



(12)发明专利申请

(10)申请公布号 CN 109829040 A

(43)申请公布日 2019.05.31

(21)申请号 201811572724.8

(22)申请日 2018.12.21

(71)申请人 深圳市元征科技股份有限公司
地址 518000 广东省深圳市龙岗区坂田街
道五和大道北4012元征工业园

(72)发明人 刘均 于海悦

(74)专利代理机构 广州三环专利商标代理有限公司 44202
代理人 郝传鑫 熊永强

(51)Int.Cl.
G06F 16/332(2019.01)
G06F 17/27(2006.01)

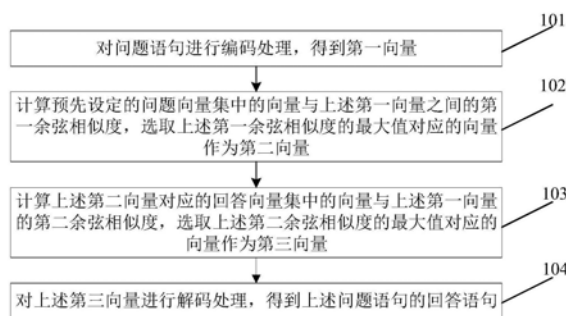
权利要求书2页 说明书10页 附图3页

(54)发明名称

一种智能对话方法及装置

(57)摘要

本申请公开了一种智能对话方法及装置。该方法包括：对问题语句进行编码处理，得到第一向量；计算预先设定的问题向量集中的向量与上述第一向量之间的第一余弦相似度，选取上述第一余弦相似度的最大值对应的向量作为第二向量；计算上述第二向量对应的回答向量集中的向量与上述第一向量的第二余弦相似度，选取上述第二余弦相似度的最大值对应的向量作为第三向量；对上述第三向量进行解码处理，得到上述问题语句的回答语句。还公开了相应的装置。通过对问题语句进行深度性语义编码，可以更好的从语料库中确定与问题语句语义相似的回答，自动生成符合问题语句的逻辑的回答。



1. 一种智能对话方法,其特征在于,包括:
 - 对问题语句进行编码处理,得到第一向量;
 - 计算预先设定的问题向量集中的向量与所述第一向量之间的第一余弦相似度,选取所述第一余弦相似度的最大值对应的向量作为第二向量;
 - 计算所述第二向量对应的回答向量集中的向量与所述第一向量的第二余弦相似度,选取所述第二余弦相似度的最大值对应的向量作为第三向量;
 - 对所述第三向量进行解码处理,得到所述问题语句的回答语句。
2. 根据权利要求1所述的方法,其特征在于,所述对问题语句进行编码处理,得到第一向量,包括:
 - 获取所述问题语句;
 - 将所述问题语句进行分词,得到多个第一词语;
 - 对所述多个第一词语进行第一特征提取处理,得到多个特征向量,其中所述特征向量与所述第一词语一一对应;
 - 按次序串联所述多个特征向量,得到所述第一向量,其中,所述次序为所述第一词语在所述问题语句中出现的顺序。
3. 根据权利要求2所述的方法,其特征在于,所述对所述多个第一词语进行第一特征提取处理,得到多个特征向量,包括:
 - 对所述多个第一词语进行第二特征提取处理,得到多个第一特征信息;
 - 将所述多个第一特征信息与所述多个第一词语的信息融合,得到多个第二特征信息,所述多个第一词语的信息包含所述第一词语的词性和所述第一词语在所述次序中的位置;
 - 对所述多个第二特征信息进行第三特征提取处理,得到所述多个特征向量。
4. 根据权利要求1所述的方法,其特征在于,所述计算预先设定的问题向量集中的向量与所述第一向量之间的第一余弦相似度,选取所述第一余弦相似度的最大值对应的向量作为第二向量,包括:
 - 确定所述第一向量在向量空间中的坐标,其中,所述向量空间由所述问题向量集得到;
 - 根据所述坐标、所述第一向量、所述问题向量集中的向量及所述问题向量集中的向量在所述向量空间中的坐标,得到向量夹角,其中,所述向量夹角为所述第一向量与所述问题向量集中的向量之间的夹角;
 - 选取所述向量夹角的最小值对应的向量作为所述第二向量。
5. 根据权利要求1所述的方法,其特征在于,所述对所述第三向量进行解码处理,得到所述问题语句的回答语句,包括:
 - 将所述第三向量按所述次序分为多个词向量;
 - 对所述多个词向量进行解码,得到多个第二词语;
 - 按所述次序串联所述多个第二词语,得到所述回答语句。
6. 一种智能对话装置,其特征在于,包括:
 - 编码单元,用于对问题语句进行编码处理,得到第一向量;
 - 第一计算单元,用于计算预先设定的问题向量集中的向量与所述第一向量之间的第一余弦相似度,选取所述第一余弦相似度的最大值对应的向量作为第二向量;
 - 第二计算单元,用于计算所述第二向量对应的回答向量集中的向量与所述第一向量的

