



(12) 发明专利申请

(10) 申请公布号 CN 114626384 A

(43) 申请公布日 2022. 06. 14

(21) 申请号 202210194466.4

(22) 申请日 2022.03.01

(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 刘航 陈蒙

(74) 专利代理机构 北京清亦华知识产权代理事
务所(普通合伙) 11201
专利代理师 杜月

(51) Int.Cl.
G06F 40/35 (2020.01)
G06F 40/289 (2020.01)
G06F 16/332 (2019.01)
G06F 16/335 (2019.01)

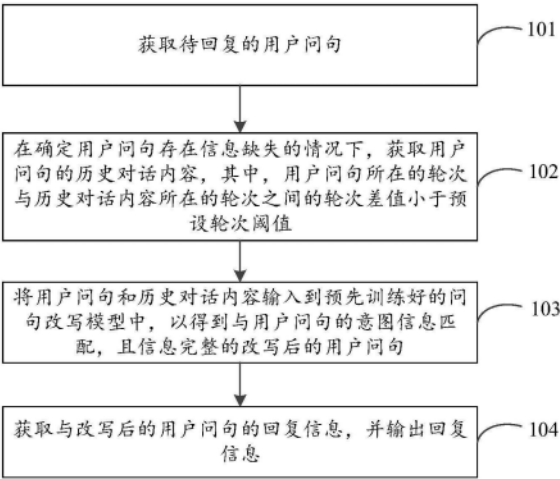
权利要求书4页 说明书15页 附图7页

(54) 发明名称

对话处理方法、问句改写模型的训练方法及其装置

(57) 摘要

本公开提出一种对话处理方法、问句改写模型的训练方法及其装置,其中,对话处理方法包括:获取待回复的用户问句;在确定用户问句存在信息缺失的情况下,获取用户问句的历史对话内容;将用户问句和历史对话内容输入到预先训练好的问句改写模型中,以得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句;获取与改写后的用户问句的回复信息,并输出回复信息。该方法在用户问句存在信息缺失的情况下,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句,由此,可准确地对用户问句中缺失的信息进行补充,实现用户问句内容的完整性,提高了用户体验。



1. 一种对话处理方法,其特征在于,所述方法包括:

获取待回复的用户问句;

在确定所述用户问句存在信息缺失的情况下,获取所述用户问句的历史对话内容,其中,所述用户问句所在的轮次与所述历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;

将所述用户问句和所述历史对话内容输入到预先训练好的问句改写模型中,以得到与所述用户问句的意图信息匹配,且信息完整的改写后的用户问句;

获取与所述改写后的用户问句的回复信息,并输出所述回复信息。

2. 如权利要求1所述的方法,其特征在于,所述问句改写模型包括编码层和解码层,所述将所述用户问句和所述历史对话内容输入到预先训练好的问句改写模型中,以得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句,包括:

将所述历史对话内容和所述用户问句输入到所述编码层中,以得到所述历史对话内容中各个词语的语义表示向量和所述用户问句的意图信息;

将所述历史对话内容中各个词语的语义表示向量和所述用户问句中各个词语的语义表示向量输入到所述解码层中,以通过所述解码层对所述用户问句进行改写,得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句。

3. 如权利要求2所述的方法,其特征在于,所述编码层包括语义表示层和意图信息识别层,所述将所述历史对话内容和所述用户问句输入到所述编码层中,以得到所述历史对话内容中各个词语的语义表示向量和所述用户问句的意图信息,包括:

将所述历史对话内容和所述用户问句进行拼接,以得到拼接文本;

将所述拼接文本输入到所述语义表示层中,以得到所述历史对话内容中各个词语的语义表示向量和所述用户问句中各个词语的语义表示向量;

将所述历史对话内容中各个词语的语义表示向量和所述用户问句中各个词语的语义表示向量输入到所述意图信息识别层,以得到所述用户问句的意图信息。

4. 如权利要求2所述的方法,其特征在于,所述将所述历史对话内容中各个词语的语义表示向量和所述用户问句中各个词语的语义表示向量输入到所述解码层中,以通过所述解码层对所述用户问句进行改写,得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句,包括:

根据所述多个词语中各个词语的语义表示向量,确定各个词语的词汇概率;

根据各个词语的词汇概率,选择词汇概率满足预设条件的词语作为目标词语;

根据所述目标词语对所述用户问句进行改写,得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句。

5. 如权利要求1-4任一项所述的方法,其特征在于,在所述将所述用户问句和所述历史对话内容输入到预先训练好的问句改写模型中,以得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句之前,所述方法还包括:

对所述历史对话内容进行分词,以得到所述历史对话内容中的各个词语;

将所述历史对话内容中的各个词语,与噪声词表中的噪声词进行匹配;

将所述历史对话内容中与所述噪声词匹配的词语删除。

6. 一种问句改写模型的训练方法,其特征在于,所述方法包括:

获取训练数据,其中,所述训练数据包括样本用户问句以及对应的样本历史对话内容,其中,所述样本用户问句存在信息缺失,所述样本用户问句所在的轮次与所述样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;

根据所述样本用户问句和所述样本历史对话内容对所述问句改写模型进行训练,直至所述问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与所述用户问句的意图信息匹配。

7.如权利要求6所述的方法,其特征在于,所述问句改写模型包括第一编码层和解码层,所述根据所述样本用户问句和所述样本历史对话内容对所述问句改写模型进行训练,直至所述问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与所述用户问句的意图信息匹配,包括:

将所述样本用户问句和所述样本历史对话内容输入到所述第一编码层中,以所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句的意图信息;

将所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句中各个词语的语义表示向量输入到所述解码层中,以通过所述第一编码层对所述样本用户问句进行改写,得到改写后的用户问句;

将所述改写后的用户问句输入到与第二编码层中,以通过所述第二编码层得到所述改写后的样本用户问句的意图信息,其中,所述第二编码层与所述第一编码层的参数是相同的;

根据所述样本用户问句的意图信息和所述改写后的样本用户问句的意图信息之间的差异信息,调整所述第一编码层和解码层的参数,直至所述问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与所述用户问句的意图信息匹配。

8.如权利要求7所述的方法,其特征在于,所述第一编码层包括语义表示层和意图信息识别层,所述将所述样本用户问句和所述样本历史对话内容输入到所述第一编码层中,以所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句的意图信息,包括:

将所述样本历史对话内容和所述样本用户问句进行拼接,以得到拼接文本;

将所述拼接文本输入到所述语义表示层中,以得到所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句中各个词语的语义表示向量;

将所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句中各个词语的语义表示向量输入到所述意图信息识别层,以得到所述样本用户问句的意图信息。

9.一种对话处理装置,其特征在于,所述装置包括:

第一获取模块,用于获取待回复的用户问句;

第二获取模块,用于在确定所述用户问句存在信息缺失的情况下,获取所述用户问句的历史对话内容,其中,所述用户问句所在的轮次与所述历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;

输入模块,用于将所述用户问句和所述历史对话内容输入到预先训练好的问句改写模型中,以得到与所述用户问句的意图信息匹配,且信息完整的改写后的用户问句;

第三获取模块,用于获取与所述改写后的用户问句的回复信息,并输出所述回复信息。

10.如权利要求9所述的装置,其特征在于,所述问句改写模型包括编码层和解码层,所

述输入模块,具体用于:

将所述历史对话内容和所述用户问句输入到所述编码层中,以得到所述历史对话内容中各个词语的语义表示向量和所述用户问句的意图信息;

将所述历史对话内容中各个词语的语义表示向量和所述用户问句中各个词语的语义表示向量输入到所述解码层中,以通过所述解码层对所述用户问句进行改写,得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句。

11.如权利要求10所述的装置,其特征在于,所述编码层包括语义表示层和意图信息识别层,所述输入模块,还用于:

将所述历史对话内容和所述用户问句进行拼接,以得到拼接文本;

将所述拼接文本输入到所述语义表示层中,以得到所述历史对话内容中各个词语的语义表示向量和所述用户问句中各个词语的语义表示向量;

将所述历史对话内容中各个词语的语义表示向量和所述用户问句中各个词语的语义表示向量输入到所述意图信息识别层,以得到所述用户问句的意图信息。

12.如权利要求10所述的装置,其特征在于,所述输入模块,还用于:

根据所述多个词语中各个词语的语义表示向量,确定各个词语的词汇概率;

根据各个词语的词汇概率,选择词汇概率满足预设条件的词语作为目标词语;

根据所述目标词语对所述用户问句进行改写,得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句。

13.如权利要求9-12任一项所述的装置,其特征在于,所述装置还包括:

分词模块,用于对所述历史对话内容进行分词,以得到所述历史对话内容中的各个词语;

匹配模块,用于将所述历史对话内容中的各个词语,与噪声词表中的噪声词进行匹配;

删除模块,用于将所述历史对话内容中与所述噪声词匹配的词语删除。

14.一种问句改写模型的训练装置,其特征在于,所述装置包括:

数据获取模块,用于获取训练数据,其中,所述训练数据包括样本用户问句以及对应的样本历史对话内容,其中,所述样本用户问句存在信息缺失,所述样本用户问句所在的轮次与所述样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;

训练模块,用于根据所述样本用户问句和所述样本历史对话内容对所述问句改写模型进行训练,直至所述问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与所述用户问句的意图信息匹配。

15.如权利要求14所述的装置,其特征在于,所述问句改写模型包括第一编码层和解码层,所述训练模块,具体用于:

将所述样本用户问句和所述样本历史对话内容输入到所述第一编码层中,以所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句的意图信息;

