



(12) 发明专利申请

(10) 申请公布号 CN 117763086 A

(43) 申请公布日 2024. 03. 26

(21) 申请号 202311771599.4

G06N 3/045 (2023.01)

(22) 申请日 2023.12.21

(71) 申请人 京东城市(北京)数字科技有限公司

地址 100086 北京市海淀区知春路76号(写字楼)1号楼9层1-7-5号

申请人 京东科技信息技术有限公司

(72) 发明人 申磊 毕冠群 陈蒙

(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038

专利代理师 曹蓓 许蓓

(51) Int. Cl.

G06F 16/33 (2019.01)

G06F 16/35 (2019.01)

G06F 16/332 (2019.01)

G06F 40/30 (2020.01)

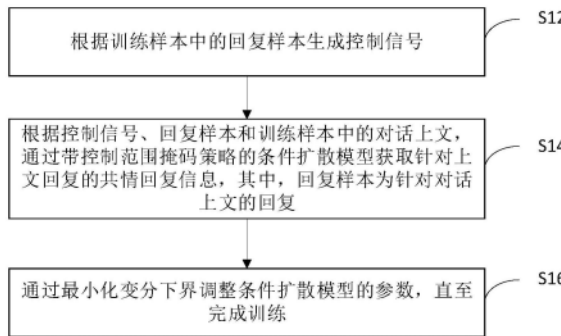
权利要求书3页 说明书11页 附图5页

(54) 发明名称

模型训练方法、回复信息确定方法及装置和
存储介质

(57) 摘要

本公开提出一种模型的训练方法、回复信息确定方法及装置和存储介质,涉及人工智能技术领域。本公开的一种模型训练方法包括:根据训练样本中的回复样本生成控制信号;根据控制信号、回复样本和训练样本中的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对上文回复的共情回复信息,其中,回复样本为针对对话上文的回复;通过最小化变分下界调整条件扩散模型的参数,直至完成训练。



1. 一种模型训练方法,包括:
根据训练样本中的回复样本生成控制信号;
根据所述控制信号、所述回复样本和训练样本中的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对所述对话上文的共情回复信息,其中,所述回复样本为针对所述对话上文的回复;
调整所述条件扩散模型的参数以使变分下界最小化,直至完成训练。
2. 根据权利要求1所述的训练方法,其中,所述根据训练样本中的回复样本生成控制信号包括:
通过分类器确定所述回复样本的分类标识;
根据所述分类标识生成所述控制信号。
3. 根据权利要求1所述的训练方法,其中,所述根据训练样本中的回复样本生成控制信号包括:
通过第一分类器确定所述回复样本的沟通机制标识;
通过第二分类器确定所述回复样本的意图标识;
通过第三分类器确定所述回复样本中每个词的语义框架标识;
根据所述沟通机制标识、所述意图标识和所述语义框架标识,生成所述控制信号。
4. 根据权利要求3所述的训练方法,其中,所述通过第一分类器确定所述回复样本的沟通机制标识包括:
根据第一子分类器确定所述回复样本的情绪反应标识;
根据第二子分类器确定所述回复样本的解释标识;
根据第三子分类器确定所述回复样本的探索标识;
根据所述情绪反应标识、所述解释标识和所述探索标识,生成所述沟通机制标识。
5. 根据权利要求1~4任意一项所述的训练方法,其中,通过带控制范围掩码策略的条件扩散模型获取针对所述对话上文的共情回复信息包括:
根据所述控制信号、所述回复样本和所述对话上文,生成第一样本信息;
通过所述条件扩散模型中的前向过程在所述第一样本信息中添加噪声,获取第二样本信息;
通过所述条件扩散模型中带控制范围掩码策略的反向过程,对所述第二样本信息去噪,获取所述共情回复信息。
6. 根据权利要求5所述的训练方法,其中,所述根据所述控制信号、所述回复样本和训练样本中的对话上文,生成第一样本信息包括:
通过拼接所述对话上文、所述控制信号和所述回复样本,生成离散文本;
将所述离散文本映射到连续表示,获取所述第一样本信息。
7. 根据权利要求5所述的训练方法,其中,所述通过所述条件扩散模型中的前向过程在所述第一样本信息中添加噪声,获取第二样本信息包括:
对所述第一样本信息中所述回复样本对应的信息添加噪声,获取所述第二样本信息。
8. 根据权利要求5所述的训练方法,其中,所述通过所述条件扩散模型中带控制范围掩码策略的反向过程,对所述第二样本信息去噪,获取所述共情回复信息包括:
基于Transformer模型对所述第二样本信息去噪,获取所述共情回复信息,其中,所述

Transformer模型的自注意力层中包括基于所述控制范围掩码策略的掩码矩阵。

9. 根据权利要求5所述的训练方法,其中,所述控制范围掩码策略包括:

通过在自注意力层设置的掩码矩阵,控制所述第二样本信息中信息之间的影响关系,
其中,在第二样本信息中的第一信息对第二信息产生影响的情况下,掩码矩阵中对应的元素为第一值,在第二样本信息中的第一信息对第二信息不产生影响的情况下,掩码矩阵中对应的元素为第二值。

10. 根据权利要求8或9所述的方法,其中,所述掩码矩阵中第1到n列和第1到n行对应所述控制信号,第n+1到第m列和第n+1到第m行对应所述回复样本,n、m为正整数,且 $m > n$ 。

11. 根据权利要求10所述的方法,其中,所述掩码矩阵中第1到k列和第1到k行对应所述控制信号中的沟通机制标识;

所述掩码矩阵中第k+1到j列和第k+1到j行对应所述控制信号中的意图标识;

所述掩码矩阵中第j+1到n列和第j+1到n行对应所述控制信号中的语义框架标识,

其中,k、j为正整数,且 $n > j > k$ 。

12. 根据权利要求9所述的方法,其中,所述通过在自注意力层设置的掩码矩阵,控制所述第二样本信息中信息之间的影响关系符合以下至少一项:

所述第二样本信息中信息对自身产生影响;

沟通机制标识对全部回复样本产生影响;

意图标识中的第一部分对回复样本中的第一部分产生影响,意图标识中的第二部分对回复样本中的第二部分产生影响;或

语义框架标识中的每个元素对回复样本中的一个或多个词语产生影响。

13. 一种回复信息确定方法,包括:

根据待处理的对话上文和训练数据集获取匹配的回复样本;

根据所述回复样本生成控制信号;

根据所述控制信号、所述回复样本和所述待处理的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对所述对话上文的共情回复信息,其中,所述条件扩散模型为通过权利要求1~12任意一项所述的模型训练方法训练生成。

14. 根据权利要求13所述的方法,其中,所述根据待处理的对话上文和训练数据集获取匹配的回复样本包括:

根据所述待处理的对话上文在所述训练数据集中查询,按照匹配度确定匹配的对话上文训练数据作为候选上文;

确定所述候选上文对应的回复样本训练数据,作为候选回复;

在所述候选回复中确定所述回复样本。

15. 根据权利要求14所述的方法,其中,所述按照匹配度确定匹配的对话上文训练数据作为候选上文包括:

根据待处理的对话上文与所述训练数据集中语句的第一相似度,按照所述第一相似度从高到低的顺序获取多个所述匹配的对话上文训练数据,作为所述候选上文;

所述在所述候选回复中确定所述回复样本包括:

获取所述待处理的对话上文与所述候选上文的情绪分布的第二相似度;

根据同一个所述候选上文对应的所述第一相似度和所述第二相似度的加权和,确定所

述候选上文对应的目标相似度；

根据所述候选文本对应的所述目标相似度，在所述候选回复中确定所述回复样本。

16. 根据权利要求14所述的方法，其中，所述在所述候选回复中确定所述回复样本包括：

在所述候选回复唯一的情况下，将所述候选回复作为所述回复样本。

17. 一种模型训练装置，包括：

控制信号生成单元，被配置为根据训练样本中的回复样本生成控制信号；

回复生成单元，被配置为根据所述控制信号、所述回复样本和训练样本中的对话上文，通过带控制范围掩码策略的条件扩散模型获取针对所述对话上文的共情回复信息，其中，所述回复样本为针对所述对话上文的回复；