- 第二余弦相似度,选取所述第二余弦相似度的最大值对应的向量作为第三向量;
解码单元,用于对所述第三向量进行解码处理,得到所述问题语句的回答语句。
7. 根据权利要求6所述的装置,其特征在于,所述编码单元,包括:
获取子单元,用于获取所述问题语句;
分词子单元,用于将所述问题语句进行分词,得到多个第一词语;
第一处理子单元,用于对所述多个第一词语进行第一特征提取处理,得到多个特征向量,其中所述特征向量与所述第一词语一一对应;
第一串联子单元,用于按次序串联所述多个特征向量,得到所述第一向量,其中,所述次序为所述第一词语在所述问题语句中出现的顺序。
8. 根据权利要求6所述的装置,其特征在于,所述第一计算单元,包括:
确定子单元,用于确定所述第一向量在向量空间中的坐标,其中,所述向量空间由所述问题向量集得到;
计算子单元,用于根据所述坐标、所述第一向量、所述问题向量集中的向量及所述问题向量集中的向量在所述向量空间中的坐标,得到向量夹角,其中,所述向量夹角为所述第一向量与所述问题向量集中的向量之间的夹角;
选取子单元,用于选取所述向量夹角的最小值对应的向量作为所述第二向量。
9. 一种智能对话装置,其特征在于,包括:处理器、存储器;所述计算机可读存储介质中存储有指令,当其在计算机上运行时,使得计算机执行如权利要求1~5任一项所述的方法。
10. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质中存储有指令,当其在计算机上运行时,使得计算机执行如权利要求1~5任一项所述的方法。

一种智能对话方法及装置

技术领域

[0001] 本申请涉及人工智能技术领域,尤其涉及一种智能对话方法及装置。

背景技术

[0002] 人机对话是自然语言处理领域衍生的一项现实应用场景,根据现实获取的大量短文本知识数据,构建单轮短对话式智能应答聊天机器人可在很多场景下节省人力成本。传统的信息检索数据挖掘的方法的主要思想是:如果某个词或短语在一篇文章中出现的频率高,并且在其他文章中很少出现,则认为此词或者短语具有很好的类别区分能力,适合用来分类,但其存在逻辑不一致的问题。

发明内容

[0003] 本申请提供一种智能对话方法及装置,以实现自动回答问题语句。

[0004] 第一方面,提供了一种智能对话方法,包括:对问题语句进行编码处理,得到第一向量;计算预先设定的问题向量集中的向量与所述第一向量之间的第一余弦相似度,选取所述第一余弦相似度的最大值对应的向量作为第二向量;计算所述第二向量对应的回答向量集中的向量与所述第一向量的第二余弦相似度,选取所述第二余弦相似度的最大值对应的向量作为第三向量;对所述第三向量进行解码处理,得到所述问题语句的回答语句。

[0005] 在一种可能实现的方式中,所述对问题语句进行编码处理,得到第一向量,包括:获取所述问题语句;将所述问题语句进行分词,得到多个第一词语;对所述多个第一词语进行第一特征提取处理,得到多个特征向量,其中所述特征向量与所述第一词语一一对应;按次序串联所述多个特征向量,得到所述第一向量,其中,所述次序为所述第一词语在所述问题语句中出现的顺序。

[0006] 在另一种可能实现的方式中,所述对所述多个第一词语进行第一特征提取处理,得到多个特征向量,包括:对所述多个第一词语进行第二特征提取处理,得到多个第一特征信息;将所述多个第一特征信息与所述多个第一词语的信息融合,得到多个第二特征信息,所述多个第一词语的信息包含所述第一词语的词性和所述第一词语在所述次序中的位置;对所述多个第二特征信息进行第三特征提取处理,得到所述多个特征向量。

[0007] 在又一种可能实现的方式中,所述计算预先设定的问题向量集中的向量与所述第一向量之间的第一余弦相似度,选取所述第一余弦相似度的最大值对应的向量作为第二向量,包括:确定所述第一向量在向量空间中的坐标,其中,所述向量空间由所述问题向量集得到;根据所述坐标、所述第一向量、所述问题向量集中的向量及所述问题向量集中的向量在所述向量空间中的坐标,得到向量夹角,其中,所述向量夹角为所述第一向量与所述问题向量集中的向量之间的夹角;选取所述向量夹角的最小值对应的向量作为所述第二向量。

[0008] 在又一种可能实现的方式中,所述对所述第三向量进行解码处理,得到所述问题语句的回答语句,包括:将所述第三向量按所述次序分为多个词向量;对所述多个词向量进行解码,得到多个第二词语;按所述次序串联所述多个第二词语,得到所述回答语句。

[0009] 第二方面,提供了一种智能对话装置,包括:编码单元,用于对问题语句进行编码处理,得到第一向量;第一计算单元,用于计算预先设定的问题向量集中的向量与所述第一向量之间的第一余弦相似度,选取所述第一余弦相似度的最大值对应的向量作为第二向量;第二计算单元,用于计算所述第二向量对应的回答向量集中的向量与所述第一向量的第二余弦相似度,选取所述第二余弦相似度的最大值对应的向量作为第三向量;解码单元,用于对所述第三向量进行解码处理,得到所述问题语句的回答语句。

[0010] 在一种可能实现的方式中,所述编码单元,包括:获取子单元,用于获取所述问题语句;分词子单元,用于将所述问题语句进行分词,得到多个第一词语;第一处理子单元,用于对所述多个第一词语进行第一特征提取处理,得到多个特征向量,其中所述特征向量与所述第一词语一一对应;第一串联子单元,用于按次序串联所述多个特征向量,得到所述第一向量,其中,所述次序为所述第一词语在所述问题语句中出现的顺序。