将所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句中各个词语的语义表示向量输入到所述解码层中,以通过所述第一编码层对所述样本用户问句进行改写,得到改写后的用户问句;

将所述改写后的用户问句输入到与第二编码层中,以通过所述第二编码层得到所述改写后的样本用户问句的意图信息,其中,所述第二编码层与所述第一编码层的参数是相同

的；

根据所述样本用户问句的意图信息和所述改写后的样本用户问句的意图信息之间的差异信息，调整所述第一编码层和解码层的参数，直至所述问句改写模型输出的改写后的用户问句的信息完整，并且意图信息与所述用户问句的意图信息匹配。

16. 如权利要求15所述的装置，其特征在于，所述第一编码层包括语义表示层和意图信息识别层，所述训练模块，还用于：

将所述样本历史对话内容和所述样本用户问句进行拼接，以得到拼接文本；

将所述拼接文本输入到所述语义表示层中，以得到所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句中各个词语的语义表示向量；

将所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句中各个词语的语义表示向量输入到所述意图信息识别层，以得到所述样本用户问句的意图信息。

17. 一种电子设备，其特征在于，包括：

存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述程序时实现如权利要求1-5中任一所述的方法，或者，实现如权利要求6-8中任一所述的方法。

18. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，该程序被处理器执行时实现如权利要求1-5中任一所述的方法，或者，实现如权利要求6-8中任一所述的方法。

对话处理方法、问句改写模型的训练方法及其装置

技术领域

[0001] 本公开涉及自然语言处理技术领域,尤其涉及一种对话处理方法、问句改写模型的训练方法及其装置。

背景技术

[0002] 目前,在多轮对话系统中,用户提出的问句中存在信息缺失的情况(如,省略和指代)。当问句中的关键成分缺失时,系统很容易错误理解用户的意图。例如。用户问句“什么意思”就是属于信息缺失的情况,系统仅根据这个问句的内容很难判断出用户在咨询哪方面的问题。

[0003] 相关技术中,对用户问句进行改写,根据对话内容抽取出省略和指代的内容补全至用户问句中,实现用户问句内容的完整性。但是,对于抽取出的省略和指代的内容是否真正是用户问句中缺失的内容,却无法判断,在抽取出的省略和指代的内容不是用户问句中缺失的内容时,将会导致用户问句出现较大偏差,影响了用户体验。

发明内容

[0004] 本公开提出一种对话处理方法,以实现在用户问句存在信息缺失的情况下,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句,由此,可准确地对用户问句中缺失的信息进行补充,实现用户问句内容的完整性,提高了用户体验。

[0005] 本公开第一方面实施例提出了一种对话处理方法,包括:获取待回复的用户问句;在确定所述用户问句存在信息缺失的情况下,获取所述用户问句的历史对话内容,其中,所述用户问句所在的轮次与所述历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;将所述用户问句和所述历史对话内容输入到预先训练好的问句改写模型中,以得到与所述用户问句的意图信息匹配,且信息完整的改写后的用户问句;获取与所述改写后的用户问句的回复信息,并输出所述回复信息。

[0006] 本公开实施例的对话处理方法,通过获取待回复的用户问句;在确定用户问句存在信息缺失的情况下,获取用户问句的历史对话内容,其中,用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;将用户问句和历史对话内容输入到预先训练好的问句改写模型中,以得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句;获取与改写后的用户问句的回复信息,并输出回复信息。该方法在用户问句存在信息缺失的情况下,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句,由此,可准确地对用户问句中缺失的信息进行补充,实现用户问句内容的完整性,提高了用户体验。

[0007] 可选地,所述问句改写模型包括编码层和解码层,所述将所述用户问句和所述历史对话内容输入到预先训练好的问句改写模型中,以得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句,包括:将所述历史对话内容和所述用户问句输入到所

述编码层中,以得到所述历史对话内容中各个词语的语义表示向量和所述用户问句的意图信息;将所述历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量输入到所述解码层中,以通过所述解码层对所述用户问句进行改写,得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句。

[0008] 可选地,所述编码层包括语义表示层和意图信息识别层,所述将所述历史对话内容和所述用户问句输入到所述编码层中,以得到所述历史对话内容中各个词语的语义表示向量和所述用户问句的意图信息,包括:将所述历史对话内容和所述用户问句进行拼接,以得到拼接文本;将所述拼接文本输入到所述语义表示层中,以得到所述历史对话内容中各个词语的语义表示向量和所述用户问句中各个词语的语义表示向量;将所述历史对话内容中各个词语的语义表示向量和所述用户问句中各个词语的语义表示向量输入到所述意图信息识别层,以得到所述用户问句的意图信息。

[0009] 可选地,所述将所述历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量输入到所述解码层中,以通过所述解码层对所述用户问句进行改写,得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句,包括:根据所述多个词语中各个词语的语义表示向量,确定各个词语的词汇概率;根据各个词语的词汇概率,选择词汇概率满足预设条件的词语作为目标词语;根据所述目标词语对所述用户问句进行改写,得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句。

[0010] 可选地,在所述将所述用户问句和所述历史对话内容输入到预先训练好的问句改写模型中,以得到与所述用户问句的意图信息一致,且信息完整的改写后的用户问句之前,所述方法还包括:对所述历史对话内容进行分词,以得到所述历史对话内容中的各个词语;将所述历史对话内容中的各个词语,与噪声词表中的噪声词进行匹配;将所述历史对话内容中与所述噪声词匹配的词语删除。

[0011] 本公开第二方面实施例提出了一种问句改写模型的训练方法,所述方法包括:获取训练数据,其中,所述训练数据包括样本用户问句以及对应的样本历史对话内容,其中,所述样本用户问句存在信息缺失,所述样本用户问句所在的轮次与所述样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;根据所述样本用户问句和所述样本历史对话内容对所述问句改写模型进行训练,直至所述问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与所述用户问句的意图信息匹配。

[0012] 本公开实施例的问句改写模型的训练方法,通过获取训练数据,其中,训练数据包括样本用户问句以及对应的样本历史对话内容,其中,样本用户问句存在信息缺失,样本用户问句所在的轮次与样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;根据样本用户问句和样本历史对话内容对问句改写模型进行训练,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。该方法根据样本用户问句和样本历史对话内容对问句改写模型进行训练,使训练好的问句改写模型可输出与用户问句的意图信息匹配,且信息完整的改写后的用户问句,以在用户问句存在信息缺失的情况下,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句,由此,可准确地对用户问句中缺失的信息进行补充,实现用户问句内容的完整性,提高了用户体验。

[0013] 可选地,所述问句改写模型包括第一编码层和解码层,所述根据所述样本用户问

句和所述样本历史对话内容对所述问句改写模型进行训练,直至所述问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与所述用户问句的意图信息匹配,包括:将所述样本用户问句和所述样本历史对话内容输入到所述第一编码层中,以所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句的意图信息;将所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句中各个词语的语义表示向量输入到所述解码层中,以通过所述第一编码层对所述样本用户问句进行改写,得到改写后的用户问句;将所述改写后的用户问句输入到与第二编码层中,以通过所述第二编码层得到所述改写后的样本用户问句的意图信息,其中,所述第二编码层与所述第一编码层的参数是相同的;根据所述样本用户问句的意图信息和所述改写后的样本用户问句的意图信息之间的差异信息,调整所述第一编码层和解码层的参数,直至所述问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与所述用户问句的意图信息匹配。

[0014] 可选地,所述第一编码层包括语义表示层和意图信息识别层,所述将所述样本用户问句和所述样本历史对话内容输入到所述第一编码层中,以所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句的意图信息,包括:将所述样本历史对话内容和所述样本用户问句进行拼接,以得到拼接文本;将所述拼接文本输入到所述语义表示层中,以得到所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句中各个词语的语义表示向量;将所述样本历史对话内容中各个词语的语义表示向量和所述样本用户问句中各个词语的语义表示向量输入到所述意图信息识别层,以得到所述样本用户问句的意图信息。

[0015] 本公开第三方面实施例提出了一种对话处理装置,所述装置包括:第一获取模块,用于获取待回复的用户问句;第二获取模块,用于在确定所述用户问句存在信息缺失的情况下,获取所述用户问句的历史对话内容,其中,所述用户问句所在的轮次与所述历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;输入模块,用于将所述用户问句和所述历史对话内容输入到预先训练好的问句改写模型中,以得到与所述用户问句的意图信息匹配,且信息完整的改写后的用户问句;第三获取模块,用于获取与所述改写后的用户问句的回复信息,并输出所述回复信息。

[0016] 本公开第四方面实施例提出了一种问句改写模型的训练装置,所述装置包括:数据获取模块,用于获取训练数据,其中,所述训练数据包括样本用户问句以及对应的样本历史对话内容,其中,所述样本用户问句存在信息缺失,所述样本用户问句所在的轮次与所述样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;训练模块,用于根据所述样本用户问句和所述样本历史对话内容对所述问句改写模型进行训练,直至所述问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与所述用户问句的意图信息匹配。

[0017] 本公开第五方面实施例提出了一种电子设备,包括:存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现上述第一方面实施例所述的方法,或者,实现上述第二方面实施例所述的方法。

[0018] 本公开第六方面实施例提出了一种计算机可读存储介质,其上存储有计算机程序,该程序被处理器执行时实现上述第一方面实施例所述的方法,或者,实现上述第二方面实施例所述的方法。

[0019] 本公开第七方面实施例提出了一种计算机程序产品,当所述计算机程序产品中的指令处理器执行时,执行上述第一方面实施例所述的方法,或者,执行上述第二方面实施例所述的方法。

