm{H_{t}}$进行变换，得到第一变换匹配特征矩阵$\mathrm{H''_{t}}$：

[0089] $\mathrm{H''_{t}}=\mathrm{u}(\mathrm{H_{t}};\mathrm{V1'},\mathrm{b'_{V}})$

[0090] 其中V1′为激活函数u函数的参数，$\mathrm{b'_{V}}$为偏置项，$H_{t}''\in\mathbb{R}^{m\times s}$，同样对第二匹配特征矩阵$\mathrm{H_{r}}$进行变换，得到第二变换匹配特征矩阵$\mathrm{H''_{r}}$：

[0091] $\mathrm{H''_{r}}=\mathrm{u}(\mathrm{H_{r}};\mathrm{U1'},\mathrm{b'_{U}})$；

[0092] 其中U1′为激活函数u函数的参数，$\mathrm{b'_{U}}$为偏置项，$H_{r}''\in\mathbb{R}^{n\times s}$。

[0093] 根据上述计算得到的交互融合的注意力权重矩阵A，结合第一变换匹配特征矩阵和第二变换匹配特征矩阵，对第一变换匹配特征矩阵和第二变换匹配特征矩阵进行交互融合。

[0094] 示例性地，对第一变换匹配特征矩阵$\mathrm{H''_{t}}$和第二变换匹配特征矩阵$\mathrm{H''_{r}}$，取第i维特征也就是$\mathrm{H''_{t}}$和$\mathrm{H''_{r}}$的第i个列向量：$h_{r}^{i}''\in\mathbb{R}^{n\times1}$，$h_{t}^{i}''\in\mathbb{R}^{m\times1}$，根据交互融合的注意力权重矩阵A，进行二阶交互融合，即：

[0095] $$f_{i}'=(h_{r}^{i}'')^{\mathrm{T}}A\,h_{t}^{i}'';$$

[0096] 其中，$\mathrm{f'_{i}}$是个标量，为了使其变为向量，仍旧引入池化矩阵$\mathrm{P}\in\mathbb{R}^{o\times s}$，那么对每一维特征如上操作后，得到上述候选应答语句response对应的融合特征表示：

[0097] $f^F = P^T f'$；

[0098] 其中，$f' \in \mathbb{R}^s$，f'_i为f^F中的一个元素，$f^F \in \mathbb{R}^o$。按照上述方法，对每个候选应答语句对应的第一匹配特征矩阵和第二匹配特征矩阵进行二阶交互，得到每个候选应答语句对应的融合特征表示。

[0099] 基于上述每个候选应答语句对应的融合特征表示，将融合特征表示输入到softmax层，进行逻辑回归处理，得到每个候选应答语句的匹配分值，例如基于上述候选应答语句response的融合特征表示f^F，其匹配得分z为：

[0100] $z = \mathrm{softmax}(W_l f^F)$；

[0101] 其中W_l为运算参数。

[0102] 步骤S104、将所述匹配分值最大的候选应答语句作为应答语句输出，以进行问答交互。

[0103] 根据本发明的一个实施例，所述方法基于预先训练的问答交互模型来实现，且所述问答交互模型在进行训练时采用交叉熵作为损失函数。

[0104] 具体地，本发明所述的方法的实现是基于预先训练的问答交互模型，按照上述问答交互的方法，建立问答交互模型，采用交叉熵函数作为损失函数L：

[0105] $L = CE(p,y)$；

[0106] 其中，CE表示交叉熵函数，y为真实标签，使用基于Adam的优化器进行梯度下降训练。

[0107] 通过上述咨询上下文和候选应答语句之间语句层面的第一匹配特征矩阵、咨询上下文和候选应答语句之间更细粒度的词汇层面的第二匹配特征矩阵，使得建立的问答交互模型兼顾了多个层面的匹配信息，基于多层面的匹配特征，采用二阶交互融合的方法，有效融合了不同体层面的匹配特征，使得预测的匹配分值更为精确，提高了多轮对话问答中应答语句的准确率。