[0011] 在另一种可能实现的方式中,所述第一处理子单元还用于:对所述多个第一词语进行第二特征提取处理,得到多个第一特征信息;以及将所述多个第一特征信息与所述多个第一词语的信息融合,得到多个第二特征信息,所述多个第一词语的信息包含所述第一词语的词性和所述第一词语在所述次序中的位置;以及对所述多个第二特征信息进行第三特征提取处理,得到所述多个特征向量。

[0012] 在又一种可能实现的方式中,所述第一计算单元,包括:确定子单元,用于确定所述第一向量在向量空间中的坐标,其中,所述向量空间由所述问题向量集得到;计算子单元,用于根据所述坐标、所述第一向量、所述问题向量集中的向量及所述问题向量集中的向量在所述向量空间中的坐标,得到向量夹角,其中,所述向量夹角为所述第一向量与所述问题向量集中的向量之间的夹角;选取子单元,用于选取所述向量夹角的最小值对应的向量作为所述第二向量。

[0013] 在又一种可能实现的方式中,所述解码单元,包括:第二处理子单元,用于将所述第三向量按所述次序分为多个词向量;解码子单元,用于对所述多个词向量进行解码,得到多个第二词语;第二串联子单元,用于按所述次序串联所述多个第二词语,得到所述回答语句。

[0014] 第三方面,提供了一种智能对话装置,包括:包括处理器、存储器;所述处理器被配置为支持所述装置执行上述第一方面及其任一种可能的实现方式的方法中相应的功能。存储器用于与处理器耦合,其保存所述装置必要的程序(指令)和数据。可选的,所述装置还可以包括输入/输出接口,用于支持所述装置与其他装置之间的通信。

[0015] 第四方面,提供了一种计算机可读存储介质,所述计算机可读存储介质中存储有指令,当其在计算机上运行时,使得计算机执行上述第一方面及其任一种可能的实现方式的方法。

[0016] 第五方面,提供了一种包含指令的计算机程序产品,当其在计算机上运行时,使得计算机执行上述第一方面及其任一种可能的实现方式的方法。

[0017] 本申请通过对用户输入的问题语句进行深度性语义编码,得到相应的向量,再从语义空间中确定语料库中与用户输入的问题语句语义相似度最高的问题语句,进而从问题语句中的回答中选取与用户输入的问题语句语义相似度最高的回答语句,自动生成符合问题语句的逻辑的回答。

附图说明

[0018] 为了更清楚地说明本申请实施例或背景技术中的技术方案,下面将对本申请实施例或背景技术中所需要使用的附图进行说明。

[0019] 图1为本申请实施例提供了一种智能对话方法的流程示意图;

[0020] 图2为本申请实施例提供的另一种智能对话方法的流程示意图;

[0021] 图3为本申请实施例提供了一种智能对话装置的结构示意图;

[0022] 图4为本申请实施例提供了一种智能对话装置的硬件结构示意图。

具体实施方式

[0023] 为了使本技术领域的人员更好地理解本申请方案,下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚、完整地描述。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0024] 本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别不同对象,而不是用于描述特定顺序。此外,术语“包括”和“具有”以及它们任何变形,意图在于覆盖不排他的包含。例如包含了一系列步骤或单元的过程、方法、系统、产品或设备没有限定于已列出的步骤或单元,而是可选地还包括没有列出的步骤或单元,或可选地还包括对于这些过程、方法、产品或设备固有的其他步骤或单元。

[0025] 在本文中提及“实施例”意味着,结合实施例描述的特定特征、结构或特性可以包含在本申请的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例,也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是,本文所描述的实施例可以与其它实施例相结合。

[0026] 为了更清楚地说明本申请实施例或背景技术中的技术方案,下面将对本申请实施例或背景技术中所需要使用的附图进行说明。

[0027] 下面结合本申请实施例中的附图对本申请实施例进行描述。

[0028] 请参阅图1,图1是本申请实施例提供了一种智能对话方法的流程示意图。

[0029] 101、对问题语句进行编码处理,得到第一向量。

[0030] 输入的问题语句的长度与输出的回答语句的长度不一定相同,对于这种输入输出次序的长度是动态变化的情况,传统的神经网络是无法处理的,首先对用户输入的问题语句的次序等特征进行编码,得到第一向量,即可得到问题语句的语义,这样,在后续的智能对话中能更好的给出符合问题语句语义的回答。

[0031] 102、计算预先设定的问题向量集中的向量与上述第一向量之间的第一余弦相似度,选取上述第一余弦相似度的最大值对应的向量作为第二向量。

[0032] 用户可根据自身的需求,自定义语料库,如:用户需要用来进行汽修,自动回答汽车故障问题,则可以用大量描述汽车故障的语句及其相应的回答语句来构建语料库。在构建完语料库后,通过上述基于RNN的编码模型对语料库中的所有语句进行深度编码,得到问题向量集,即包含所有问题语句编码后得到的向量。

[0033] 通过计算第一向量与问题向量集中任意一个向量之间的余弦相似度,即可得到第一向量与问题向量集中任意一个向量之间的相似度,即第一余弦相似度。余弦相似度越大,