[0020] 本公开附加的方面和优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本公开的实践了解到。

附图说明

[0021] 图1为本公开实施例一所提供的对话处理方法的流程示意图;

[0022] 图2为本公开实施例二所提供的对话处理方法的流程示意图;

[0023] 图3为本公开实施例三所提供的对话处理方法的流程示意图;

[0024] 图4为本公开实施例所提供的对话处理方法的流程示意图;

[0025] 图5为本公开实施例四所提供的问句改写模型的训练方法的流程示意图;

[0026] 图6为本公开实施例五所提供的问句改写模型的训练方法的流程示意图;

[0027] 图7为本公开实施例所提供的问句改写模型的训练方法的流程示意图;

[0028] 图8为本公开实施例六所提供的一种对话处理装置的结构示意图;

[0029] 图9为本公开实施例七所提供的一种问句改写模型的训练装置的结构示意图;

[0030] 图10是根据一示例性实施例示出的一种用于对话处理方法或问句改写模型的训练方法的电子设备的框图。

具体实施方式

[0031] 下面详细描述本发明的实施例,所述实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的,旨在用于解释本发明,而不能理解为对本发明的限制。

[0032] 下面参考附图描述本公开实施例的对话处理方法、问句改写模型的训练方法及其装置。

[0033] 图1为本公开实施例一所提供的对话处理方法的流程示意图,需要说明的是,该对话处理方法可应用于对话处理装置。该装置可被配置于电子设备中。其中,该电子设备可以是移动终端,例如,手机、平板电脑、个人数字助理等具有各种操作系统的硬件设备。

[0034] 如图1所示,对话处理方法包括如下步骤:

[0035] 步骤101,获取待回复的用户问句。

[0036] 在本公开实施例中,可从多轮对话中,获取用户提出的具有疑问形式的句子,将用户提出的具有疑问形式的句子作为待回复的用户问句,比如,用户提出“什么意思”,该问句可作为待回复的用户问句。

[0037] 步骤102,在确定用户问句存在信息缺失的情况下,获取用户问句的历史对话内容,其中,用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值。

[0038] 进一步地,可对待回复的用户问句进行自然语言处理,确定用户问句是否存在信息缺失。比如,对待回复的用户问句进行分词识别,获取待回复的用户问句对应的各个分词,进而,对各个分词进行词性识别,确定待回复问句是否存在信息缺失。又比如,对待回复

的用户问句进行语义识别,根据语义识别结果确定待回复问句是否存在信息缺失。

[0039] 进而,在确定用户问句存在信息缺失的情况下,可获取用户问句的历史对话内容,其中,用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值。

[0040] 需要理解的是,历史对话内容可保存用户和系统的多轮聊天内容,随着对话轮次的不断增加,用户问句的意图在不断的发生变化。当历史对话内容与用户问句距离越远时,它们具有相同意图的可能性越低;当历史对话内容与用户问句距离越近时,它们具有相同意图的概率越高。为了获取与用户问句的意图一致的历史对话内容,在本公开实施例中,对历史对话内容进行了对话轮次限制,用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值,如,用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于6。

[0041] 步骤103,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,以得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句。

[0042] 在本公开实施例中,可将用户问句和历史对话内容输入到预先训练好的问句改写模型中,问句改写模型可输出与用户问句的意图信息匹配,且信息完整的改写后的用户问句。

[0043] 步骤104,获取与改写后的用户问句的回复信息,并输出回复信息。

[0044] 进而,可根据改写后的用户问句的内容判断用户所咨询的问题,针对用户所咨询的问题,进行回复,并将回复的信息进行输出。

[0045] 综上,通过获取待回复的用户问句;在确定用户问句存在信息缺失的情况下,获取用户问句的历史对话内容,其中,用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;将用户问句和历史对话内容输入到预先训练好的问句改写模型中,以得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句;获取与改写后的用户问句的回复信息,并输出回复信息,由此,在用户问句存在信息缺失的情况下,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句,可准确地对用户问句中缺失的信息进行补充,实现用户问句内容的完整性,提高了用户体验。

[0046] 为了准确地得到与用户问句的意图信息一致,且信息完整的改写后的用户问句,如图2所示,图2为本公开实施例二所提供的对话处理方法的流程示意图。在本公开实施例中,可根据历史对话内容和用户问句,得到历史对话内容中各个词语的语义表示向量和用户问句的意图信息,进而,根据历史对话内容中各个词语的语义表示向量,获取与用户意图信息一致,且信息完整的改写后的用户问句。图2所示实施例可包括如下步骤:

[0047] 步骤201,获取待回复的用户问句。

[0048] 步骤202,在确定用户问句存在信息缺失的情况下,获取用户问句的历史对话内容,其中,用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值。

[0049] 步骤203,将历史对话内容和用户问句输入到编码层中,以得到历史对话内容中各个词语的语义表示向量和用户问句的意图信息。

[0050] 在本公开实施例中,问句改写模型可包括编码层和解码层,其中,编码层可包括语

义表示层和意图信息识别层。可选地,将历史对话内容和用户问句进行拼接,以得到拼接文本;将拼接文本输入到语义表示层中,以得到历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量;将历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量输入到意图信息识别层,以得到用户问句的意图信息。

[0051] 也就是说,可将历史对话内容和用户问句对应的两个文本进行拼接,得到拼接文本,进而,将拼接文本输入到编码层的语义表示层中,语义表示层可输出历史对话内容中各个词语的语义表示向量序列和用户问句中各个词语的语义表示向量序列。

[0052] 比如,以历史对话内容为context,用户问句为query为例,可将历史对话内容与用户问句拼接为一个长文本,将该长文本输入至编码层中的语义表示层中,得到历史对话内容中各个词语的语义表示向量序列和用户问句中各个词语的语义表示向量序列,进而得到历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量。具体可表现为如下公式: $H_c, H_q = \text{Encoder}(\text{context}, \text{query})$;

[0053] 其中, H_c 表示历史对话内容中各个词语的语义表示向量序列, H_q 表示用户问句中各个词语的语义表示向量序列。

[0054] 需要说明的是,为了获取用户问句的意图信息,可在历史对话内容和用户问句对应的两个文本前面增加一个特殊的分类标签“[CLS]”,并与历史对话内容和用户问句对应的两个文本一起输入至编码层的语义表示层中,该标签会与用户问句的各个词语的进行自注意力权重计算,同时也会与历史对话内容的各个词语的进行自注意力权重计算,得到标签的语义表示向量、历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量,进一步地,将该标签的语义表示向量、历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量输入到意图信息识别层,标签的语义表示向量与用户问句的各个词语的语义表示向量进行自注意力权重计算,同时也会与历史对话内容的各个词语的语义表示向量进行自注意力权重计算,最终输出用户问句的意图信息。步骤204,将历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量输入到解码层中,以通过解码层对用户问句进行改写,得到与用户问句的意图信息一致,且信息完整的改写后的用户问句。

[0055] 可选地,根据多个词语中各个词语的语义表示向量,确定各个词语的词汇概率;根据各个词语的词汇概率,选择词汇概率满足预设条件的词语作为目标词语;根据目标词语对用户问句进行改写,得到与用户问句的意图信息一致,且信息完整的改写后的用户问句。

[0056] 也就是说,将历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量输入至解码层中,解码层进行逐步解码时,可输出各个词语以及各个词语的词汇概率,进而,可将词汇概率足预设条件的词语作为目标词语,如,可将词汇概率最高的词语作为目标词语;进一步地,根据目标词语对用户进行改写,可得到与用户问句的意图信息一致,且信息完整的改写后的用户问句。

[0057] 另外,需要说明的是,解码层可使用copy机制(复制机制)进行逐步解码,为了减小解码层的词语搜索空间,减少资源消耗,可根据用户问句是否存在信息缺失,确定目标词语的来源。比如,在用户问句存在信息缺失时,解码层可从历史对话内容中复制用户问句缺失的信息;在用户问句不存在信息缺失时,解码层可从用户问句中复制词语,具体可表现为如下公式:

[0058] $P = \lambda P(H_c) + (1 - \lambda) P(H_q)$;

[0059] 其中, $P(\cdot)$ 表示词汇概率计算函数, λ 表示门控开关, 在用户问句不存在信息缺失时, $\lambda = 0$; 在用户问句存在信息缺失时, $\lambda = 1$ 。

[0060] 步骤205, 获取与改写后的用户问句的回复信息, 并输出回复信息。

[0061] 本公开实施例中, 步骤201-202、步骤205可以分别采用本公开的各实施例中的任一种方式实现, 本公开实施例并不对此作出限定, 也不再赘述。

[0062] 综上, 通过将历史对话内容和用户问句输入到编码层中, 以得到历史对话内容中各个词语的语义表示向量和用户问句的意图信息; 将历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量输入到解码层中, 以通过解码层对用户问句进行改写, 得到与用户问句的意图信息一致, 且信息完整的改写后的用户问句, 由此, 可准确地得到与用户问句的意图信息一致, 且信息完整的改写后的用户问句。

[0063] 为了问句改写模型对重要信息更好地进行识别, 使问句改写模型更加准确地输出与用户问句的意图信息一致, 且信息完整的改写后的用户问句, 如图3所示, 图3为本公开实施例三所提供的对话处理方法的流程示意图。在本公开实施例中, 由于历史对话内容中可存在噪声词, 比如, 邀评词、催拍催付词, 寒暄词等。这些噪声词没有为问句改写模型提供任何有效的信息, 反而干扰问句改写模型对重要信息的识别。因此, 在将所述用户问句和历史对话内容输入到预先训练好的问句改写模型中, 以得到与用户问句的意图信息一致, 且信息完整的改写后的用户问句之前, 可将历史对话内容中的噪声词删除。图3所示实施例可包括如下步骤:

[0064] 步骤301, 获取待回复的用户问句。

[0065] 步骤302, 在确定用户问句存在信息缺失的情况下, 获取用户问句的历史对话内容, 其中, 用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值。

[0066] 步骤303, 对历史对话内容进行分词, 以得到历史对话内容中的各个词语。

[0067] 在本公开实施例中, 可通过相关分词算法对历史对话内容进行分词, 得到历史对话内容中的各个词语。其中, 需要说明的是, 相关分词算法可包括: 基于词典的分词算法、基于机械学习的分词算法或基于神经网络的分词算法等。

[0068] 步骤304, 将历史对话内容中的各个词语, 与噪声词表中的噪声词进行匹配。

[0069] 进一步地, 可将历史对话内容的各个词语, 与预设的噪声词表中的噪声词进行匹配。

[0070] 步骤305, 将历史对话内容中与噪声词匹配的词语删除。

[0071] 进而, 可将历史对话内容中与噪声词匹配的词语删除, 比如, 可将“欢迎光临”词语删除。

[0072] 步骤306, 将用户问句和历史对话内容输入到预先训练好的问句改写模型中, 以得到与用户问句的意图信息匹配, 且信息完整的改写后的用户问句。

[0073] 步骤307, 获取与改写后的用户问句的回复信息, 并输出回复信息。

[0074] 本公开实施例中, 步骤301-302、步骤306-307可以分别采用本公开的各实施例中的任一种方式实现, 本公开实施例并不对此作出限定, 也不再赘述。

[0075] 综上, 通过将历史对话内容中的各个词语, 与噪声词表中的噪声词进行匹配; 对历

史对话内容进行分词,以得到历史对话内容中的各个词语;将历史对话内容中与噪声词匹配的词语删除,由此,可使问句改写模型对重要信息更好地进行识别,使问句改写模型更加准确地输出与用户问句的意图信息一致,且信息完整的改写后的用户问句。

[0076] 为了更加清楚地说明上述实施例,现举例进行说明。

[0077] 如图4所示,首先对历史对话内容进行噪声过滤,尽可能保留与用户问句意图相关的句子,然后编码器对历史对话内容和用户问句进行向量表示,获取词汇的语义信息;最后解码器根据词编码器的输出进行解码,通过多次复制输入中的词汇生成新的问句。为了保持改写后问句的意图不发生改变,可通过意图一致性约束模块,将改写后的问句进行重新编码,获得意图分布。通过保持意图分布的一致性,促使编码器更多的关注意图相关的信息,进而约束模型生成意图一致的问句。

[0078] 本公开实施例的对话处理方法,通过获取待回复的用户问句;在确定用户问句存在信息缺失的情况下,获取用户问句的历史对话内容,其中,用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;将用户问句和历史对话内容输入到预先训练好的问句改写模型中,以得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句;获取与改写后的用户问句的回复信息,并输出回复信息。该方法在用户问句存在信息缺失的情况下,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句,由此,可准确地对用户问句中缺失的信息进行补充,实现用户问句内容的完整性,提高了用户体验。

[0079] 为了实现上述实施例,本公开还提出一种问句改写模型的训练方法。

[0080] 图5为本公开实施例四所提供的问句改写模型的训练方法的流程示意图。需要说明的是,该问句改写模型的训练方法可应用于问句改写模型的训练装置。该装置可被配置于电子设备中。其中,该电子设备可以是移动终端,例如,手机、平板电脑、个人数字助理等具有各种操作系统的硬件设备。

[0081] 如图5所示,问句改写模型的训练方法可包括如下步骤:

[0082] 步骤501,获取训练数据,其中,训练数据包括样本用户问句以及对应的样本历史对话内容,其中,样本用户问句存在信息缺失,样本用户问句所在的轮次与样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值。

[0083] 在本公开实施例中,可将存在信息缺失的用户问句作为样本用户问句,并将获取样本用户问句的历史对话内容,为了获取与样本用户问句的意图一致的样本历史对话内容,在本公开实施例中,对样本历史对话内容进行了对话轮次限制。样本用户问句所在的轮次与样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值,如,样本用户问句所在的轮次与样本历史对话内容所在的轮次之间的轮次差值小于6。

[0084] 步骤502,根据样本用户问句和样本历史对话内容对问句改写模型进行训练,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。

[0085] 进一步地,根据样本用户问句和样本历史对话内容对问句改写模型进行训练,根据问句改写模型的输出结果对问句改写模型的参数进行调整,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。

[0086] 综上,通过获取训练数据,其中,训练数据包括样本用户问句以及对应的样本历史

对话内容,其中,样本用户问句存在信息缺失,样本用户问句所在的轮次与样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;根据样本用户问句和样本历史对话内容对问句改写模型进行训练,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。由此,根据样本用户问句和样本历史对话内容对问句改写模型进行训练,使训练好的问句改写模型可输出与用户问句的意图信息匹配,且信息完整的改写后的用户问句,以在用户问句存在信息缺失的情况下,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句,可准确地对用户问句中缺失的信息进行补充,实现用户问句内容的完整性,提高了用户体验。

[0087] 为了使训练好的问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配,如图6所示,图6为本公开实施例五所提供的问句改写模型的训练方法的流程示意图。在本公开实施例中,可通过第一编码层对历史对话内容和用户问句进行向量表示,得到历史对话内容的语义向量以及用户问句的意图信息,通过解码层对历史对话内容的语义向量进行解码,得到改写后的用户问句,进而,通过第二解码层获取改写后的样本用户问句的意图信息,根据样本用户问句的意图信息和改写后的样本用户问句的意图信息进行问句改写模型的参数调整,直至问句改写模型输出的改写后的样本用户问句的信息完整,并且意图信息与样本用户问句的意图信息匹配。图6所示实施例可包括如下步骤:

[0088] 步骤601,获取训练数据,其中,训练数据包括样本用户问句以及对应的样本历史对话内容,其中,样本用户问句存在信息缺失,样本用户问句所在的轮次与样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值。

[0089] 步骤602,将样本用户问句和样本历史对话内容输入到第一编码层中,以样本历史对话内容中各个词语的语义表示向量和样本用户问句的意图信息。

[0090] 在本公开实施例中,第一编码层包括语义表示层和意图信息识别层。

[0091] 可选地,将样本历史对话内容和样本用户问句进行拼接,以得到拼接文本;将拼接文本输入到语义表示层中,以得到样本历史对话内容中各个词语的语义表示向量和样本用户问句中各个词语的语义表示向量;将样本历史对话内容中各个词语的语义表示向量和样本用户问句中各个词语的语义表示向量输入到意图信息识别层,以得到样本用户问句的意图信息。

[0092] 可将样本历史对话内容和样本用户问句对应的两个文本进行拼接,得到拼接文本,进而,将拼接文本输入到第一编码层的语义表示层中,语义表示层可输出样本历史对话内容中各个词语的语义表示向量序列和样本用户问句中各个词语的语义表示向量序列。进一步地,将样本历史对话内容中各个词语的语义表示向量序列和样本用户问句中各个词语的语义表示向量序列输入到第一编码层的意图信息识别层,意图信息识别层可输出样本用户问句的意图信息。

[0093] 步骤603,将样本历史对话内容中各个词语的语义表示向量和样本用户问句中各个词语的语义表示向量输入到解码层中,以通过第一编码层对样本用户问句进行改写,得到改写后的用户问句。

[0094] 进一步地,将样本历史对话内容中各个词语的语义表示向量和样本用户问句中各

个词语的语义表示向量输入到解码层中,解码层进行逐步解码时,可输出各个词语以及各个词语的词汇概率,进而,可将词汇概率足预设条件的词语作为目标词语,如,可将词汇概率最高的词语作为目标词语,进一步地,第一编码层根据目标词语对样本用户问句进行改写,得到改写后的用户问句。

[0095] 步骤604,将改写后的用户问句输入到与第二编码层中,以通过第二编码层得到改写后的样本用户问句的意图信息,其中,第二编码层与第一编码层的参数是相同的。

[0096] 在本公开实施例中,可将改写后的用户问句输入到与第二编码层中,第二编码层可输出改写后的样本用户问句的意图信息,其中,需要说明的是,第二编码层与第一编码层具有相同的参数。

[0097] 步骤605,根据样本用户问句的意图信息和改写后的样本用户问句的意图信息之间的差异信息,调整第一编码层和解码层的参数,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。

[0098] 进而,将样本用户问句的意图信息和改写后的样本用户问句的意图信息进行对比,获取样本用户问句的意图信息和改写后的样本用户问句的意图信息之间的差异信息,根据该差异信息对第一编码层和解码层的参数进行调整,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。

[0099] 本公开实施例中,步骤601可以分别采用本公开的各实施例中的任一种方式实现,本公开实施例并不对此作出限定,也不再赘述。

[0100] 综上,通过将样本用户问句和样本历史对话内容输入到第一编码层中,以样本历史对话内容中各个词语的语义表示向量和样本用户问句的意图信息;将样本历史对话内容中各个词语的语义表示向量输入到解码层中,以通过第一编码层对样本用户问句进行改写,得到改写后的用户问句;将改写后的用户问句输入到与第二编码层中,以通过第二编码层得到改写后的样本用户问句的意图信息,其中,第二编码层与第一编码层的参数是相同的;根据样本用户问句的意图信息和改写后的样本用户问句的意图信息之间的差异信息,调整第一编码层和解码层的参数,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。由此,可使训练好的问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。