调整单元，被配置为调整所述条件扩散模型的参数以使变分下界最小化，直至完成训练。

18. 一种回复信息确定装置，包括：

回复样本确定单元，被配置为根据待处理的对话上文和训练数据集获取匹配的回复样本；

控制信号确定单元，被配置为根据所述回复样本生成控制信号；

回复确定单元，被配置为根据所述控制信号、所述回复样本和所述待处理的对话上文，通过带控制范围掩码策略的条件扩散模型获取针对所述对话上文的共情回复信息，其中，所述条件扩散模型为通过权利要求1~12任意一项所述的模型训练方法训练生成。

19. 一种数据处理装置，包括：

存储器；以及

耦接至所述存储器的处理器，所述处理器被配置为基于存储在所述存储器的指令执行如权利要求1至16任一项所述的方法。

20. 一种计算机可读存储介质，其上存储有计算机程序指令，该指令被处理器执行时实现权利要求1至16任一项所述的方法的步骤。

模型训练方法、回复信息确定方法及装置和存储介质

技术领域

[0001] 本公开涉及人工智能技术领域,特别是一种模型训练方法、回复信息确定方法及装置和存储介质。

背景技术

[0002] 共情是开放域对话中的关键因素,展示了一个人对他人处境和心理状况的关心。共情回复生成作为一种条件文本生成任务,旨在赋予模型理解说话者并准确表达共情的能力。

[0003] 共情回复生成往往会面临“单调共情”的问题,即表现出浅层的、通用的共情并与对话上文的相关性较弱。例如,“对此我感到很抱歉”经常被用作对各种带有负面情绪的上文的回复。为了缓解这个问题,相关技术中主要将情绪或认知模块融入到编码器-解码器框架中,并使用最大似然估计训练模型。

发明内容

[0004] 本公开的一个目的在于提高共情回复的灵活度和准确度。

[0005] 根据本公开的一些实施例的一个方面,提出一种模型训练方法,包括:根据训练样本中的回复样本生成控制信号;根据控制信号、回复样本和训练样本中的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息,其中,回复样本为针对对话上文的回复;调整条件扩散模型的参数以使变分下界最小化,直至完成训练。

[0006] 在一些实施例中,根据训练样本中的回复样本生成控制信号包括:通过分类器确定回复样本的分类标识;根据分类标识生成控制信号。

[0007] 在一些实施例中,根据训练样本中的回复样本生成控制信号包括:通过第一分类器确定回复样本的沟通机制标识;通过第二分类器确定回复样本的意图标识;通过第三分类器确定回复样本中每个词的语义框架标识;根据沟通机制标识、意图标识和语义框架标识,生成控制信号。

[0008] 在一些实施例中,通过第一分类器确定回复样本的沟通机制标识包括:根据第一子分类器确定回复样本的情绪反应标识;根据第二子分类器确定回复样本的解释标识;根据第三子分类器确定回复样本的探索标识;根据情绪反应标识、解释标识和探索标识,生成沟通机制标识。

[0009] 在一些实施例中,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息包括:根据控制信号、回复样本和对话上文,生成第一样本信息;通过条件扩散模型中的前向过程在第一样本信息中添加噪声,获取第二样本信息;通过条件扩散模型中带控制范围掩码策略的反向过程,对第二样本信息去噪,获取共情回复信息。

[0010] 在一些实施例中,根据控制信号、回复样本和训练样本中的对话上文,生成第一样本信息包括:通过拼接对话上文、控制信号和回复样本,生成离散文本;将离散文本映射到连续表示,获取第一样本信息。

[0011] 在一些实施例中,通过条件扩散模型中的前向过程在第一样本信息中添加噪声,获取第二样本信息包括:对第一样本信息中回复样本对应的信息添加噪声,获取第二样本信息。

[0012] 在一些实施例中,通过条件扩散模型中带控制范围掩码策略的反向过程,对第二样本信息去噪,获取共情回复信息包括:基于Transformer模型对第二样本信息去噪,获取共情回复信息,其中,Transformer模型的自注意力层中包括基于控制范围掩码策略的掩码矩阵。

[0013] 在一些实施例中,控制范围掩码策略包括:通过在自注意力层设置的掩码矩阵,控制第二样本信息中信息之间的影响关系,其中,在第二样本信息中的第一信息对第二信息产生影响的情况下,掩码矩阵中对应的元素为第一值,在第二样本信息中的第一信息对第二信息不产生影响的情况下,掩码矩阵中对应的元素为第二值。

[0014] 在一些实施例中,掩码矩阵中第1到n列和第1到n行对应控制信号,第n+1到第m列和第n+1到第m行对应回复样本,n、m为正整数,且 $m > n$ 。

[0015] 在一些实施例中,掩码矩阵中第1到k列和第1到k行对应控制信号中的沟通机制标识;掩码矩阵中第k+1到j列和第k+1到j行对应控制信号中的意图标识;掩码矩阵中第j+1到n列和第j+1到n行对应控制信号中的语义框架标识,其中,k、j为正整数,且 $n > j > k$ 。

[0016] 在一些实施例中,通过在自注意力层设置的掩码矩阵,控制第二样本信息中信息之间的影响关系符合以下至少一项:第二样本信息中信息对自身产生影响;沟通机制标识对全部回复样本产生影响;意图标识中的第一部分对回复样本中的第一部分产生影响,意图标识中的第二部分对回复样本中的第二部分产生影响;或语义框架标识中的每个元素对回复样本中的一个或多个词语产生影响。

[0017] 根据本公开的一些实施例的一个方面,提出一种回复信息确定方法,包括:根据待处理的对话上文和训练数据集获取匹配的回复样本;根据回复样本生成控制信号;根据控制信号、回复样本和待处理的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息,其中,条件扩散模型为通过上文中提到的任意一种模型训练方法训练生成。

[0018] 在一些实施例中,根据待处理的对话上文和训练数据集获取匹配的回复样本包括:根据待处理的对话上文在训练数据集中查询,按照匹配度确定匹配的对话上文训练数据作为候选上文;确定候选上文对应的回复样本训练数据,作为候选回复;在候选回复中确定回复样本。

[0019] 在一些实施例中,按照匹配度确定匹配的对话上文训练数据作为候选上文包括:根据待处理的对话上文与训练数据集中语句的第一相似度,按照第一相似度从高到低的顺序获取多个匹配的对话上文训练数据,作为候选上文;在候选回复中确定回复样本包括:获取待处理的对话上文与候选上文的情绪分布的第二相似度;根据同一个候选上文对应的第一相似度和第二相似度的加权和,确定候选上文对应的目标相似度;根据候选文本对应的目标相似度,在候选回复中确定回复样本。

[0020] 在一些实施例中,在候选回复中确定回复样本包括:在候选回复唯一的情况下,将候选回复作为回复样本。

[0021] 根据本公开的一些实施例的一个方面,提出一种模型训练装置,包括:控制信号生

成单元,被配置为根据训练样本中的回复样本生成控制信号;回复生成单元,被配置为根据控制信号、回复样本和训练样本中的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息,其中,回复样本为针对对话上文的回复;调整单元,被配置为调整条件扩散模型的参数以使变分下界最小化,直至完成训练。

[0022] 根据本公开的一些实施例的一个方面,提出一种回复信息确定装置,包括:回复样本确定单元,被配置为根据待处理的对话上文和训练数据集获取匹配的回复样本;控制信号确定单元,被配置为根据回复样本生成控制信号;回复确定单元,被配置为根据控制信号、回复样本和待处理的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息,其中,条件扩散模型为通过上文中任意一种模型训练方法训练生成。