[0108] 图2是本发明实施例问答交互模型的整体框架示意图，图中以一个候选应答语句response为例，根据咨询语句获取咨询上下文，咨询上下文由咨询语句1咨询语句2和咨询语句3组成，将候选应答语句response、咨询语句1、咨询语句2、咨询语句3和3个咨询语句拼接组成的咨询上下文进行编码，得到各自的语义表示，如图中的空心圆形；分别计算咨询语句1、2、3与候选应答语句response的匹配度，得到第一匹配特征矩阵$\mathbf{h}_t^1$、$\mathbf{h}_t^2$、$\mathbf{h}_t^3$，如图中的深色实心圆形，计算咨询上下文与候选应答语句response的匹配度，得到第二匹配特征矩阵$\mathbf{h}_r^1$、$\mathbf{h}_r^2$、$\mathbf{h}_r^3$，如图中的浅色实心圆形；将第一匹配特征矩阵进行组合，得到$H_t = \{h_t^1, h_t^2, h_t^3\}$，第二匹配特征矩阵进行组合，得到$H_r = \{h_r^1, h_r^2, h_r^3\}$，将$H_t$和$H_r$的二阶交互特征表示中的交互权重矩阵降秩拆分为V和U两个低秩的二维矩阵，通过$H_t V$与$H_r U$的二阶交互计算，得到交互融合的注意力权重矩阵A，使用激活函数对H_t和H_r进行变换，得到$H_t V'$和$H_r U'$，对变换后的$H_t V'$和$H_r U'$，结合交互融合的注意力权重矩阵A，进行二阶交互融合，得到候选应答语句response对应的融合特征表示，最后将融合特征表示输入到softmax层，预测最终的匹配得分。

[0109] 图3是根据本发明实施例的问答交互装置的主要模块示意图。如图3所示，问答交互装置300主要包括语义表示获取模块301、匹配特征向量获取模块302、交互融合模块303

和应答语句获取模块304。

[0110] 语义表示获取模块301，用于根据接收到的咨询语句，获取与所述咨询语句相关的咨询上下文，对所述咨询上下文进行编码，得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示；

[0111] 匹配特征向量获取模块302，用于根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵，根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵，所述候选应答语句的语义表示通过对候选应答语句进行编码得到；

[0112] 交互融合模块303，用于将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值；

[0113] 应答语句获取模块304，用于将所述匹配分值最大的候选应答语句作为应答语句输出，以进行问答交互。

[0114] 具体地，所述语义表示获取模块301还可以用于：基于所述咨询上下文中的每个咨询语句的词汇，通过词嵌入层，得到所述咨询上下文中的每个咨询语句的词向量表示；将所述咨询上下文中的每个咨询语句的词向量表示进行拼接，得到所述咨询上下文的词向量表示；根据所述每个咨询语句的词向量表示和所述咨询上下文的词向量表示，通过语义分析模型，分别得到所述咨询上下文中的每个咨询语句的语义表示和所述咨询上下文的语义表示。

[0115] 具体地，所述匹配特征向量获取模块302还可以用于：根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示，通过计算所述每个咨询语句和候选应答语句的词汇匹配度，得到每个咨询语句和候选应答语句的交互矩阵；通过对所述交互矩阵进行卷积和池化操作，得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵。

[0116] 具体地，所述匹配特征向量获取模块302还可以用于：根据所述咨询上下文的语义表示和每个候选应答语句的语义表示，采用注意力机制，通过提取所述咨询上下文和每个候选应答语句之间的关键信息匹配度，得到注意力权重；根据所述注意力权重，对所述咨询上下文进行加权表示，得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵。

[0117] 具体地，所述交互融合模块303还可以用于：根据所述第一匹配特征矩阵和所述第二匹配特征矩阵，计算交互融合的注意力权重矩阵；根据所述交互融合的注意力权重矩阵，对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，得到每个候选应答语句对应的融合特征表示；将所述融合特征表示进行逻辑回归处理，得到每个候选应答语句的匹配分值。

[0118] 具体地，所述问答交互装置300还可以包括变换模块(图中未示出)，用于：对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合之前，使用激活函数分别对所述第一匹配特征矩阵和所述第二匹配特征矩阵进行变换，得到第一变换匹配特征矩阵和第二变换匹配特征矩阵；所述交互融合模块303还可以用于：对所述第一变换匹配特征矩阵和所述第二变换匹配特征矩阵进行交互融合。