说明两个向量之间的相似度越大,两个向量对应的语句的语义相似度也就越大,因此,选取第一余弦相似度的最大值对应的向量作为第二向量。

[0034] 103、计算上述第二向量对应的回答向量集中的向量与上述第一向量的第二余弦相似度,选取上述第二余弦相似度的最大值对应的向量作为第三向量。

[0035] 语料库中每个问题语句的向量都对应有包含多个回答向量的回答向量集,与102得到第二向量一样,通过计算第一向量与第二向量对应的回答向量集中任意一个向量之间的余弦相似度,即可得到第一向量与第二向量对应的回答向量集中任意一个向量之间的相似度,即第二余弦相似度,并选取第二余弦相似度的最大值对应的向量作为第三向量。

[0036] 104、对上述第三向量进行解码处理,得到上述问题语句的回答语句。

[0037] 确定第三向量后,只需将第三向量进行解码,将其转化成语句输出即可得到问题语句的回答语句,完成智能对话。

[0038] 本申请实施例通过对用户输入的问题语句进行深度性语义编码,得到相应的向量,再从语义空间中确定语料库中与用户输入的问题语句语义相似度最高的问题语句,进而从问题语句中的回答中选取与用户输入的问题语句语义相似度最高的回答语句,自动生成符合问题语句的逻辑的回答。

[0039] 请参阅图2,图2是本申请实施例提供的另一种智能对话方法的流程示意图。

[0040] 201、获取问题语句。

[0041] 本申请可针对用户输入的任意问题语句实现智能对话,用户输入的方式可以为以下任意一种:通过智能终端(如:手机、平板电脑等)的屏幕手动输入、利用智能终端的语音采集功能语音输入、通过移动互联网进行远程输入。需要理解的是,无论用户是以何种方式完成输入的,智能终端都需将其转化为文字。如:当用户向智能终端说了一句:“刹车时有异响”,终端的语音处理系统会自动将用户的这句话转换成文字,即“刹车时有异响”这六个字。也就是说,本申请适用于对文字进行处理,获取的语句可以是任意问题语句,这与用户的需求和语料库有关,如本申请可应用于汽车维修领域,对用户输入的汽车故障语句自动给出回答语句,从而省去人工维修的时间和成本。

[0042] 202、将上述问题语句进行分词,得到多个第一词语。

[0043] 获取问题语句后,首先对该问题语句进行分词,具体地,通过分词工具对问题语句进行分词,得到多个第一词语。该分词工具是基于理解的分词方法完成分词的,即通过让计算机模拟人对句子的理解,达到识别词的效果。其基本思想就是在分词的同时进行句法、语义分析,利用句法信息和语义信息来处理歧义现象。需要理解的是,该分词工具可以是市面上任意一款成熟的分词产品,此处将不做限定。

[0044] 203、对上述多个第一词语进行第一特征提取处理,得到多个特征向量。

[0045] 将“我吃苹果”这句话分为:我、吃、苹果这三个词,且我、吃、苹果这三个词语的词性分别标注为:我(名词)、吃(动词)、苹果(名词)。显然,一个句子中,前一个词语其实对于当前词语的词性预测是有很大影响的,比如预测苹果的时候,由于前面的吃是一个动词,那么很显然苹果作为名词的概率就会远大于动词的概率,因为动词后面接名词很常见,而动词后面接动词很少见。

[0046] 传统的神经网络在对语句进行处理时,往往只能单独处理一个个的输入,也就是说,前一个输入和后一个输入是完全没有关系的。显然,当我们在理解一句话意思时,孤立

的理解这句话的每个词是不够的,我们需要处理这些词连接起来的整个序列,即前面的输入和后面的输入是有关系的。循环神经网络(recurrent neural network,RNN)通过内部的记忆来处理任意时序的输入序列,可很好的传统神经网络再处理语句时出现的上述问题。

[0047] 通过基于RNN构建的句子编码模型对上述任意一个第一词语进行第一特征提取处理,并得到一个对应的特征向量,完成对所有第一词语进行编码处理后,将得到多个特征向量。在一个具体实现的例子中,对上述多个第一词语进行第二特征提取处理,得到多个包含特征的序列在内的第一特征信息,将上述多个第一特征信息与上述多个第一词语的信息融合,得到多个第二特征信息,上述多个第一词语的信息包含上述第一词语的词性和上述第一词语在次序中的位置,如:将“刹车时有异响”这句话分为:刹车、时、有、异响这四个词语,那么这四个词语在次序中的位置(即正常人对该问题语句时的阅读顺序)分别为:刹车为1(代表问题语句中的第一个词语)、时为2、有为3、异响为4。对上述多个第二特征信息进行第三特征提取处理,得到多个特征向量。

[0048] 对于训练好的基于RNN的编码模型,只需分好词的问题语句输入模型即可,编码模型的具体结构如下:

[0049] 对于一句话 $X = (x_1, x_2, \dots, x_t)$,第 t 步中 x_t 进入隐层单元,具体地,首先进入重置门 r_t ,其中 σ 是sigmoid函数, W_r 是重置门的权重矩阵, $[h_{t-1}, x_t]$ 表示把上一个隐层的输出向量和本层的输入向量联合在一起构成新的输入向量,重置门的计算公式如下:

$$[0050] \quad r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

[0051] 接着进入更新门 z_t , W_z 为更新门的权重矩阵,更新门的计算公式如下:

$$[0052] \quad z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

[0053] 最后通过隐层输出对问题语句深度性编码后得到的向量,隐层 h_t 的输出计算公式如下:

$$[0054] \quad h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

[0055] \tilde{h}_t 的计算公式如下:

$$[0056] \quad \tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$

[0057] 其中, \tanh 是双曲正切函数, W 为权重矩阵。在第 T 步得到的 h_t 即为句子的语义向量。

[0058] 通过基于RNN的编码模型对上述多个第一词语进行第一特征提取处理,可得到多个特征向量,再按次序串联上述多个特征向量,即可得到第一向量。第一向量为对问题语句进行深度性编码后得到的,深度匹配问题语句的语义。