[0101] 为了更加清楚地说明上述实施例,现举例进行说明。

[0102] 如图7所示,问句改写模型为基于Transformer网络结构,问句改写模型中包括第一编码层和第二编码层,第二编码层与第一编码层的参数相同,为了获取用户问句的意图,在将历史对话容和用户问句的拼接文本输入第一编码层之前,在拼接文本前面增加一个特殊的分类标签“[CLS]”,并与拼接文本一起编码。在Transformer编码层的每一层,该标签会与用户问句与历史对话内容进行注意力权重计算,最终该标签的输出将会包含用户问句的缺失信息。该标签的输出向量(如,维度信息)可以被视为用户问句的意图分布或者完整的语义表示,即图7中的 Intent_q 。对于改写后的用户问句,在没有历史对话内容的情况下,内容比较完整,不存在信息缺失(如,省略和指代)的现象,其意图是明确的。进而,将改写后的问句与“[CLS]”标签重新输入至第二编码层中进行编码,获取改写后问句的意图分布,即图7中的 Intent_r 。最后,采用均方误差(MSE)让两个分布保持一致,促使编码层更关注于用户问句的缺失的内容,从而提升解码层复制缺失的词语的概率。具体可表现为如下公式:

[0103] $\text{loss}_{\text{intent}} = \text{MES}(\text{Int ent}_q, \text{Int ent}_r)$;

[0104] 该意图一致性约束方法仅在问句改写模型的训练阶段进行重编码计算,在真实应用场景中,仍然保持问句改写模型的编码-解码结构,没有带来额外的计算资源消耗。同时用户问句的意图一致性约束利用了神经网络的表征能力,结合KL散步进行分布约束,无需任何标注样本即可完成问句改写模型的训练。这种无监督的训练方式极大地节省了标注成本。

[0105] 本公开实施例的问句改写模型的训练方法,通过获取训练数据,其中,训练数据包括样本用户问句以及对应的样本历史对话内容,其中,样本用户问句存在信息缺失,样本用户问句所在的轮次与样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;根据样本用户问句和样本历史对话内容对问句改写模型进行训练,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。该方法根据样本用户问句和样本历史对话内容对问句改写模型进行训练,使训练好的问句改写模型可输出与用户问句的意图信息匹配,且信息完整的改写后的用户问句,以在用户问句存在信息缺失的情况下,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句,由此,可准确地对用户问句中缺失的信息进行补充,实现用户问句内容的完整性,提高了用户体验。

[0106] 为了实现图1至图4所述实施例,本公开还提出一种对话处理装置。图8为本公开实施例六所提供的一种对话处理装置的结构示意图。

[0107] 如图8所示,对话处理装置800包括:第一获取模块810、第二获取模块820、输入模块830和第三获取模块840。

[0108] 其中,第一获取模块810,用于获取待回复的用户问句;第二获取模块820,用于在确定用户问句存在信息缺失的情况下,获取用户问句的历史对话内容,其中,用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;输入模块830,用于将用户问句和历史对话内容输入到预先训练好的问句改写模型中,以得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句;第三获取模块840,用于获取与改写后的用户问句的回复信息,并输出回复信息。

[0109] 作为本公开实施例的一种可能实现方式,所述问句改写模型包括编码层和解码层,输入模块830,具体用于:将历史对话内容和用户问句输入到编码层中,以得到历史对话内容中各个词语的语义表示向量和用户问句的意图信息;将历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量输入到所述解码层中,以通过解码层对用户问句进行改写,得到与用户问句的意图信息一致,且信息完整的改写后的用户问句。

[0110] 作为本公开实施例的一种可能实现方式,编码层包括语义表示层和意图信息识别层,输入模块830,还用于:将历史对话内容和用户问句进行拼接,以得到拼接文本;将拼接文本输入到语义表示层中,以得到历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量;将历史对话内容中各个词语的语义表示向量和用户问句中各个词语的语义表示向量输入到意图信息识别层,以得到用户问句的意图信息。

[0111] 作为本公开实施例的一种可能实现方式,输入模块830,还用于:根据各个词语的词汇概率,选择词汇概率满足预设条件的词语作为目标词语;根据目标词语对用户问句进行改写,得到与用户问句的意图信息一致,且信息完整的改写后的用户问句。

[0112] 作为本公开实施例的一种可能实现方式,对话处理装置800还包括:分词模块、匹配模块和删除模块。

[0113] 其中,分词模块,用于对历史对话内容进行分词,以得到历史对话内容中的各个词语;匹配模块,用于将历史对话内容中的各个词语,与噪声词表中的噪声词进行匹配;删除模块,用于将历史对话内容中与噪声词匹配的词语删除。

[0114] 本公开实施例的对话处理装置,通过获取待回复的用户问句;在确定用户问句存在信息缺失的情况下,获取用户问句的历史对话内容,其中,用户问句所在的轮次与历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;将用户问句和历史对话内容输入到预先训练好的问句改写模型中,以得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句;获取与改写后的用户问句的回复信息,并输出回复信息。该装置可实现在用户问句存在信息缺失的情况下,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句,由此,可准确地对用户问句中缺失的信息进行补充,实现用户问句内容的完整性,提高了用户体验。

[0115] 为了实现图5至图7所述实施例,本公开还提出一种问句改写模型的训练装置。图9为本公开实施例七所提供的一种问句改写模型的训练装置的结构示意图。

[0116] 如图9所示,问句改写模型的训练装置900包括:数据获取模块910和训练模920。

[0117] 其中,数据获取模块910,用于获取训练数据,其中,训练数据包括样本用户问句以及对应的样本历史对话内容,其中,样本用户问句存在信息缺失,样本用户问句所在的轮次与样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;训练模块920,用于根据样本用户问句和样本历史对话内容对问句改写模型进行训练,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。

[0118] 作为本公开实施例的一种可能实现方式,问句改写模型包括第一编码层和解码层,训练模块920,具体用于:将样本用户问句和样本历史对话内容输入到第一编码层中,以样本历史对话内容中各个词语的语义表示向量和样本用户问句的意图信息;将样本历史对话内容中各个词语的语义表示向量和样本用户问句中各个词语的语义表示向量输入到解码层中,以通过第一编码层对样本用户问句进行改写,得到改写后的用户问句;将改写后的用户问句输入到与第二编码层中,以通过第二编码层得到改写后的样本用户问句的意图信息,其中,第二编码层与第一编码层的参数是相同的;根据样本用户问句的意图信息和改写后的样本用户问句的意图信息之间的差异信息,调整第一编码层和解码层的参数,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。

[0119] 作为本公开实施例的一种可能实现方式,第一编码层包括语义表示层和意图信息识别层,训练模块920,还用于:将样本历史对话内容和样本用户问句进行拼接,以得到拼接文本;将拼接文本输入到语义表示层中,以得到样本历史对话内容中各个词语的语义表示向量和样本用户问句中各个词语的语义表示向量;将样本历史对话内容中各个词语的语义表示向量和样本用户问句中各个词语的语义表示向量输入到意图信息识别层,以得到样本用户问句的意图信息。

[0120] 本公开实施例的问句改写模型的训练装置,通过获取训练数据,其中,训练数据包括样本用户问句以及对应的样本历史对话内容,其中,样本用户问句存在信息缺失,样本用

户问句所在的轮次与样本历史对话内容所在的轮次之间的轮次差值小于预设轮次阈值;根据样本用户问句和样本历史对话内容对问句改写模型进行训练,直至问句改写模型输出的改写后的用户问句的信息完整,并且意图信息与用户问句的意图信息匹配。该装置可实现根据样本用户问句和样本历史对话内容对问句改写模型进行训练,使训练好的问句改写模型可输出与与用户问句的意图信息匹配,且信息完整的改写后的用户问句,以在用户问句存在信息缺失的情况下,将用户问句和历史对话内容输入到预先训练好的问句改写模型中,得到与用户问句的意图信息匹配,且信息完整的改写后的用户问句,由此,可准确地对用户问句中缺失的信息进行补充,实现用户问句内容的完整性,提高了用户体验。

[0121] 为了实现上述实施例,本申请还提出一种电子设备,如图10所示,图10是根据一示例性实施例示出的一种用于对话处理方法或问句改写模型的训练方法的电子设备的框图。

[0122] 如图10所示,上述电子设备1000包括:

[0123] 存储器1010及处理器1020,连接不同组件(包括存储器1010和处理器1020)的总线1030,存储器1010存储有计算机程序,当处理器1020执行所述程序时实现本公开实施例所述的对话处理方法或问句改写模型的训练方法。

[0124] 总线1030表示几类总线结构中的一种或多种,包括存储器总线或者存储器控制器,外围总线,图形加速端口,处理器或者使用多种总线结构中的任意总线结构的局域总线。举例来说,这些体系结构包括但不限于工业标准体系结构(ISA)总线,微通道体系结构(MAC)总线,增强型ISA总线、视频电子标准协会(VESA)局域总线以及外围组件互连(PCI)总线。

[0125] 电子设备1000典型地包括多种电子设备可读介质。这些介质可以是任何能够被电子设备1000访问的可用介质,包括易失性和非易失性介质,可移动的和不可移动的介质。