[0119] 具体地，所述交互融合模块303还可以用于：根据所述第一匹配特征矩阵和第二匹

配特征矩阵，得到二阶交互特征表示；对所述二阶交互特征表示中的交互权重矩阵进行降秩拆分，以将所述交互权重矩阵表示为两个二维交互子权重矩阵相乘的形式；基于所述降秩拆分后的交互权重矩阵，将所述第一匹配特征矩阵和所述第二匹配特征矩阵中的对应元素两两进行二阶交互，并进行池化处理，得到交互融合的注意力权重矩阵。

[0120] 具体地，所述问答交互装置300基于预先训练的问答交互模型来实现，且所述问答交互模型在进行训练时采用交叉熵作为损失函数。

[0121] 图4是本发明实施例可以应用于其中的示例性系统架构图。

[0122] 如图4所示，系统架构400可以包括终端设备401、402、403，网络404和服务器405。网络404用以在终端设备401、402、403和服务器405之间提供通信链路的介质。网络404可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

[0123] 用户可以使用终端设备401、402、403通过网络404与服务器405交互，以接收或发送消息等。终端设备401、402、403上可以安装有各种通讯客户端应用，例如问答交互应用、人机对话应用等(仅为示例)。

[0124] 终端设备401、402、403可以是具有显示屏并且支持网页浏览的各种电子设备，包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。

[0125] 服务器405可以是提供各种服务的服务器，例如对用户利用终端设备401、402、403所进行的问答交互提供支持的后台管理服务器(仅为示例)。后台管理服务器可以根据接收到的咨询语句，获取与所述咨询语句相关的咨询上下文，对所述咨询上下文进行编码，得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示；根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵，根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵，所述候选应答语句的语义表示通过对候选应答语句进行编码得到；将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值；将所述匹配分值最大的候选应答语句作为应答语句输出，以进行问答交互等处理，并将处理结果(例如应答语句等--仅为示例)反馈给终端设备。

[0126] 需要说明的是，本发明实施例所提供的问答交互的方法一般由服务器405执行，相应地，问答交互的装置一般设置于服务器405中。

[0127] 应该理解，图4中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务器。

[0128] 下面参考图5，其示出了适于用来实现本发明实施例的终端设备或服务器的计算机系统500的结构示意图。图5示出的终端设备或服务器仅仅是一个示例，不应对本发明实施例的功能和使用范围带来任何限制。

[0129] 如图5所示，计算机系统500包括中央处理单元(CPU)501，其可以根据存储在只读存储器(ROM)502中的程序或者从存储部分508加载到随机访问存储器(RAM)503中的程序而执行各种适当的动作和处理。在RAM 503中，还存储有系统500操作所需的各种程序和数据。CPU 501、ROM 502以及RAM 503通过总线504彼此相连。输入/输出(I/O)接口505也连接至总线504。

[0130] 以下部件连接至I/O接口505：包括键盘、鼠标等的输入部分506；包括诸如阴极射

线管(CRT)、液晶显示器(LCD)等以及扬声器等的输出部分507;包括硬盘等的存储部分508;以及包括诸如LAN卡、调制解调器等的网络接口卡的通信部分509。通信部分509经由诸如因特网的网络执行通信处理。驱动器510也根据需要连接至I/O接口505。可拆卸介质511,诸如磁盘、光盘、磁光盘、半导体存储器等等,根据需要安装在驱动器510上,以便于从其上读出的计算机程序根据需要被安装入存储部分508。

[0131] 特别地,根据本发明公开的实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本发明公开的实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信部分509从网络上被下载和安装,和/或从可拆卸介质511被安装。在该计算机程序被中央处理单元(CPU)501执行时,执行本发明的系统中限定的上述功能。

[0132] 需要说明的是,本发明所示的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是所述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者所述的任意合适的组合。在本发明中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本发明中,计算机可读的信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或所述的任意合适的组合。计算机可读的信号介质还可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:无线、电线、光缆、RF等等,或者所述的任意合适的组合。

[0133] 附图中的流程图和框图,图示了按照本发明各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分,所述模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意,在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个接连地表示的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的是,框图或流程图中的每个方框、以及框图或流程图中的方框的组合,可以用执行规定的功能或操作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0134] 描述于本发明实施例中所涉及到的单元可以通过软件的方式实现,也可以通过硬件的方式来实现。所描述的单元也可以设置在处理器中,例如,可以描述为:一种处理器包括:语义表示获取模块、匹配特征向量获取模块、交互融合模块和快照模块。