[0023] 根据本公开的一些实施例的一个方面,提出一种数据处理装置,包括:存储器;以及耦接至存储器的处理器,处理器被配置为基于存储在存储器的指令执行上文中提到的任意一种方法。

[0024] 根据本公开的一些实施例的一个方面,提出一种计算机可读存储介质,其上存储有计算机程序指令,该指令被处理器执行时实现上文中提到的任意一种方法的步骤。

附图说明

[0025] 此处所说明的附图用来提供对本公开的进一步理解,构成本公开的一部分,本公开的示意性实施例及其说明用于解释本公开,并不构成对本公开的不当限定。在附图中:

[0026] 图1为本公开的模型训练方法的一些实施例的流程图。

[0027] 图2为本公开的模型训练方法的另一些实施例的示意图。

[0028] 图3为本公开的模型训练方法中掩码矩阵的一些实施例的示意图。

[0029] 图4为本公开的回复内容确定方法的一些实施例的流程图。

[0030] 图5为本公开的回复内容确定方法中确定回复样本的一些实施例的流程图。

[0031] 图6为本公开的模型训练装置的一些实施例的示意图。

[0032] 图7为本公开的回复内容确定装置的一些实施例的示意图。

[0033] 图8为本公开的数据处理装置的一些实施例的示意图。

[0034] 图9为本公开的数据处理装置的另一些实施例的示意图。

具体实施方式

[0035] 下面通过附图和实施例,对本公开的技术方案做进一步的详细描述。

[0036] 发明人发现,相关技术中的共情回复生成模型存在如下问题:

[0037] 1) 相关技术中共情回复生成的思路是,首先对对话上文进行向量表示,接着融合情绪和认知因素到解码器初始状态或者利用融合因素平等地影响每个解码步。共情内容的产生依赖基于对话上文的预测使得回复生成缺乏灵活性,且对每个解码步相同程度的干预会对生成内容的流畅性造成损失。

[0038] 2) 相关技术中共情回复生成模型主要基于编码器-解码器框架和最大似然估计,天然容易对训练数据中的高频通用回复有更多的生成倾向。

[0039] 针对上述问题,本公开提出一种模型训练方法、回复信息确定方法,及对应的装

置、数据处理装置和存储介质,以解决上述问题,提高共情回复的灵活度和准确度。

[0040] 本公开的模型训练方法的一些实施例的流程图如图1所示。

[0041] 在步骤S12中,根据训练样本中的回复样本生成控制信号。

[0042] 在一些实施例中,训练样本可以包括基于人与人之间的真实对话采集的样本,从而提高后续共情分析的可靠性。在一些实施例中,训练样本中包括对话上文和回复样本,对话上文与回复样本一一对应。在一些实施例中,训练样本中包括对话内容,根据对话内容中信息的先后顺序确定相对应的对话上文和回复样本。

[0043] 在一些实施例中,可以通过对回复样本的情感、数据分析获取控制信号。在一些实施例中,可以通过分类器确定回复样本的分类标识,再根据分类标识生成控制信号。

[0044] 在一些实施例中,通过第一分类器确定回复样本的CM (Communication Mechanism,沟通机制) 标识,控制信号中包括CM标识。在一些实施例中,CM标识可以由多个部分组成,第一分类器中包括多个子分类器,根据第一子分类器确定回复样本的ER (Emotional Response,情绪反应) 标识,根据第二子分类器确定回复样本的IP (Interpretations,解释) 标识,根据第三子分类器确定回复样本的EX (Explorations,探索) 标识,根据情绪反应标识、解释标识和探索标识,生成沟通机制标识。ER表示如温暖,同情和关心等情绪,IP代表对说话者描述的感觉和经历的理解,EX代表探索在之前的对话中未提到的感觉和经历。通过这样的方法,能够从情绪反应、说话者描述的感觉和经历的理解,以及探索在之前的对话中未提到的感觉和经历三个维度进行回复样本的分析,使控制信号具备从上述三个维度进行控制的功能,提高控制的准确度和灵活控制能力。

[0045] 在一些实施例中,考虑到人们在表达共情时表现出了广泛的意图,即倾听者更有可能以特定的共情意图,如“承认”,“安慰”和“鼓励”来回复积极或消极的情绪,而不仅仅是表达相似或相反的情绪,因此通过第二分类器对回复样本的意图进行识别。在一些实施例中,可以基于数据集(例如EmpatheticIntents数据集)对构建的分类器(例如BERT分类器)进行预训练,利用生成的分类器处理回复样本,得到IT (Intent,意图) 标识。控制信号中包括IT标识。通过这样的方法,能够从说话者的意图的角度进行回复样本分析,使得控制信号具备控制共情回复的意图的功能,提高控制的准确度和灵活控制能力。

[0046] 在一些实施例中,一个词的框架代表其所属事件,概念或关系的通用类别,可以被视为词汇意义的抽象。在一些实施例中,语义框架基于FrameNet,一个包含词汇和谓词-论元语义信息的语言知识图。例如,“鸟”,“猫”和“狗”这样的词共享相同的框架标签“动物”。在一些实施例中,可以利用open-SESAME模型对回复样本中的每个词提取语义框架,生成每个词的SF (Semantic Frame,语义框架) 标识。控制信号中包括SF标识。通过这样的方法,能够从每个词的语义框架的角度进行回复样本分析,使得控制信号具备控制每个词的语义框架的功能,提高控制的准确度和灵活控制能力。

[0047] 在一些实施例中,可以通过多个维度的分析确定每个维度的分类标识,进而综合各个维度的分类标识,生成控制信号。例如,通过第一分类器确定回复样本的CM标识;通过第二分类器确定回复样本的IT标识;通过第三分类器确定回复样本中每个词的SF标识;根据CM标识、IT标识和SF标识,生成控制信号。

[0048] 通过这样的方法,能够对回复样本进行多维度分析,提高控制信号的准确度和灵活控制能力。

[0049] 在步骤S14中,根据控制信号、回复样本和训练样本中的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息。

[0050] 在一些实施例中,条件扩散模型可以包括前向过程和反向过程,前向过程对输入的信息添加噪声,反向过程对添加了噪声的信息进行去噪,得到去噪后的结果作为共情回复信息,其中,在反向过程中执行控制范围掩码策略,实现对生成的共情回复信息的控制。

[0051] 在一些实施例中,可以先根据控制信号、回复样本和对话上文,生成第一样本信息,在一些实施例中,先通过拼接对话上文、控制信号和回复样本,生成离散文本,进而将离散文本映射到连续表示,获取第一样本信息。进一步的,通过条件扩散模型中的前向过程在第一样本信息中添加噪声,获取第二样本信息。在一些实施例中,可以仅对第一样本信息中回复样本对应的信息添加噪声,不对对话上文、控制信号对应的信息添加噪声,获取第二样本信息。通过这样的方法,能够提高噪声添加的针对性,降低需要处理的信息量,避免对处理中需要参考的对话上文和控制信号产生影响,提高生成的共情回复信息的准确度和生成速度。在获得第二样本信息后,通过条件扩散模型中带控制范围掩码策略的反向过程,对第二样本信息去噪,获取共情回复信息。

[0052] 通过这样的方法,能够将控制信号、回复样本和对话上文均纳入条件扩散模型的考虑范围,提高处理的信息的全面性,进而提高生成的共情回复信息的准确度。

[0053] 在一些实施例中,可以通过在自注意力层设置的掩码矩阵,控制第二样本信息中信息之间的影响关系,从而实现对生成共情回复信息的控制。在一些实施例中,条件扩散模型发反向过程采用Transformer模型,Transformer模型的自注意力层中包括基于控制范围掩码策略的掩码矩阵,基于Transformer模型对第二样本信息去噪,获取共情回复信息。