[0126] 存储器1010还可以包括易失性存储器形式的计算机系统可读介质,例如随机存取存储器(RAM) 1040和/或高速缓存存储器1050。电子设备1000可以进一步包括其它可移动/不可移动的、易失性/非易失性计算机系统存储介质。仅作为举例,存储系统1060可以用于读写不可移动的、非易失性磁介质(图10未显示,通常称为“硬盘驱动器”)。尽管图10中未示出,可以提供用于对可移动非易失性磁盘(例如“软盘”)读写的磁盘驱动器,以及对可移动非易失性光盘(例如CD-ROM, DVD-ROM或者其它光介质)读写的光盘驱动器。在这些情况下,每个驱动器可以通过一个或者多个数据介质接口与总线1030相连。存储器1010可以包括至少一个程序产品,该程序产品具有一组(例如至少一个)程序模块,这些程序模块被配置以执行本公开各实施例的功能。

[0127] 具有一组(至少一个)程序模块1070的程序/实用工具1080,可以存储在例如存储器1010中,这样的程序模块1070包括但不限于操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。程序模块1070通常执行本公开所描述的实施例中的功能和/或方法。

[0128] 电子设备1000也可以与一个或多个外部设备1090(例如键盘、指向设备、显示器等)通信,还可与一个或者多个使得用户能与该电子设备1000交互的设备通信,和/或使得该电子设备1000能与一个或多个其它计算设备进行通信的任何设备(例如网卡,调制解调器等等)通信。这种通信可以通过输入/输出(I/O)接口1092进行。并且,电子设备1000还可以通过网络适配器1093与一个或者多个网络(例如局域网(LAN),广域网(WAN)和/或公共

网络,例如因特网)通信。如图10所示,网络适配器1093通过总线1030与电子设备1000的其它模块通信。应当明白,尽管图10中未示出,可以结合电子设备1000使用其它硬件和/或软件模块,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID系统、磁带驱动器以及数据备份存储系统等。

[0129] 处理器1020通过运行存储在存储器1010中的程序,从而执行各种功能应用以及数据处理。

[0130] 需要说明的是,本实施例的电子设备的实施过程和技术原理参见图1至图4对本公开实施例的对话处理的解释说明,或图5至图7对本公开实施例的问句改写模型的训练的解释说明,此处不再赘述。

[0131] 为了实现上述实施例,本申请还提出一种存储有计算机指令的非瞬时计算机可读存储介质,计算机指令用于使计算机执行上述图1至图4实施例所述的对话处理方法,或图5至图7实施例所述的问句改写模型的训练方法。

[0132] 为了实现上述实施例,本公开还提供一种计算机程序产品,当所述计算机程序产品中的指令处理器执行时,执行上述图1至图4实施例所述的对话处理方法,或图5至图7实施例所述的问句改写模型的训练方法。

[0133] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本公开的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不必针对的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示例中以合适的方式结合。此外,在不相互矛盾的情况下,本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

[0134] 此外,术语“第一”、“第二”仅用于描述目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本公开的描述中,“多个”的含义是至少两个,例如两个,三个等,除非另有明确具体的限定。

[0135] 流程图中或在此以其他方式描述的任何过程或方法描述可以被理解为,表示包括一个或更多个用于实现定制逻辑功能或过程的步骤的可执行指令的代码的模块、片段或部分,并且本公开的优选实施方式的范围包括另外的实现,其中可以不按所示出或讨论的顺序,包括根据所涉及的功能按基本同时的方式或按相反的顺序,来执行功能,这应被本公开的实施例所属技术领域的技术人员所理解。

[0136] 在流程图中表示或在此以其他方式描述的逻辑和/或步骤,例如,可以被认为用于实现逻辑功能的可执行指令的定序列表,可以具体实现在任何计算机可读介质中,以供指令执行系统、装置或设备(如基于计算机的系统、包括处理器的系统或其他可以从指令执行系统、装置或设备取指令并执行指令的系统)使用,或结合这些指令执行系统、装置或设备而使用。就本说明书而言,“计算机可读介质”可以是任何可以包含、存储、通信、传播或传输程序以供指令执行系统、装置或设备或结合这些指令执行系统、装置或设备而使用的装置。计算机可读介质的更具体的示例(非穷尽性列表)包括以下:具有一个或多个布线的电连接部(电子装置),便携式计算机盘盒(磁装置),随机存取存储器(RAM),只读存储器

(ROM),可擦除可编程只读存储器 (EPROM或闪速存储器),光纤装置,以及便携式光盘只读存储器 (CDROM)。另外,计算机可读介质甚至可以是可在其上打印所述程序的纸或其他合适的介质,因为可以例如通过对纸或其他介质进行光学扫描,接着进行编辑、解译或必要时以其他合适方式进行处理来以电子方式获得所述程序,然后将其存储在计算机存储器中。

[0137] 应当理解,本公开的各部分可以用硬件、软件、固件或它们的组合来实现。在上述实施方式中,多个步骤或方法可以用存储在存储器中且由合适的指令执行系统执行的软件或固件来实现。如,如果用硬件来实现和在另一实施方式中一样,可用本领域公知的下列技术中的任一项或他们的组合来实现:具有用于对数据信号实现逻辑功能的逻辑门电路的离散逻辑电路,具有合适的组合逻辑门电路的专用集成电路,可编程门阵列 (PGA),现场可编程门阵列 (FPGA) 等。

[0138] 本技术领域的普通技术人员可以理解实现上述实施例方法携带的全部或部分步骤是可以通程序来指令相关的硬件完成,所述的程序可以存储于一种计算机可读存储介质中,该程序在执行时,包括方法实施例的步骤之一或其组合。

[0139] 此外,在本公开各个实施例中的各功能单元可以集成在一个处理模块中,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模块中。上述集成的模块既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。所述集成的模块如果以软件功能模块的形式实现并作为独立的产品销售或使用,也可以存储在一个计算机可读取存储介质中。

[0140] 上述提到的存储介质可以是只读存储器,磁盘或光盘等。尽管上面已经示出和描述了本公开的实施例,可以理解的是,上述实施例是示例性的,不能理解为对本公开的限制,本领域的普通技术人员在本公开的范围内可以对上述实施例进行变化、修改、替换和变型。