[0135] 其中,这些模块的名称在某种情况下并不构成对该模块本身的限定,例如,交互融

合模块还可以被描述为“用于将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值的模块”。

[0136] 另一方面，本发明还提供了一种计算机可读介质，该计算机可读介质可以是所述实施例中描述的设备中所包含的；也可以是单独存在，而未装配入该设备中。所述计算机可读介质承载有一个或者多个程序，当所述一个或者多个程序被一个该设备执行时，使得该设备包括：根据接收到的咨询语句，获取与所述咨询语句相关的咨询上下文，对所述咨询上下文进行编码，得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示；根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵，根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵，所述候选应答语句的语义表示通过对候选应答语句进行编码得到；将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值；将所述匹配分值最大的候选应答语句作为应答语句输出，以进行问答交互。

[0137] 根据本发明实施例的技术方案，具有如下优点或有益效果：通过根据接收到的咨询语句，获取与咨询语句相关的咨询上下文，对咨询上下文进行编码，得到咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示；根据咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到每个咨询语句与候选应答语句的第一匹配特征矩阵，根据咨询上下文的语义表示和每个候选应答语句的语义表示得到咨询上下文与候选应答语句的第二匹配特征矩阵，候选应答语句的语义表示通过对候选应答语句进行编码得到；将第一匹配特征矩阵和第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值；将匹配分值最大的候选应答语句作为应答语句输出，以进行问答交互的技术方案，实现了基于咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示，得到表征每个咨询语句与每个候选应答语句的第一匹配特征矩阵和表征咨询上下文与每个候选应答语句的第二匹配特征矩阵，通过将第一匹配特征矩阵和第二匹配特征矩阵进行交互融合，计算并选择候选应答语句中匹配分值最大的候选应答语句作为应答语句，实现了既考虑了咨询语句层面的匹配特征，又兼顾了咨询上下文层面的匹配特征的候选应答语句的筛选，提高了智能人机多轮对话问答中应答语句的准确率，更好的满足了实际应用。

[0138] 所述具体实施方式，并不构成对本发明保护范围的限制。本领域技术人员应该明白的是，取决于设计要求和其他因素，可以发生各种各样的修改、组合、子组合和替代。任何在本发明的精神和原则之内所作的修改、等同替换和改进等，均应包含在本发明保护范围之内。