[0054] 在一些实施例中,在第二样本信息中的第一信息对第二信息产生影响的情况下,掩码矩阵中对应的元素为第一值,在第二样本信息中的第一信息对第二信息不产生影响的情况下,掩码矩阵中对应的元素为第二值。在一些实施例中,第一值为0,第二值为负无穷。通过这样的方法,能够灵活控制样本信息中各类信息对生成共情回复信息的控制能力,提高控制灵活性。

[0055] 在一些实施例中,掩码矩阵中第1到n列和第1到n行对应控制信号,第n+1到第m列和第n+1到第m行对应回复样本,n、m为正整数,且 $m > n$ 。在一些实施例中,掩码矩阵中第1到k列和第1到k行对应控制信号中的沟通机制标识;掩码矩阵中第k+1到j列和第k+1到j行对应控制信号中的意图标识;掩码矩阵中第j+1到n列和第j+1到n行对应控制信号中的语义框架标识,其中,k、j为正整数,且 $n > j > k$ 。通过设置掩码矩阵中的值,设置各类信息的相互影响。在一些实施例中,掩码矩阵中的元素 A_{ij} 标识第i行对应的信息对第j列的信息的影响。

[0056] 通过这样的方法,能够使用掩码矩阵对控制信号、回复样本和对话上文进行全面且灵活的参考,提高生成共情回复信息的准确度和控制灵活性。

[0057] 在一些实施例中,可以通过设置掩码矩阵中的元素的值,设置第二样本信息中信息对自身产生影响。在一些实施例中,设置掩码矩阵中的元素的值使其符合以下至少一项:沟通机制标识对全部回复样本产生影响;意图标识中的第一部分对回复样本中的第一部分产生影响,意图标识中的第二部分对回复样本中的第二部分产生影响;语义框架标识中的每个元素对回复样本中的一个或多个词语产生影响。通过这样的方法,同一粒度/级别的词汇内部,包括IT信号词、SF信号词和回复,被设计为相互控制,从而确保生成回复的流畅性。

[0058] 在步骤S16中,通过最小化变分下界调整条件扩散模型的参数,直至完成训练。在一些实施例中,在训练过程中,条件扩散模型的反向过程的参数得到调整。

[0059] 基于上文所示实施例中的方法,将共情回复生成转化为一个可控文本生成任务,引入显式的多粒度控制信号来解决单调共情问题,在条件扩散模型中统一利用对话上文信息和控制信号,实现了对共情回复内容的指导,提高了共情回复的灵活度和准确度。

[0060] 本公开的模型训练方法的另一些实施例的示意图如图2所示。

[0061] 扩散模型23的输入包括三部分:对话上文211,控制信号212和回复样本213。其中,对话上文是一个来自说话者和倾听者的交替话语 u_i 序列,即 $w^u = \{u_1, u_2, \dots, u_n\}$ 。控制信号用 w^c 表示,回复样本用 $w^y = \{y_1, y_2, \dots, y_n\}$ 表示。扩散模型23由一个正向过程和一个反向过程构成,反向过程的Transformer块中包含了控制范围掩码策略,用于精细指导回复内容的生成。在训练过程中,控制信号从标准回复中提取,而在推理过程中,通过一个情绪增强的匹配方法获得恰当的回复候选去替代标准回复。

[0062] (1) 获取控制信号

[0063] 共情是一个多维度的信息,一般认为受多个因素共同影响。本公开中将这些因素显式地表示为不同粒度的控制信号,从而引导共情回复生成。在一些实施例中,如图2所示,对回复样本213进行信息提取,获取控制信号。在一些实施例中,从“沟通机制”、“意图”和“语义框架”三个角度进行分析,这三个角度分别关注话语、句子和词三种从粗到细的粒度,从而提高控制信号的全面性。这样的信号提取角度仅作为示例,不构成对本公开的不当限制。

[0064] I、沟通机制CM

[0065] 共情回复相关的沟通机制包括“情绪反应(ER)”,“解释(IP)”和“探索(EX)”。其中,ER表示如温暖,同情和关心等情绪,IP代表对说话者描述的感觉和经历的理解,EX代表探索在之前的对话中未提到的感觉和经历。在一些实施例中,可以使用三个预训练的RoBERTa分类器来分别识别整个回复是否暗含了某种机制。

[0066] II、意图IT

[0067] 已有分析表明,人们在表达共情时表现出了广泛的意图[3],即倾听者更有可能以特定的共情意图,如“承认”,“安慰”和“鼓励”来回复积极或消极的情绪,而不仅仅是表达相似或相反的情绪。在一些实施例中,通过在EmpatheticIntent数据集上训练的BERT分类器来对回复中句子的意图进行识别。

[0068] III、语义框架SF

[0069] 语义框架基于FrameNet,一个包含词汇和谓词-论元语义信息的语言知识图。一个词的框架代表其所属事件,概念或关系的通用类别,可以被视为词汇意义的抽象。例如,“鸟”,“猫”和“狗”这样的词共享相同的框架标签“动物”。在一些实施例中使用open-SESAME模型为回复中每个词提取语义框架。

[0070] 以上三种标记模型的性能如表1所示,表1中的数据仅作为标识示例,不构成对本公开的不当限制。

[0071] 表1用于获取控制信号中的标识的模型的性能

[0072]	标识类型	准确率	F1 指标	等级
	CM-ER	79.43	74.46	2
	CM-IP	84.04	62.60	2
	CM-EX	92.61	72.58	2
	IT	87.85	87.71	9
	SF		86.55	1222

[0073] 在得到各个维度的标识后,将标识按照预定顺序拼接为一个序列,如按照从粗到细粒度拼接为一个序列,生成控制信号。

[0074] (2) 带控制范围掩码策略的条件扩散模型

[0075] 扩散模型包含一个前向过程和一个反向过程。首先将对话上文 w^u 、控制信号 w^c 和相应的回复样本 w^y 拼接起来,即 $w = w^u \oplus w^c \oplus w^y$ 。使用嵌入函数 $\text{Emb}(\cdot)$ 将离散文本 w 映射到连续表示 x_0 , $x_0 = u_0 \oplus c_0 \oplus y_0$,其中, u_0 , c_0 和 y_0 分别代表 x_0 中属于 w^u , w^c 和 w^y 的部分。 x_0 作为第一样本信息。

[0076] I、前向过程

[0077] 在前向过程中,模型基于公式(1)逐步向 x_0 添加噪声:

$$[0078] \quad q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t I) \quad (1)$$

[0079] $x_1, \dots, x_T$ 构成一个马尔可夫链, $x_T \sim \mathcal{N}(0, I)$ 。 $\beta_t \in (0, 1)$ 是一个噪声调度,控制每一步添加的噪声规模。

[0080] 在一些实施例中,考虑到共情回复生成是一个条件文本生成(序列到序列)任务,只关心回复信息的生成效果。因此,可以使用“部分噪声”,只对 x_t 中 w^y 对应的部分 y_t 施加噪声。

[0081] II、反向过程

[0082] 在前向过程完成后,反向过程的基于公式(2)逐步通过对 x_T 去噪来恢复 x_0 。

$$[0083] \quad p_\theta(x_{t-1}|x_t, t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \sigma_\theta(x_t, t)) \quad (2)$$

[0084] $\mu_\theta(\cdot)$ 和 $\sigma_\theta(\cdot)$ 是前向过程 $q(x_{t-1}|x_t)$ 的预测均值和标准差,可以通过Transformer模型 f_θ 实现。在一些实施例中,在反向过程中,添加一个舍入步骤,参数化为 $p_\theta(w|x_0) = \prod_{i=1}^n p_\theta(w_i|x_i)$, $\prod_{i=1}^n p_\theta(w_i|x_i)$ 是一个softmax分布。