[0059] 204、确定上述第一向量在向量空间中的坐标。

[0060] 如上所述,用户可根据自身的需求,自定义语料库,如:用户需要用来进行汽修,自动回答汽车故障问题,则可以用大量描述汽车故障的语句及其相应的回答语句来构建语料库。在构建完语料库后,通过上述基于RNN的编码模型对语料库中的所有语句进行深度编码,得到问题向量集,即包含所有问题语句编码后得到的向量,再通过问题向量集中的所有向量可构建出一个向量空间,即可得到问题向量集中任意一个向量在向量空间中的坐标以及第一向量在向量空间中的坐标。

[0061] 205、根据上述坐标、上述第一向量、上述问题向量集中的向量及上述问题向量集

中的向量在上述向量空间中的坐标,得到向量夹角。

[0062] 根据第一向量在向量空间中的坐标、第一向量、问题向量集中的向量及问题向量集中的向量在上述向量空间中的坐标,计算第一向量与问题向量集中任意一个向量之间的夹角,具体计算公式如下:

$$[0063] \quad \cos \theta = \frac{a \cdot b}{\|a\| \|b\|}$$

[0064] 其中,a为第一向量,b为问题向量集中任意一个向量。

[0065] 这样,即可求得第一向量与问题向量集中任意一个向量之间的夹角,即第一向量与问题向量集中任意一个向量之间的余弦相似度,也就是说,通过计算两个向量的夹角即可得到第一向量与问题向量集中任意一个向量之间的相似度。

[0066] 206、选取上述向量夹角的最小值对应的向量作为第二向量。

[0067] 如上所述,通过计算两个向量的夹角即可得到第一向量与问题向量集中任意一个向量之间的相似度,上述向量夹角的最小值对应的向量即为与第一向量语义相似度最高的向量,因此,选取上述向量夹角的最小值对应的向量作为第二向量。

[0068] 基于RNN的编码器对可以深度匹配用户输入的问题语句和语料库中的问题语句的语义相似度,提升问题语句的检索效果。

[0069] 207、计算上述第二向量对应的回答向量集中的向量与上述第一向量的第二余弦相似度,选取上述第二余弦相似度的最大值对应的向量作为第三向量。

[0070] 语料库中包含问题语句以及这些问题语句相应的回答语句,且一句问题语句可能有多句回答语句,通过编码器对回答语句进行编码后,每个问题语句向量就对应有一个包含多个回答向量的回答向量集。显然同一句话在不同语境下的语义可能并不一样,这样,相同的问题语句在不同的语境下可能有不同的回答,因此需要对回答向量集中的向量与问题向量之间的语义相似度进行判断,从回答向量集中选取与问题语句语义最相似的向量,作为回答语句向量。

[0071] 同205~206确定第二向量的方式一样,通过计算第一向量与第二向量对应的回答向量集中任意一个向量之间的余弦相似度(即向量夹角),即可得到第一向量与第二向量对应的回答向量集中任意一个向量之间的相似度,上述向量夹角的最小值对应的向量即为与第一向量语义相似度最高的向量,因此,选取向量夹角的最小值对应的向量作为第三向量。

[0072] 208、将第三向量按上述次序分为多个词向量。

[0073] 确定第三向量后,只需将第三向量进行解码,将其转化成语句输出即可完成智能对话。

[0074] 上述第三向量是一句话对应的向量,与编码过程一样,需要先将其分解成多个词向量,分解次序和对问题语句进行分词的次序一样,这样,可保证问题语句和回答语句的逻辑一致性。

[0075] 209、对上述多个词向量进行解码,得到多个第二词语。

[0076] 对分解后得到的多个词向量依次进行解码,即可得到多个第二词语,需要指出的是,此处对每个词向量进行解码时都会结合当前被解码的词向量的上一个词向量的词性等信息进行解码,以得到最符合问题语句语义的词语。

[0077] 210、按上述次序串联上述多个第二词语,得到回答语句。

[0078] 按分词的次序将上述多个第二词语串联,即可得到问题语句对应的回答语句,最后,智能终端再将回答语句输出,即完成智能对话。输出的方式可以是一下任意一种:将回答语句转化成文字,通过智能终端的显示屏呈现给用户、将回答语句转化为语音对用户进行提示、通过移动互联网远程发送给用户。

[0079] 本申请的编码器和解码器均可通过训练提高编码和解码的精度,当用户觉得智能终端给出的回答不能解决自身的问题时,可向智能终端进行错误报备,同时也可将正确的回答通过智能终端输入到语料库中,这样,语料库可不断更新,智能对话的精度也将越来越高。

[0080] 本申请实施例通过对用户输入的问题语句进行深度性语义编码,得到相应的向量,再从语义空间中确定语料库中与用户输入的问题语句语义相似度最高的问题语句,进而从问题语句中的回答中选取与用户输入的问题语句语义相似度和逻辑关联最高的回答语句,自动生成符合问题语句的逻辑的回答。