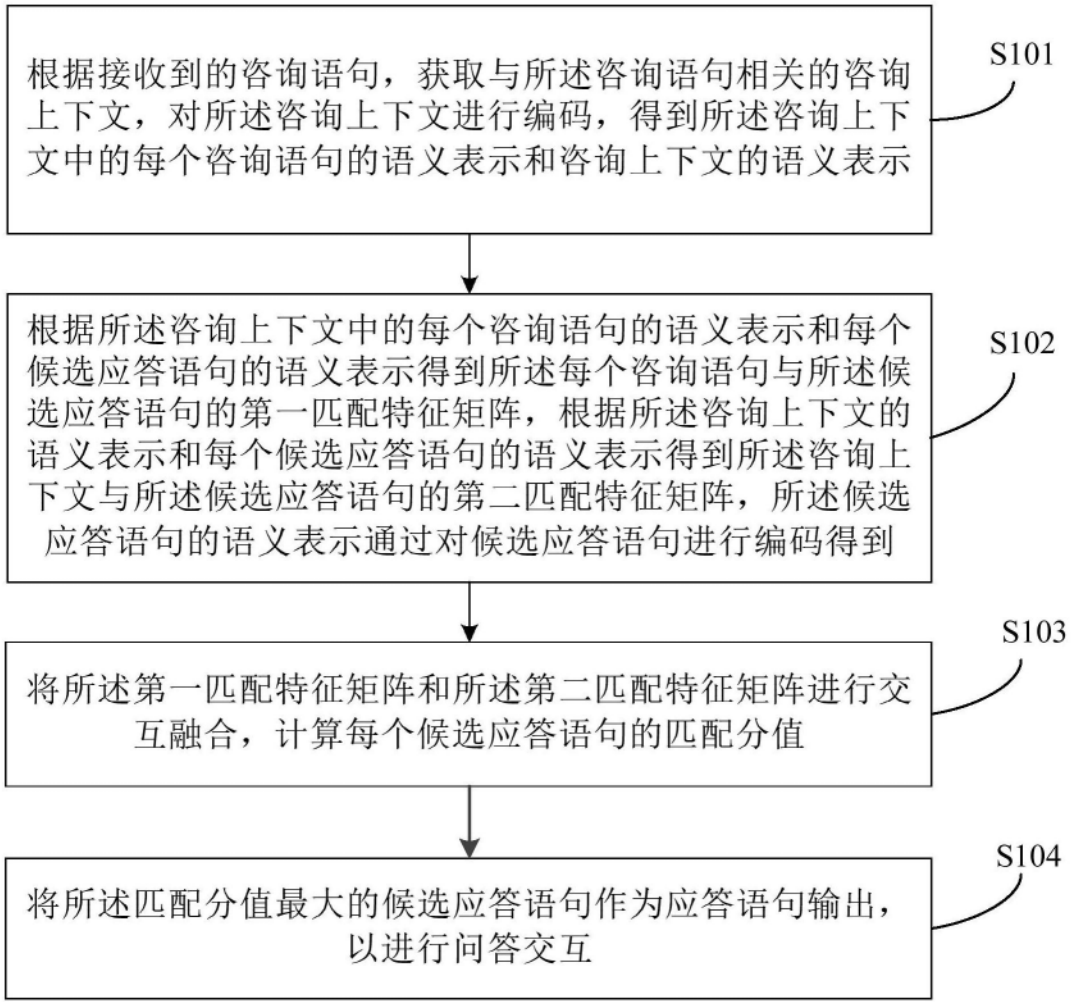

根据接收到的咨询语句，获取与所述咨询语句相关的咨询上下文，对所述咨询上下文进行编码，得到所述咨询上下文中的每个咨询语句的语义表示和咨询上下文的语义表示
S101
根据所述咨询上下文中的每个咨询语句的语义表示和每个候选应答语句的语义表示得到所述每个咨询语句与所述候选应答语句的第一匹配特征矩阵，根据所述咨询上下文的语义表示和每个候选应答语句的语义表示得到所述咨询上下文与所述候选应答语句的第二匹配特征矩阵，所述候选应答语句的语义表示通过对候选应答语句进行编码得到
S102
将所述第一匹配特征矩阵和所述第二匹配特征矩阵进行交互融合，计算每个候选应答语句的匹配分值
S103
将所述匹配分值最大的候选应答语句作为应答语句输出，以进行问答交互
S104