[0085] III、控制范围掩码策略

[0086] 本公开的模型训练方法中掩码矩阵的一些实施例的示意图如图3所示,该图中的语句和矩阵中的值仅作为示例,不构成对本公开的不当限制。图3中左上的语句部分显示了控制信号和回复样本(“-”表示SF信号为空),以回复样本为“Sounds great!What is your gift?”为例,对其进行控制信号分析后,得到的控制信号如左上部分1-11所示,该示例仅为控制信号的标识,不代表控制信号中的具体内容。左上部分的12-19为回复样本中的词语。图3中左下部分表示每个信号在回复不同粒度上的控制范围,图3中右侧是相应的掩码矩阵M。对于意图信号Acknowledging(第2行),对Questioning(第3行)和相应的第一句回复

Sounds great! (第12-14行) 可见。同时, 由于great (第13行) 受到Exploration (第1行)、Acknowledge (第2行)、Desirability (第5行) 和其余回复词 (第12-19行) 的控制, 在掩码矩阵中会关注到这些内容。

[0087] 相关技术中的扩散模型的非自回归性质使得一个输入词在表示更新的时候, 可以通过完全的自注意力机制关注所有其他的词。本公开的扩散模型中区分控制信号和回复, 将掩码矩阵M整合到Transformer自注意力层中:

$$[0088] \quad Q^{i+1}, K^{i+1}, V^{i+1} = h^i W_q, h^i W_k, h^i W_v,$$

$$[0089] \quad S^{i+1} = \text{softmax} \left(\frac{Q^{i+1} K^{i+1 T} + M}{\sqrt{d_k}} \right),$$

$$[0090] \quad h^{i+1} = S^{i+1} V^{i+1},$$

[0091] 其中, W_q, W_k 和 W_v 是可训练的参数, h^i 是第i层Transformer的隐状态。

[0092] 针对控制范围掩码策略的设计, 如果信号i对应的词控制词汇j的生成, 那么j的计算就会受到i的影响。在更新j的表示时不会掩蔽i。同一粒度/级别的词汇内部, 包括IT信号词、SF信号词和回复, 也被设计为相互控制, 从而确保生成回复的流畅性。通过控制范围掩码策略, 并将其用于Transformer模型 f_θ 的自注意力层。对于一个掩码矩阵M, 如果词汇j被词汇i控制, 那么其元素 A_{ij} 的值为0; 否则为负无穷:

$$[0093] \quad A_{ij} = \begin{cases} 0, & i \Rightarrow j \\ -\text{inf}, & i \nRightarrow j \end{cases}$$

[0094] (3) 训练过程

[0095] 在训练过程中, 根据“获取控制信号”中描述的方式从标准回复中提取控制信号。最小化变分下界 $\mathcal{L}_{\text{vib}}$, 使得反向过程的模型 f_θ 得到训练。

$$[0096] \quad \mathcal{L}_{\text{vib}} = \sum_{t=2}^T \|y_0 - \widetilde{f}_\theta(x_t, t)\|^2 + \|\text{EMB}(w^y - \widetilde{f}_\theta(x_1, 1))\|^2 + \mathcal{R}(\|x_0\|^2),$$

[0097] 其中, $\widetilde{f}_\theta(x_t, t)$ 表示恢复的 x_0 , $\mathcal{R}(\cdot)$ 为正则化项, $\text{EMB}(\cdot)$ 为词嵌入模块, 负责把文本表示为向量, y_0 为经过 $\text{EMB}(\cdot)$ 编码后的回复表示。

[0098] 在实现模型训练后, 可以采用训练完成后的模型进行测试, 或投入使用, 确定待处理的对话上文的共情回复信息。本公开的回复内容确定方法的一些实施例的流程图如图4所示。

[0099] 在步骤S41中, 根据待处理的对话上文和训练数据集获取匹配的回复样本。

[0100] 在一些实施例中, 考虑到待处理的对话上文所对应的回复为未知信息, 可以基于训练数据集获得与待处理的对话上文相匹配的回复样本, 用于生成控制信号。

[0101] 在一些实施例中, 可以将待处理的对话上文在训练数据集中进行查询, 按照匹配度确定匹配的对话上文训练数据, 将其作为候选上文。进一步的, 将候选上文对应的回复样本训练数据候选回复, 在候选回复中确定回复样本。在一些实施例中, 候选上文为匹配度最高的唯一对话上文训练数据, 则候选回复作为回复样本, 执行后续操作。

[0102] 在一些实施例中, 可以基于如图5所示的方法确定回复样本。

[0103] 在步骤511中, 根据待处理的对话上文与训练数据集中语句的第一相似度 (语义相

似度),按照第一相似度从高到低的顺序获取多个匹配的对话上文训练数据,作为候选上文。在一些实施例中,可以获得第一相似度最高的多条对话上文训练数据作为候选上文;在一些实施例中,可以按照第一相似度从高到低的顺序,获取预定数量、预定比例或获取第一相似度高于预定阈值的对话上文训练数据作为候选上文。

[0104] 在步骤512中,确定候选上文对应的回复样本训练数据,作为候选回复。在一些实施例中,步骤512可以省略,在步骤515中执行。

[0105] 在步骤513中,获取待处理的对话上文与候选上文的情绪分布的第二相似度。在一些实施例中,可以预训练机器学习模型(例如BERT模型)用于对对话上文进行情绪分类,使用训练完成的模型获取待处理的对话上文与候选上文的情绪分布,确定待处理的对话上文与各条候选上文的情绪分布的第二相似度(如余弦相似度)。

[0106] 在步骤514中,根据同一个候选上文对应的第一相似度 $SIM_{semantic}$ 和第二相似度 $SIM_{emotional}$ 的加权和,如公式(3)所示,确定候选上文对应的目标相似度Score,对候选上文按照目标相似度从高到低的顺序进行排序。

[0107] $Score = SIM_{semantic} + \gamma SIM_{emotional}$ (3)

[0108] γ 是一个超参数,用于平衡语义和情绪相似性

[0109] 在步骤515中,根据候选文本对应的目标相似度,在候选回复中确定回复样本。在一些实施例中,可以基于步骤512中确定的候选文本对应的候选回复,对候选回复进行排序,确定目标相似度最高的候选文本对应的候选回复作为回复样本。在一些实施例中,可以跳过步骤512,确定目标相似度最高的候选文本对应的候选回复作为回复样本。

[0110] 基于上文所示实施例中的方法,能够基于语义相似性和情绪一致性两个维度确定回复样本,从而提高控制信号的适配度,进而提高后续确定的共情回复信息的准确度。

[0111] 在步骤S43中,根据回复样本生成控制信号。在一些实施例中,根据回复样本生成控制信号的操作可以如上文步骤S12中所示。

[0112] 在步骤S45中,根据控制信号、回复样本和待处理的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息。条件扩散模型为通过上文中提到的任意一种训练方法训练生成。在一些实施例中,根据控制信号、回复样本和待处理的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息的操作可以如上文步骤S14中所示。

[0113] 通过上文所示实施例中的方法,将共情回复生成转化为可控文本生成任务,并引入显式的多粒度控制信号来解决单调共情问题,在条件扩散模型中统一利用对话上文信息和控制信号;通过控制范围掩码策略实现了对共情回复内容的指导,并整合了一个情绪增强的匹配方法来为给定的上文产生多样的回复,提高了共情回复的灵活度和准确度。

[0114] 本公开的模型训练装置61的一些实施例的示意图如图6所示。

[0115] 控制信号生成单元611能够根据训练样本中的回复样本生成控制信号。在一些实施例中,训练样本可以包括基于人与人之间的真实对话采集的样本,从而提高后续共情分析的可靠性。在一些实施例中,训练样本中包括对话上文和回复样本,对话上文与回复样本一一对应。在一些实施例中,训练样本中包括对话内容,根据对话内容中信息的先后顺序确定相对应的对话上文和回复样本。在一些实施例中,控制信号生成单元611能够执行上文步骤S12中任一实施例中的操作。