[0081] 上述详细阐述了本申请实施例的方法,下面提供了本申请实施例的装置。

[0082] 请参阅图3,图3为本申请实施例提供的一种智能对话装置的结构示意图,该装置1包括:编码单元11、第一计算单元12、第二计算单元13及解码单元14。其中:

[0083] 编码单元11,用于对问题语句进行编码处理,得到第一向量;

[0084] 第一计算单元12,用于计算预先设定的问题向量集中的向量与所述第一向量之间的第一余弦相似度,选取所述第一余弦相似度的最大值对应的向量作为第二向量;

[0085] 第二计算单元13,用于计算所述第二向量对应的回答向量集中的向量与所述第一向量的第二余弦相似度,选取所述第二余弦相似度的最大值对应的向量作为第三向量;

[0086] 解码单元14,用于对所述第三向量进行解码处理,得到所述问题语句的回答语句。

[0087] 进一步地,所述编码单元11,包括:获取子单元111,用于获取所述问题语句;分词子单元112,用于将所述问题语句进行分词,得到多个第一词语;第一处理子单元113,用于对所述多个第一词语进行第一特征提取处理,得到多个特征向量,其中所述特征向量与所述第一词语一一对应;第一串联子单元114,用于按次序串联所述多个特征向量,得到所述第一向量,其中,所述次序为所述第一词语在所述问题语句中出现的顺序。

[0088] 进一步地,所述第一处理子单元113还用于:对所述多个第一词语进行第二特征提取处理,得到多个第一特征信息;以及将所述多个第一特征信息与所述多个第一词语的信息融合,得到多个第二特征信息,所述多个第一词语的信息包含所述第一词语的词性和所述第一词语在所述次序中的位置;以及对所述多个第二特征信息进行第三特征提取处理,得到所述多个特征向量。

[0089] 进一步地,所述第一计算单元12,包括:确定子单元121,用于确定所述第一向量在向量空间中的坐标,其中,所述向量空间由所述问题向量集得到;计算子单元122,用于根据所述坐标、所述第一向量、所述问题向量集中的向量及所述问题向量集中的向量在所述向量空间中的坐标,得到向量夹角,其中,所述向量夹角为所述第一向量与所述问题向量集中的向量之间的夹角;选取子单元123,用于选取所述向量夹角的最小值对应的向量作为所述第二向量。

[0090] 进一步地,所述解码单元14,包括:第二处理子单元141,用于将所述第三向量按所述次序分为多个词向量;解码子单元142,用于对所述多个词向量进行解码,得到多个第二

词语;第二串联子单元143,用于按所述次序串联所述多个第二词语,得到所述回答语句。

[0091] 本申请实施例通过对用户输入的问题语句进行深度性语义编码,得到相应的向量,再从语义空间中确定语料库中与用户输入的问题语句语义相似度最高的问题语句,进而从问题语句中的回答中选取与用户输入的问题语句语义相似度最高的回答语句,可实现自动、高效、准确的智能对话。

[0092] 图4为本申请实施例提供的一种智能对话的硬件结构示意图。该对话装置2包括处理器21,还可以包括输入装置22、输出装置23和存储器24。该输入装置22、输出装置23、存储器24和处理器21之间通过总线相互连接。

[0093] 存储器包括但不限于是随机存储记忆体(random access memory,RAM)、只读存储器(read-only memory,ROM)、可擦除可编程只读存储器(erasable programmable read only memory,EPR0M)、或便携式只读存储器(compact disc read-only memory,CD-ROM),该存储器用于相关指令及数据。

[0094] 输入装置用于输入数据和/或信号,以及输出装置用于输出数据和/或信号。输出装置和输入装置可以是独立的器件,也可以是一个整体的器件。

[0095] 处理器可以包括是一个或多个处理器,例如包括一个或多个中央处理器(central processing unit,CPU),在处理器是一个CPU的情况下,该CPU可以是单核CPU,也可以是多核CPU。

[0096] 存储器用于存储网络设备的程序代码和数据。

[0097] 处理器用于调用该存储器中的程序代码和数据,执行如下步骤:

[0098] 在一个实现方式中,所述处理器用于执行如下步骤:对问题语句进行编码处理,得到第一向量;计算预先设定的问题向量集中的向量与所述第一向量之间的第一余弦相似度,选取所述第一余弦相似度的最大值对应的向量作为第二向量;计算所述第二向量对应的回答向量集中的向量与所述第一向量的第二余弦相似度,选取所述第二余弦相似度的最大值对应的向量作为第三向量;对所述第三向量进行解码处理,得到所述问题语句的回答语句。

[0099] 在另一个实现方式中,所述处理器用于执行如下步骤:获取所述问题语句;将所述问题语句进行分词,得到多个第一词语;对所述多个第一词语进行第一特征提取处理,得到多个特征向量,其中所述特征向量与所述第一词语一一对应;按次序串联所述多个特征向量,得到所述第一向量,其中,所述次序为所述第一词语在所述问题语句中出现的顺序。

[0100] 在又一个实现方式中,所述处理器用于执行如下步骤:对所述多个第一词语进行第二特征提取处理,得到多个第一特征信息;将所述多个第一特征信息与所述多个第一词语的信息融合,得到多个第二特征信息,所述多个第一词语的信息包含所述第一词语的词性和所述第一词语在所述次序中的位置;对所述多个第二特征信息进行第三特征提取处理,得到所述多个特征向量。