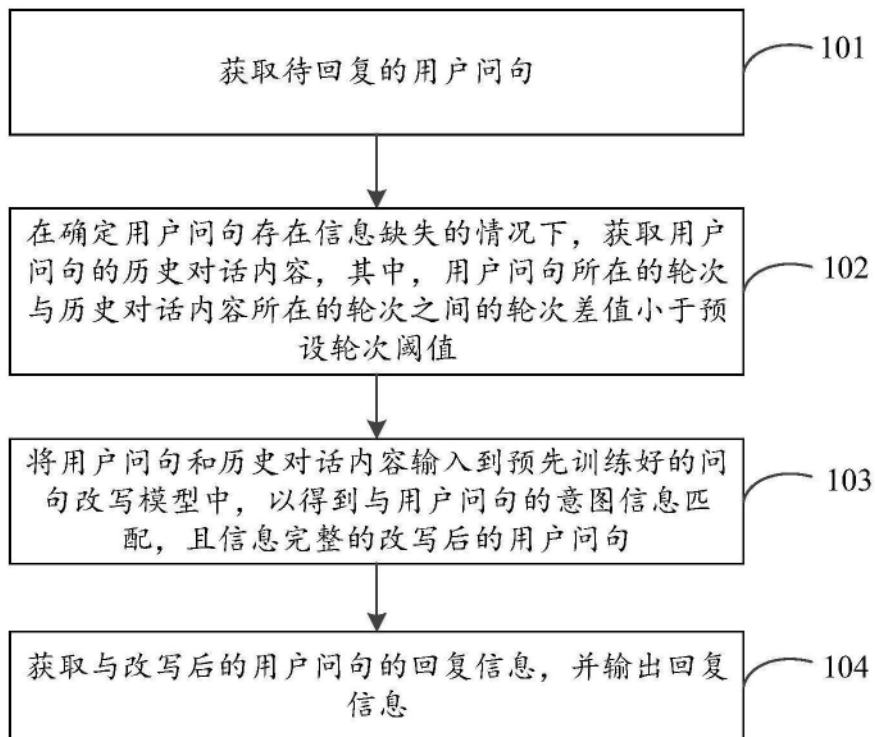


图1

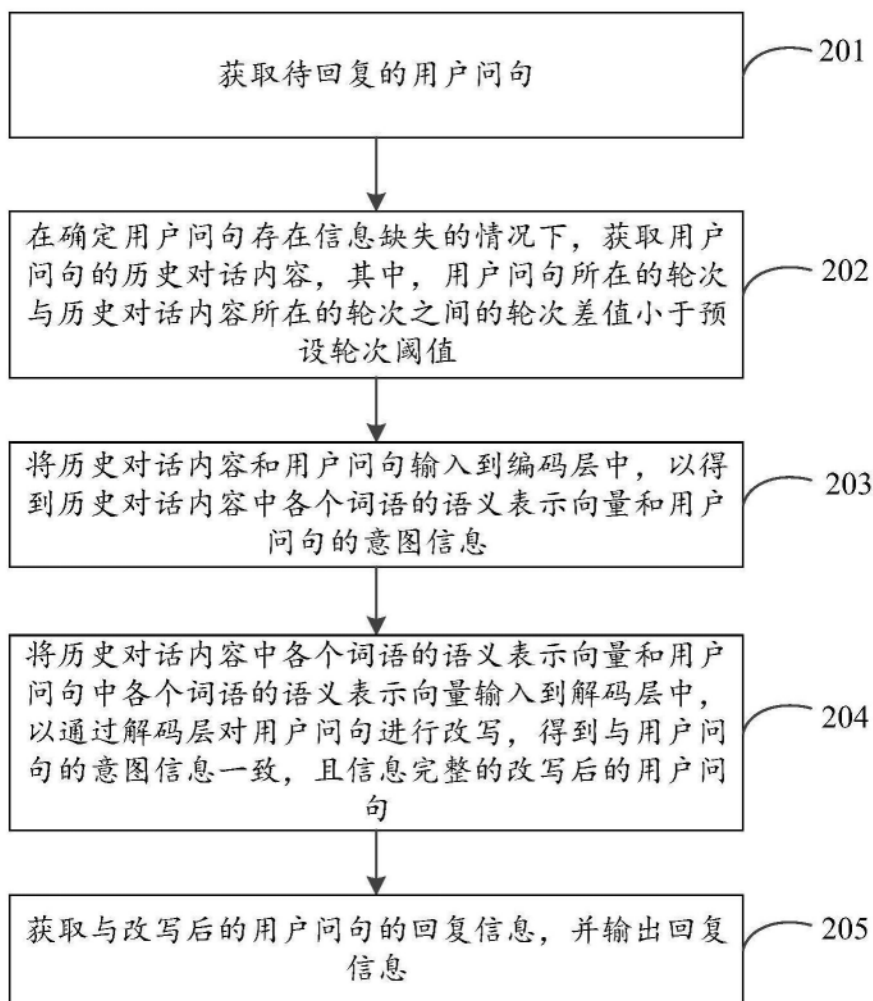


图2

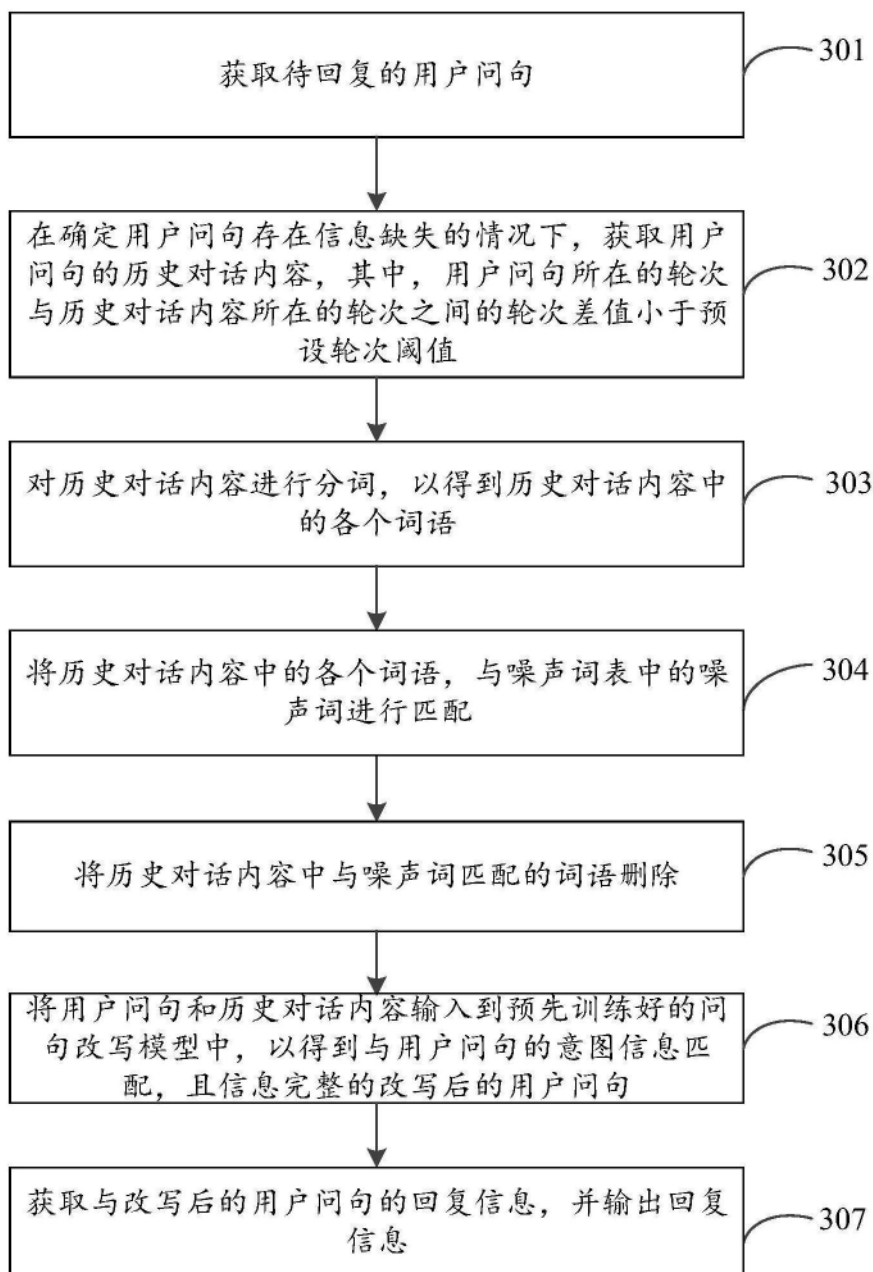


图3

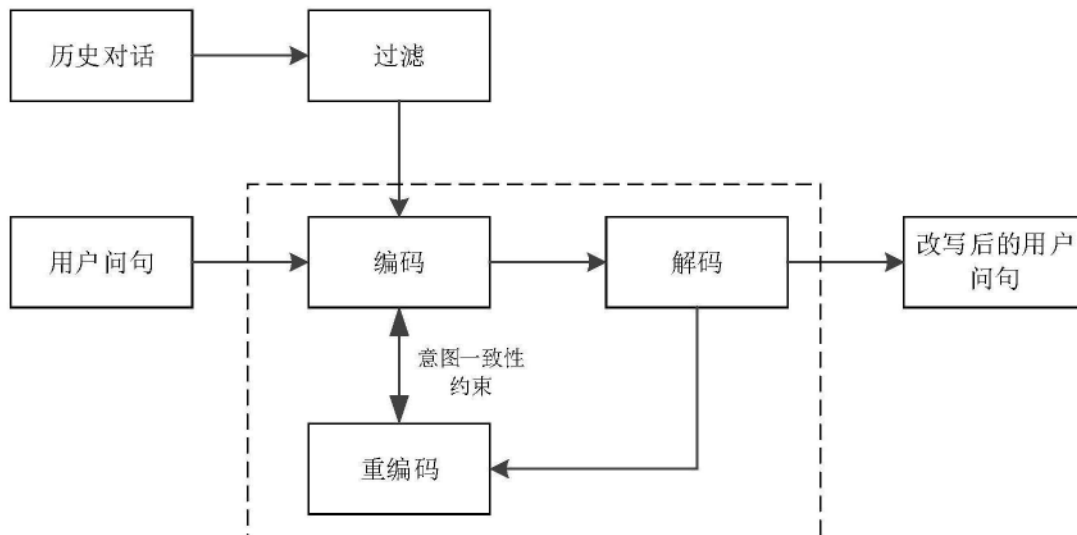


图4

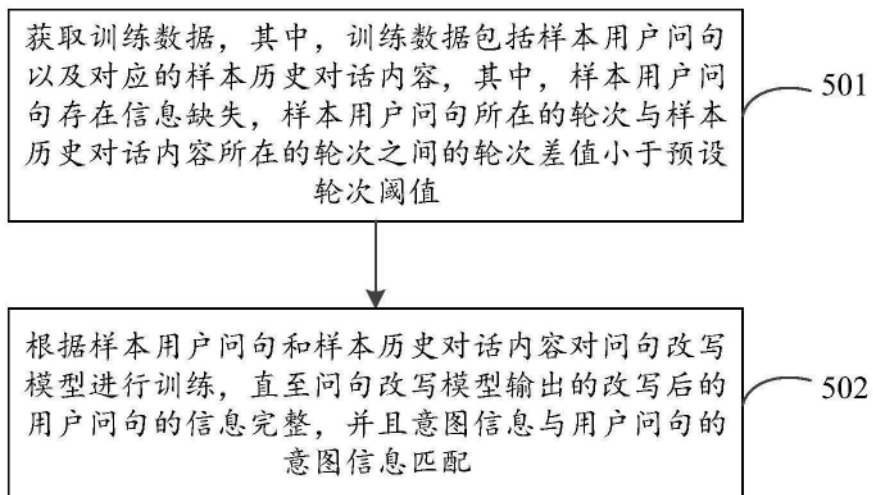


图5