[0116] 回复生成单元612能够根据控制信号、回复样本和训练样本中的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息。在一些实施例中,回复生成单元612能够执行上文步骤S14中任一实施例中的操作。

[0117] 调整单元613能够通过最小化变分下界调整条件扩散模型的参数,直至完成训练。在一些实施例中,在训练过程中,条件扩散模型的反向过程的参数得到调整。

[0118] 这样的装置能够将共情回复生成转化为一个可控文本生成任务,引入显式的多粒度控制信号来解决单调共情问题,在条件扩散模型中统一利用对话上文信息和控制信号,实现了对共情回复内容的指导,提高了共情回复的灵活度和准确度。

[0119] 本公开的回复内容确定装置72的一些实施例的示意图如图7所示。

[0120] 回复样本确定单元721能够根据待处理的对话上文和训练数据集获取匹配的回复样本。在一些实施例中,回复样本确定单元721能够执行上述步骤S41和图5对应的实施例中任一实施例中的操作。

[0121] 控制信号确定单元722能够根据回复样本生成控制信号。在一些实施例中,根据回复样本生成控制信号的操作可以如上文步骤S12中所示。

[0122] 回复确定单元723能够根据控制信号、回复样本和待处理的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息。条件扩散模型为通过上文中提到的任意一种训练方法训练生成。在一些实施例中,根据控制信号、回复样本和待处理的对话上文,通过带控制范围掩码策略的条件扩散模型获取针对对话上文的共情回复信息的操作可以如上文步骤S14中所示。

[0123] 这样的装置将共情回复生成转化为可控文本生成任务,并引入显式的多粒度控制信号来解决单调共情问题,在条件扩散模型中统一利用对话上文信息和控制信号;通过控制范围掩码策略实现了对共情回复内容的指导,并整合了一个情绪增强的匹配方法来为给定的上文产生多样的回复,提高了共情回复的灵活度和准确度。

[0124] 本公开数据处理装置的一个实施例的结构示意图如图8所示。数据处理装置包括存储器801和处理器802。其中:存储器801可以是磁盘、闪存或其它任何非易失性存储介质。存储器用于存储上文中模型训练方法或回复信息确定方法的对应实施例中的指令。处理器802耦接至存储器801,可以作为一个或多个集成电路来实施,例如微处理器或微控制器。该处理器802用于执行存储器中存储的指令,能够提高了共情回复的灵活度和准确度。

[0125] 在一个实施例中,还可以如图9所示,数据处理装置900包括存储器901和处理器902。处理器902通过BUS总线903耦合至存储器901。该数据处理装置900还可以通过存储接口904连接至外部存储装置905以便调用外部数据,还可以通过网络接口906连接至网络或者另外一台计算机系统(未标出)。此处不再进行详细介绍。

[0126] 在该实施例中,通过存储器存储数据指令,再通过处理器处理上述指令,能够提高了共情回复的灵活度和准确度。

[0127] 在另一个实施例中,一种计算机可读存储介质,其上存储有计算机程序指令,该指令被处理器执行时实现模型训练方法或回复信息确定方法对应实施例中的方法的步骤。本领域内的技术人员应明白,本公开的实施例可提供为方法、装置、或计算机程序产品。因此,本公开可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本公开可采用在一个或多个其中包含有计算机可用程序代码的计算机可用非瞬时性

存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0128] 本公开是参照根据本公开实施例的方法、设备(系统)和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0129] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0130] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0131] 至此,已经详细描述了本公开。为了避免遮蔽本公开的构思,没有描述本领域所公知的一些细节。本领域技术人员根据上面的描述,完全可以明白如何实施这里公开的技术方案。

[0132] 可能以许多方式来实现本公开的方法以及装置。例如,可通过软件、硬件、固件或者软件、硬件、固件的任何组合来实现本公开的方法以及装置。用于所述方法的步骤的上述顺序仅是为了进行说明,本公开的方法的步骤不限于以上具体描述的顺序,除非以其它方式特别说明。此外,在一些实施例中,还可将本公开实施为记录在记录介质中的程序,这些程序包括用于实现根据本公开的方法的机器可读指令。因而,本公开还覆盖存储用于执行根据本公开的方法的程序的记录介质。

[0133] 需要说明的是,本公开的说明书和权利要求书及附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换,以便这里描述的本公开的实施例能够以除了在这里图示或描述的那些以外的顺序实施。此外,术语“包括”和“具有”以及他们的任何变形,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、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0134] 最后应当说明的是:以上实施例仅用以说明本公开的技术方案而非对其限制;尽管参照较佳实施例对本公开进行了详细的说明,所属领域的普通技术人员应当理解:依然可以对本公开的具体实施方式进行修改或者对部分技术特征进行等同替换;而不脱离本公开技术方案的精神,其均应涵盖在本公开请求保护的技术方案范围当中。