[0101] 在又一个实现方式中,所述处理器用于执行如下步骤:确定所述第一向量在向量空间中的坐标,其中,所述向量空间由所述问题向量集得到;根据所述坐标、所述第一向量、所述问题向量集中的向量及所述问题向量集中的向量在所述向量空间中的坐标,得到向量夹角,其中,所述向量夹角为所述第一向量与所述问题向量集中的向量之间的夹角;选取所述向量夹角的最小值对应的向量作为所述第二向量。

[0102] 在又一个实现方式中,所述处理器用于执行如下步骤:将所述第三向量按所述次序分为多个词向量;对所述多个词向量进行解码,得到多个第二词语;按所述次序串联所述多个第二词语,得到所述回答语句。

[0103] 可以理解的是,图4仅仅示出了一种智能对话装置的简化设计。在实际应用中,智能对话装置还可以分别包含必要的其他元件,包含但不限于任意数量的输入/输出装置、处理器、控制器、存储器等,而所有可以实现本申请实施例的智能对话装置都在本申请的保护范围之内。

[0104] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本申请的范围。

[0105] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统、装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0106] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统、装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0107] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0108] 另外,在本申请各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。

[0109] 在上述实施例中,可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时,可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时,全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中,或者通过所述计算机可读存储介质进行传输。所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线(digital subscriber line,DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质,(例如,软盘、硬盘、磁带)、光介质(例如,数字通用光盘(digital versatile disc,DVD))、或者半导体介质(例如固态硬盘(solid state disk,SSD))等。

[0110] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程,该流程

可以由计算机程序来指令相关的硬件完成,该程序可存储于计算机可读取存储介质中,该程序在执行时,可包括如上述各方法实施例的流程。而前述的存储介质包括:只读存储器(read-only memory,ROM)或随机存储存储器(random access memory,RAM)、磁碟或者光盘等各种可存储程序代码的介质。