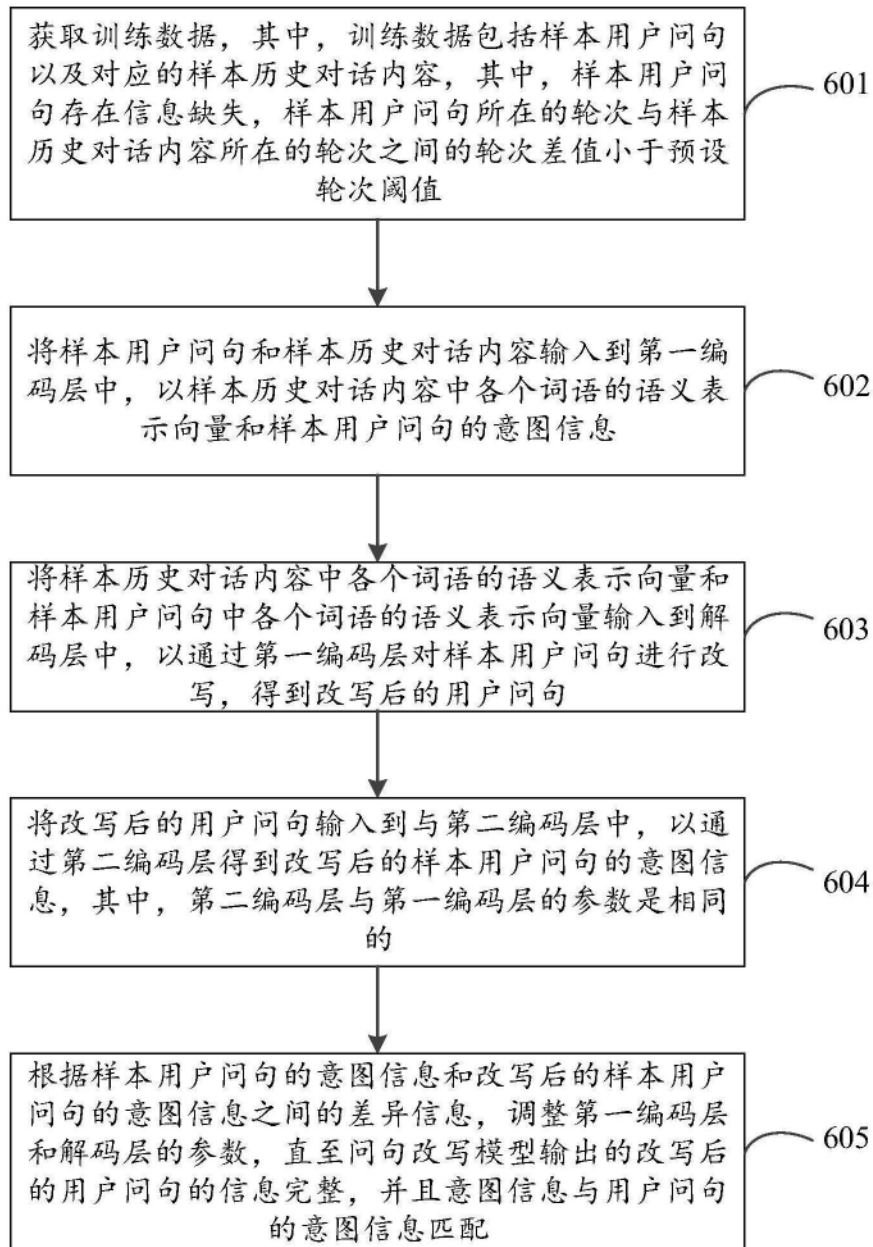


图6

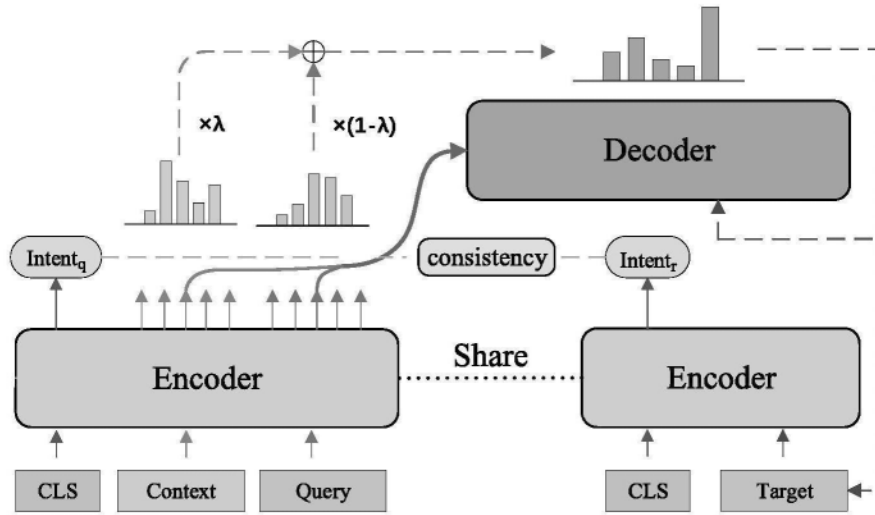


图7

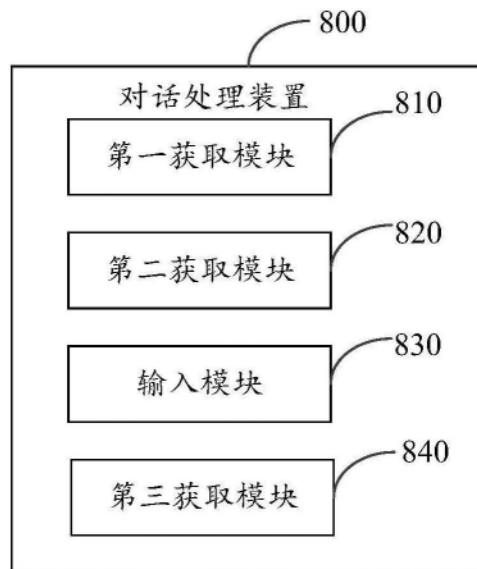


图8

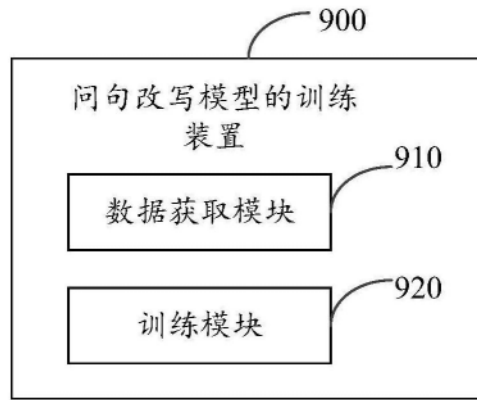


图9

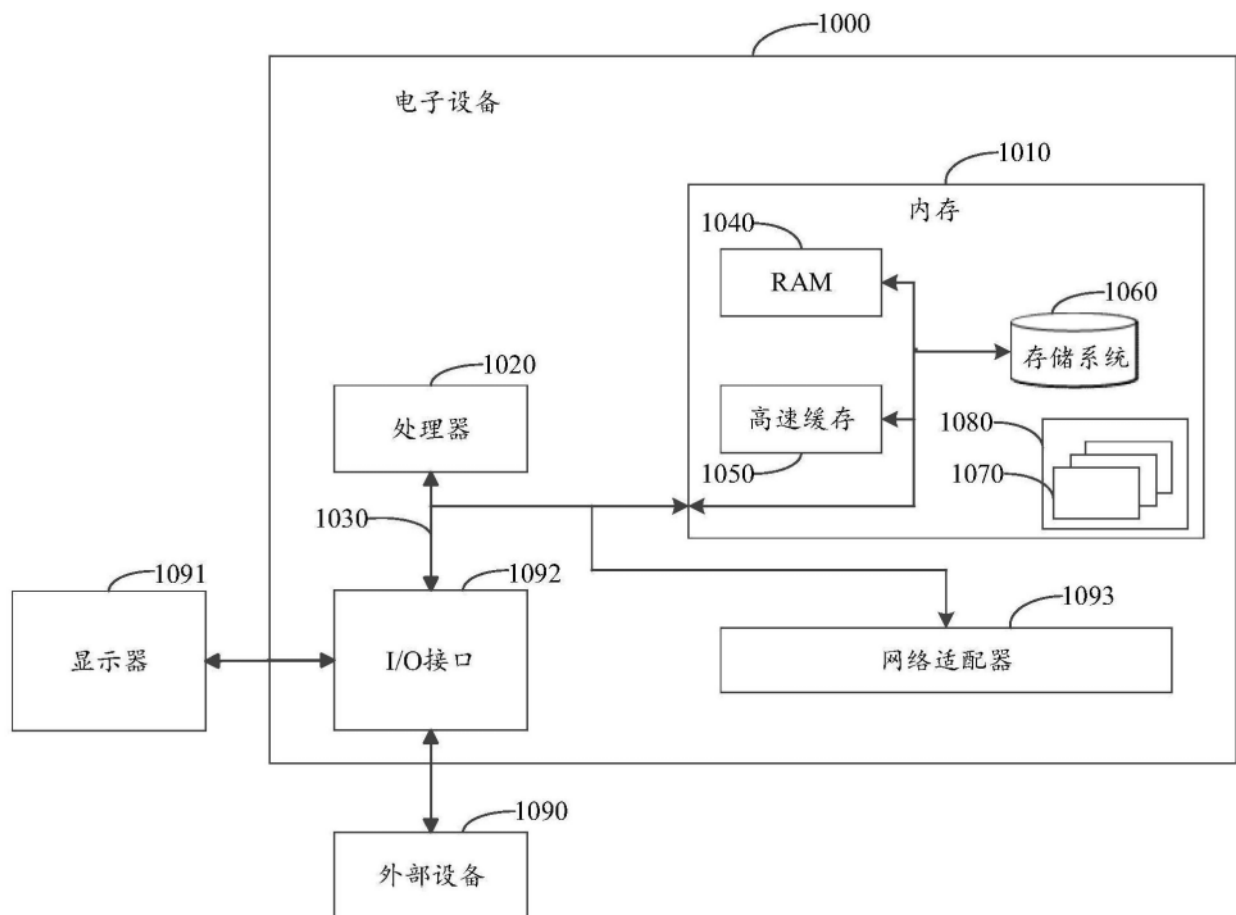


图10