图1

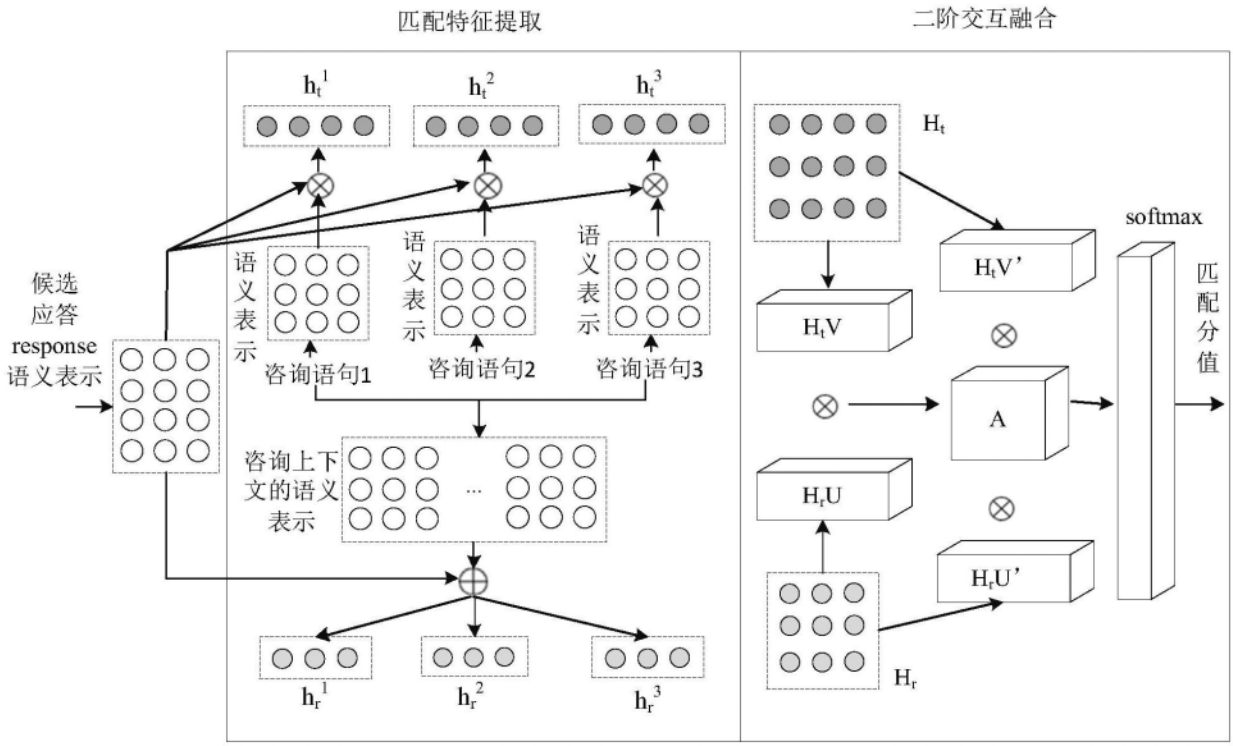

匹配特征提取
二阶交互融合
候选
应答
response
语义表示
语义表示
语义表示
语义表示
咨询语句1
咨询语句2
咨询语句3
咨询上下文的语义表示
softmax
匹配分值
A