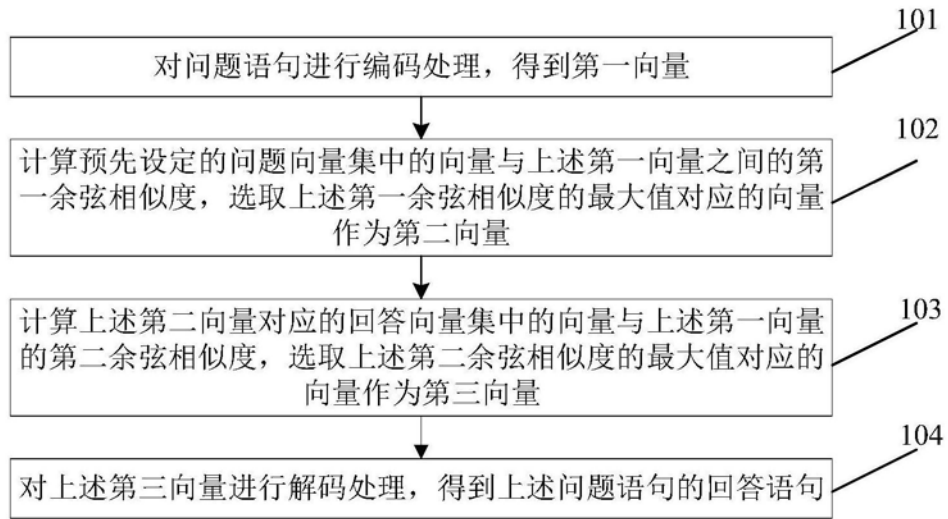


图1

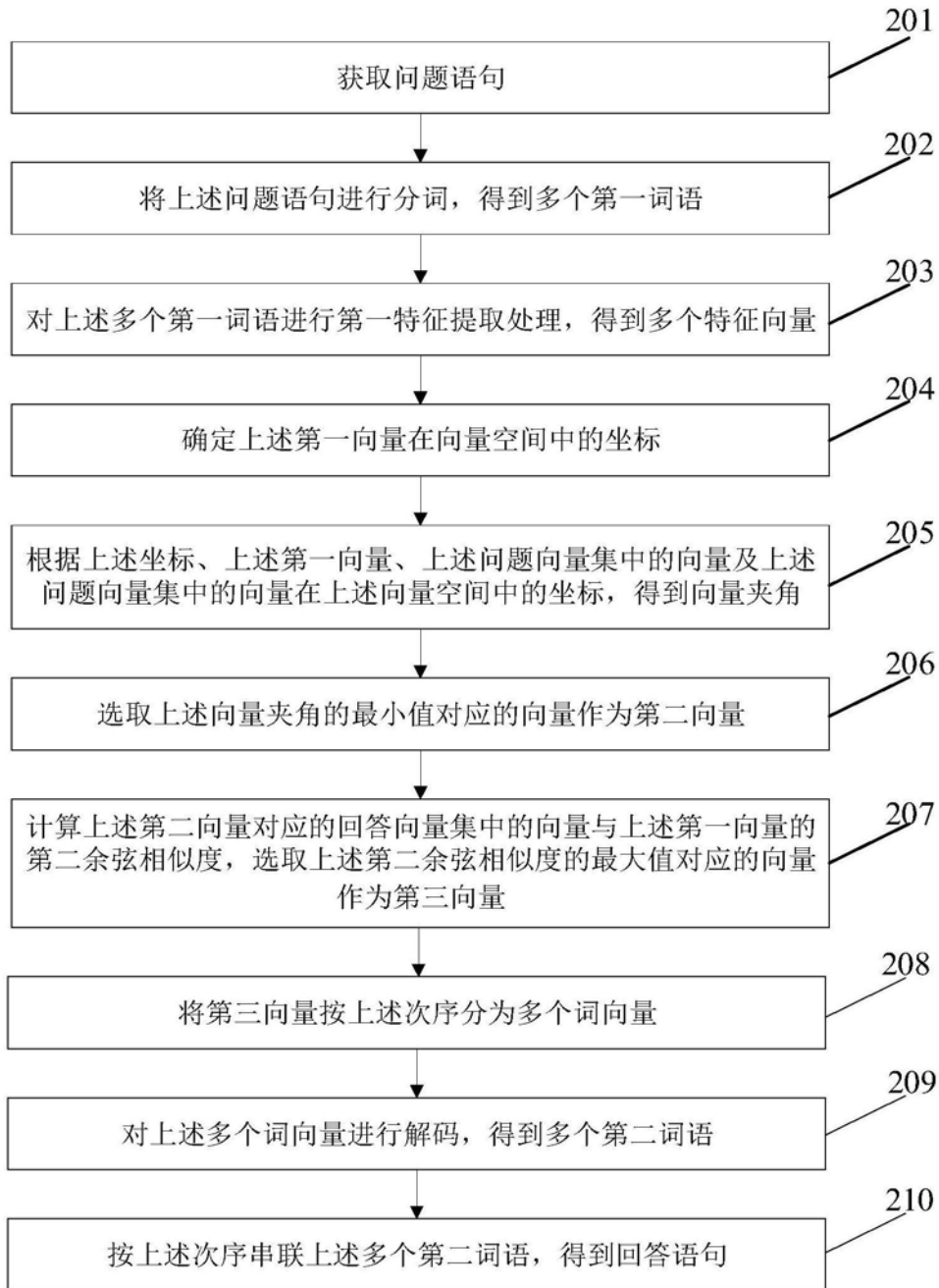


图2

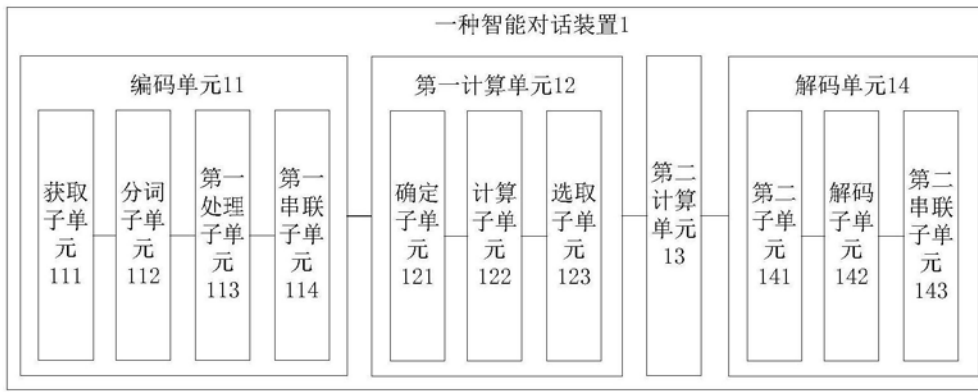


图3

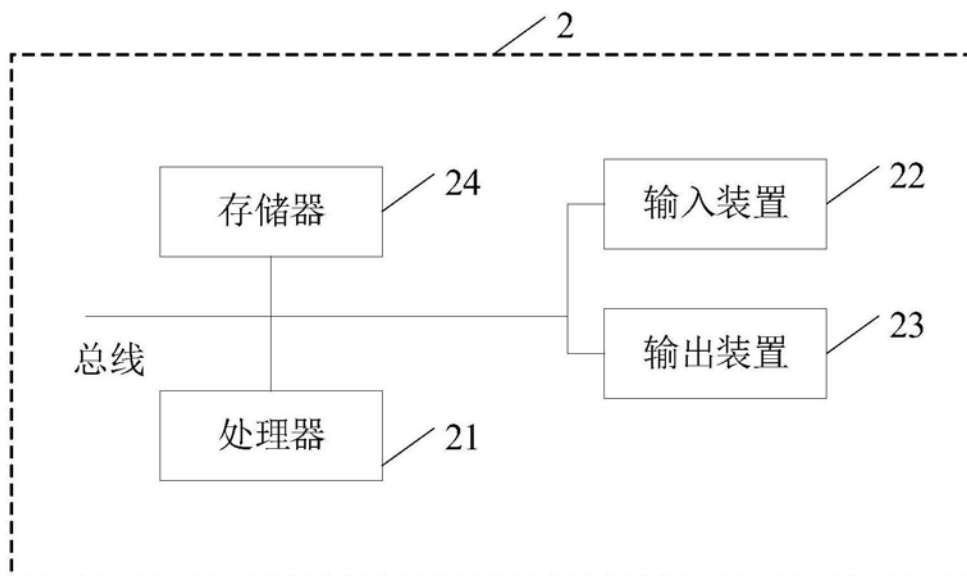


图4