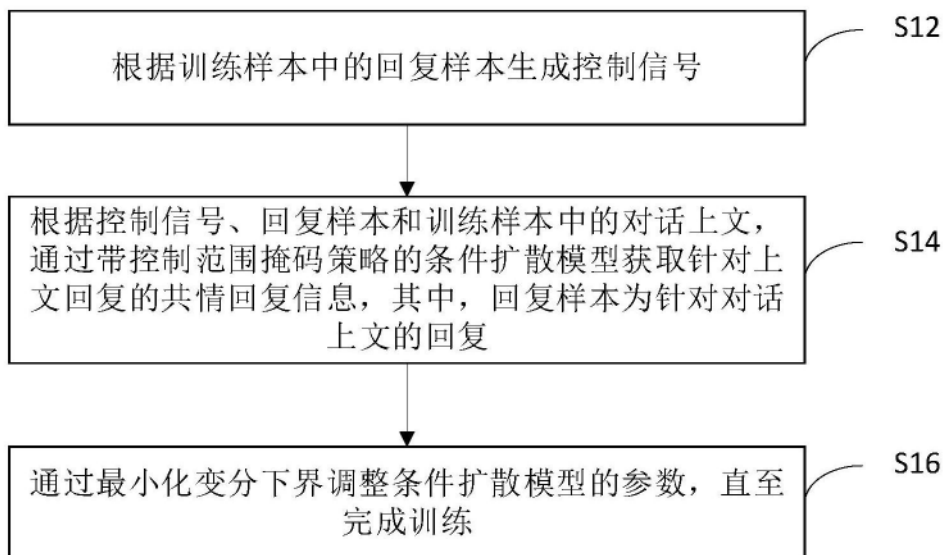


图1

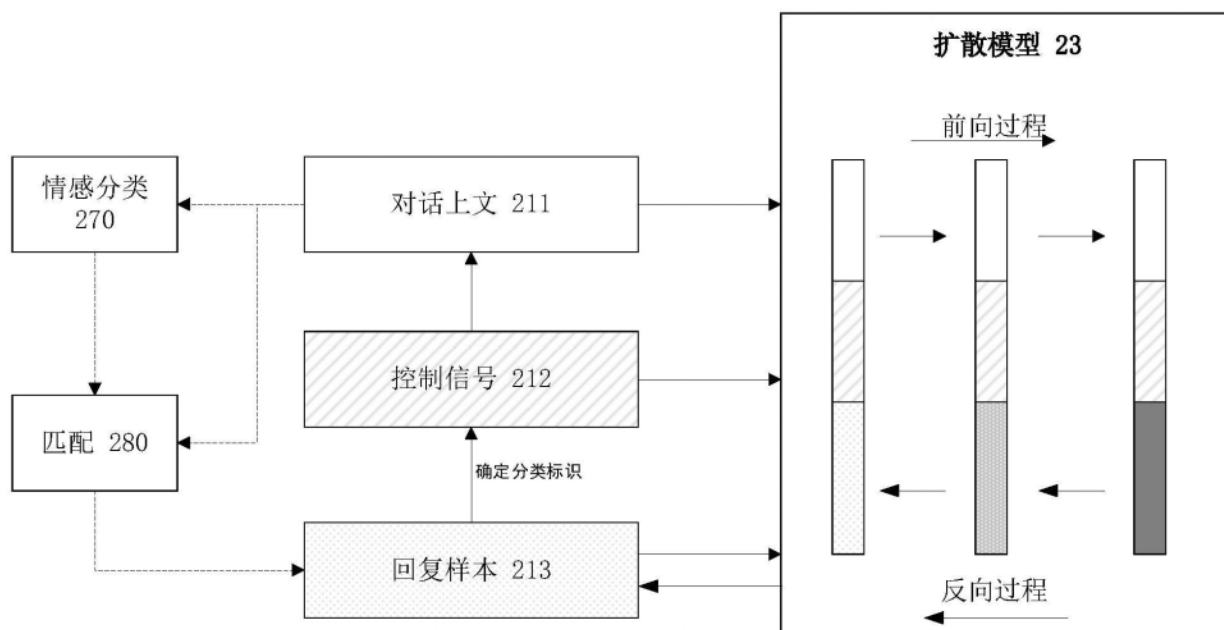


图2

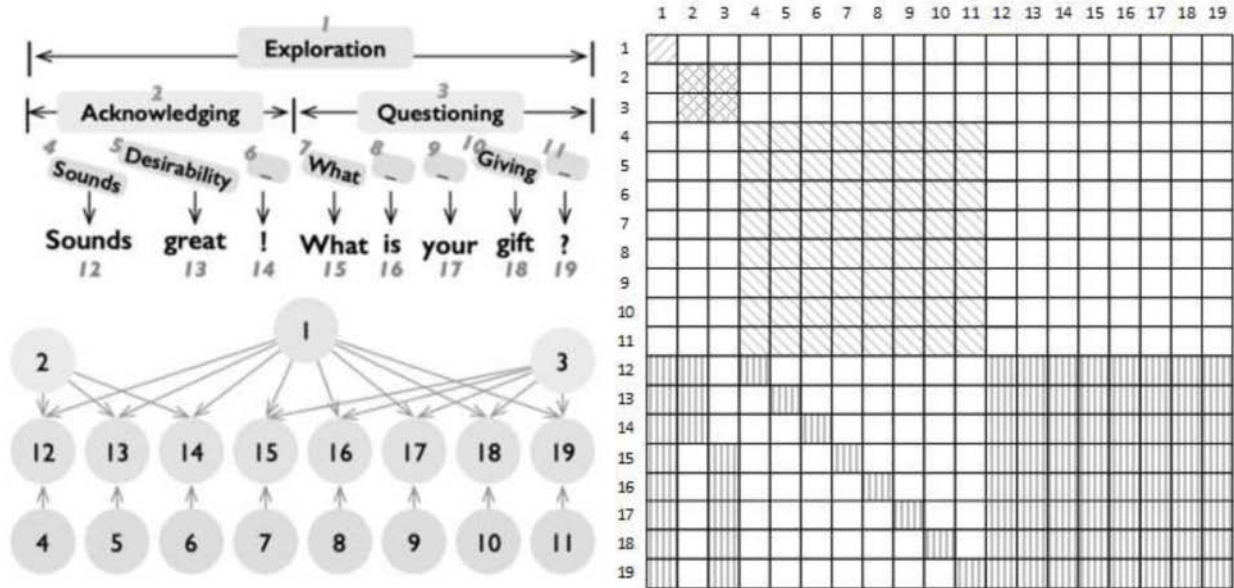


图3

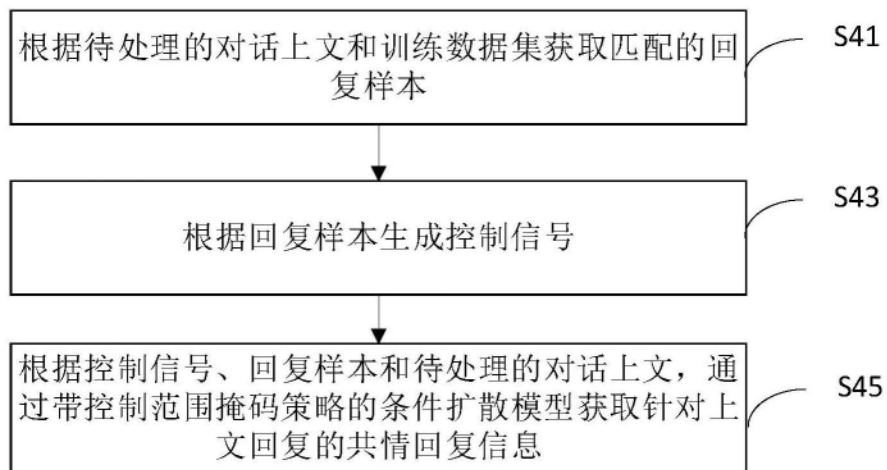


图4

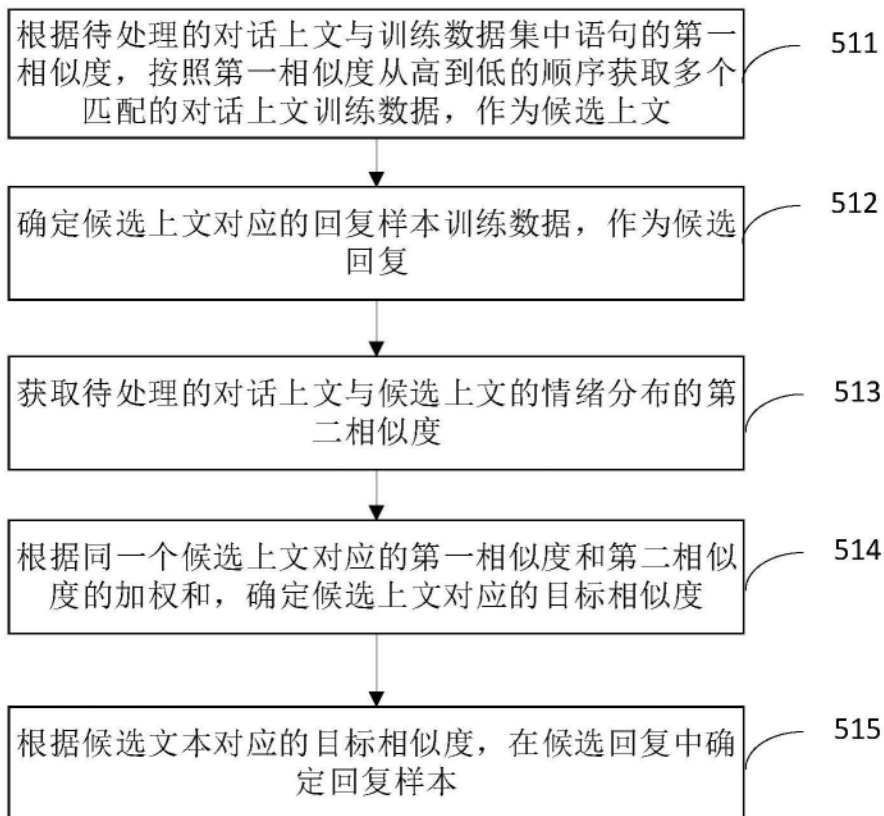