图2

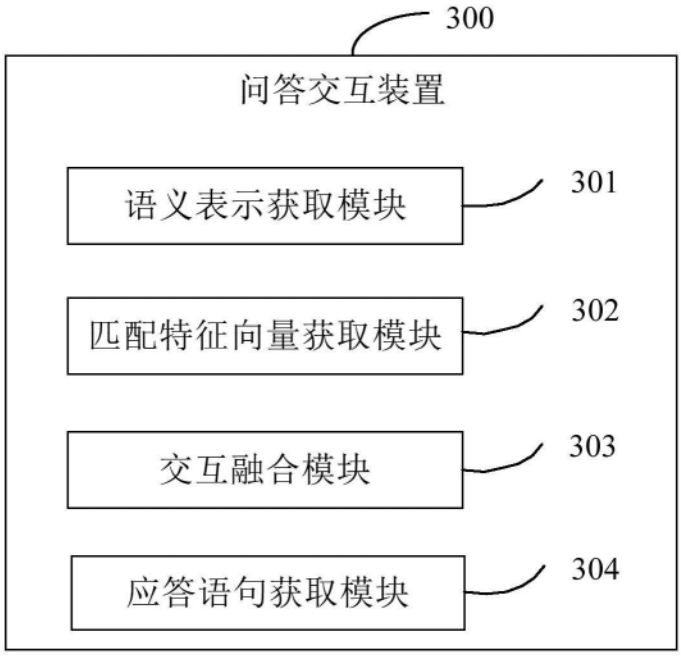

300
问答交互装置
语义表示获取模块
301
匹配特征向量获取模块
302
交互融合模块
303
应答语句获取模块
304

图3

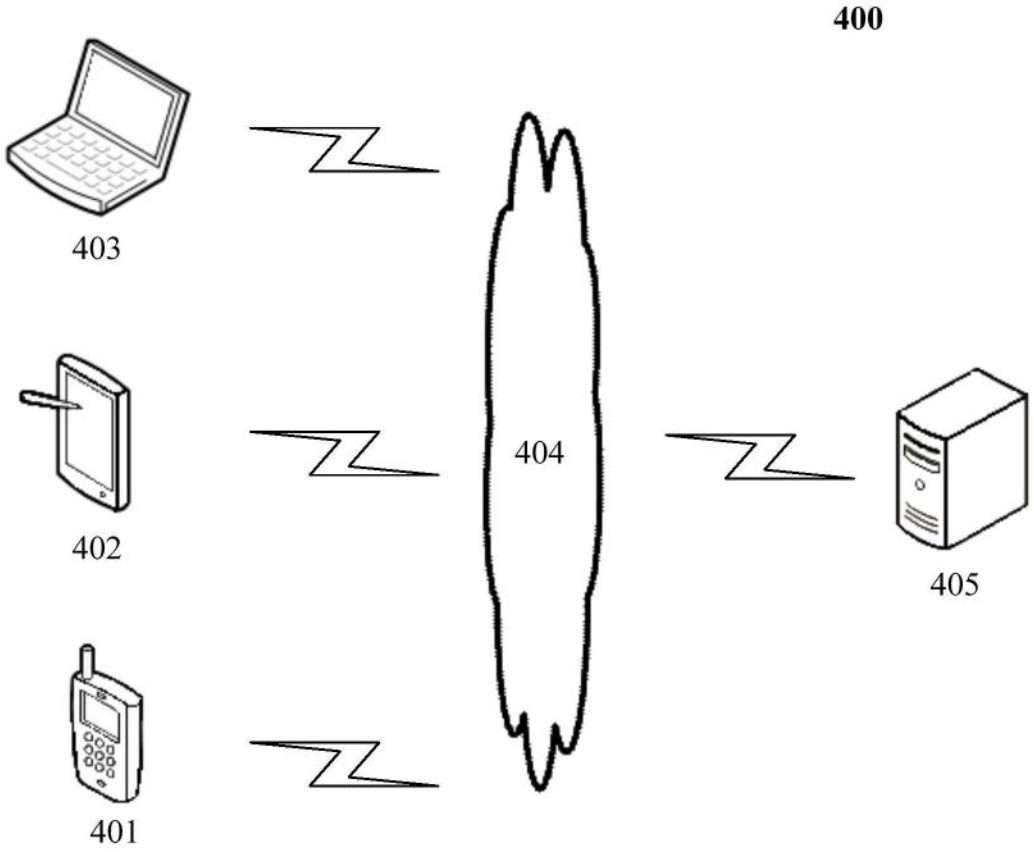

400
403
402
401
404
405

图4

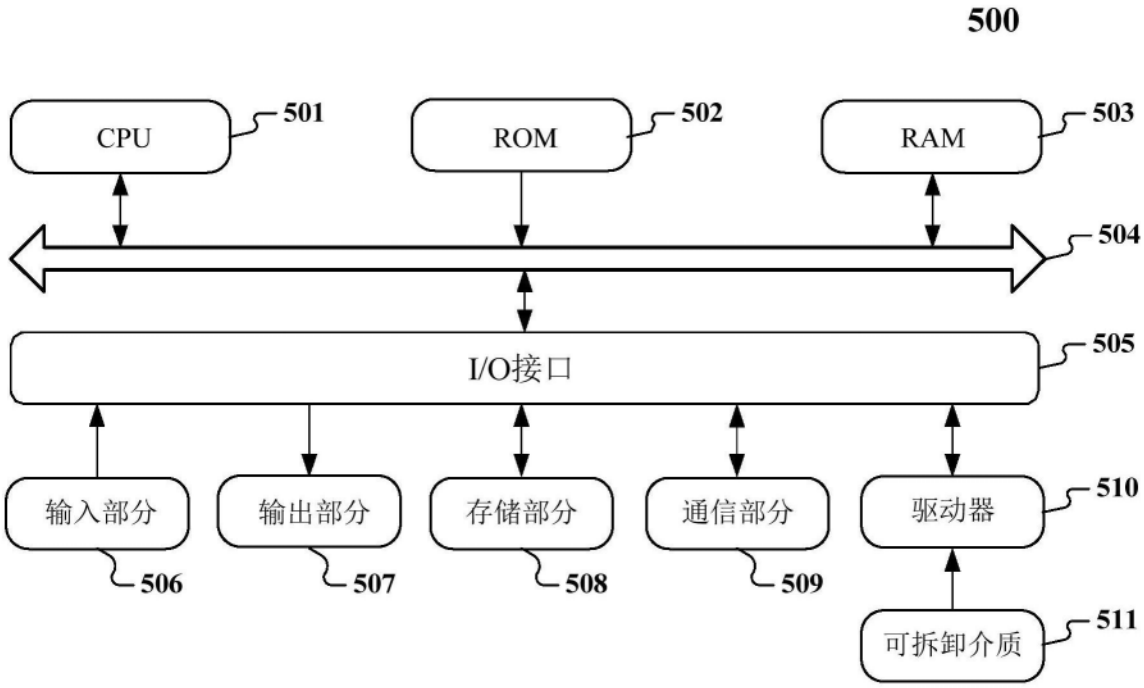

500
CPU
501
ROM
502
RAM
503
504
I/O接口
505
输入部分
506
输出部分
507
存储部分
508
通信部分
509
驱动器
510
可拆卸介质
511

图5