图5

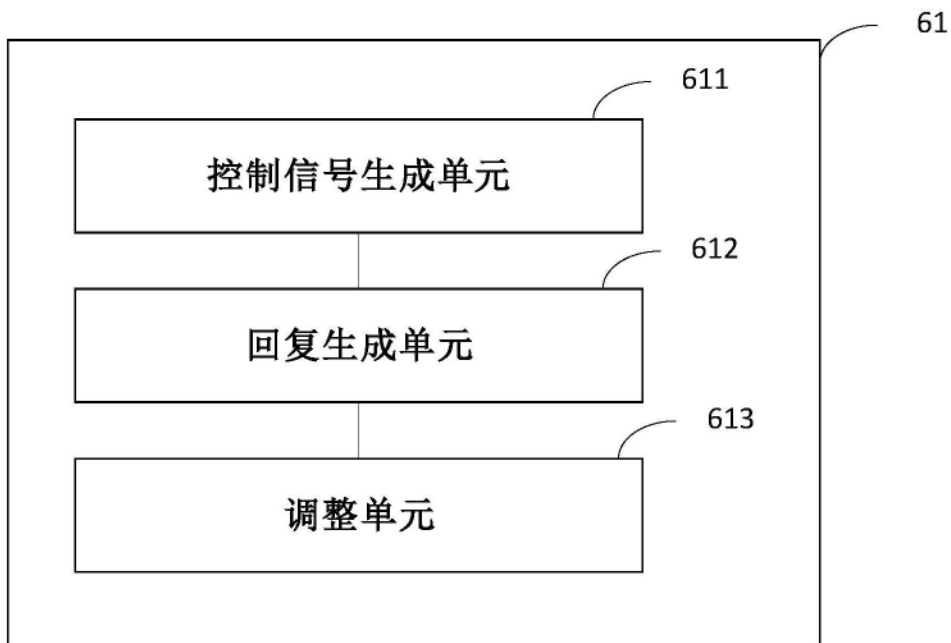


图6

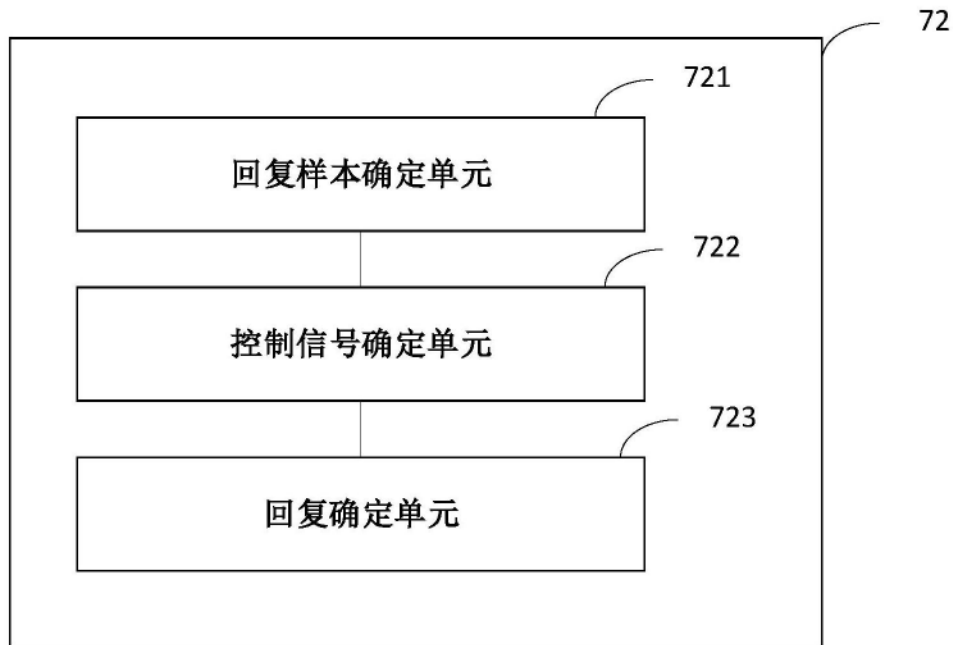


图7

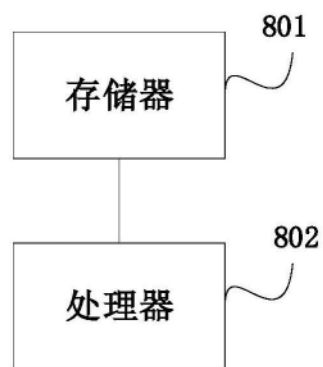


图8

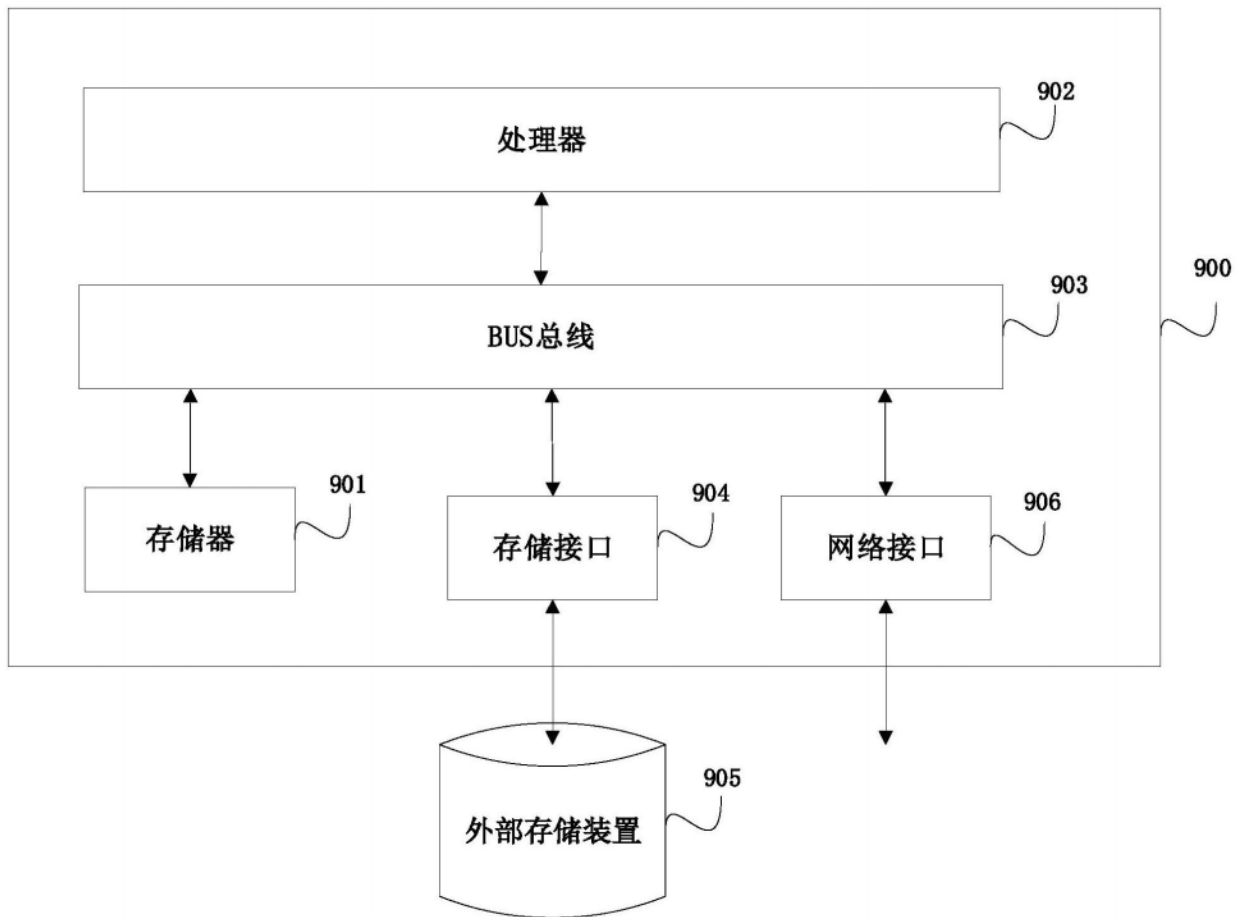


图9