



(12) 发明专利申请

(10) 申请公布号 CN 118069813 A

(43) 申请公布日 2024. 05. 24

(21) 申请号 202410371837.0

(22) 申请日 2024.03.29

(71) 申请人 广东电网有限责任公司

地址 510000 广东省广州市越秀区东风东路757号

申请人 广东电网有限责任公司客户服务中心

(72) 发明人 苏立伟 谭火超 陶飞达 刘振华
康峰 曹彦朝 杨秋勇 李庭磊
梁瑞莹 张艳 苏林峰 叶枝平
舒畅 李文虎 皮伟丰

(74) 专利代理机构 广州三环专利商标代理有限公司 44202

专利代理师 江嘉玲

(51) Int.Cl.

G06F 16/332 (2019.01)

G06F 40/284 (2020.01)

G06F 16/36 (2019.01)

G06F 18/22 (2023.01)

G06F 40/30 (2020.01)

G06F 40/289 (2020.01)

G06F 40/295 (2020.01)

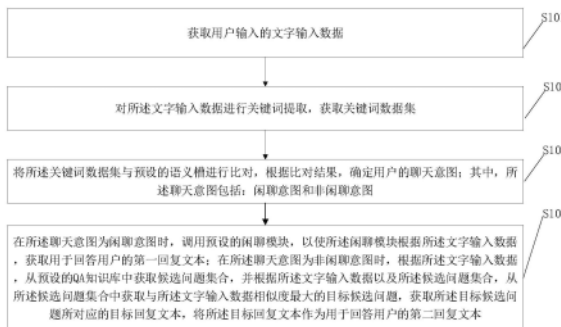
权利要求书2页 说明书10页 附图4页

(54) 发明名称

一种对话交互处理方法、装置、终端设备及存储介质

(57) 摘要

本发明公开了一种对话交互处理方法、装置、终端设备及存储介质,所述方法包括:对用户输入的文字输入数据进行关键词提取,获取关键词数据集,接着将关键词数据集与预设的语义槽进行比对,确定用户的聊天意图;若聊天意图为闲聊意图,则直接根据预设的闲聊模块输出用于闲聊的第一回复文本;若聊天意图为非闲聊意图,那么需要从预设的QA知识库中获取候选问题集合并计算相似度,获取相似度最大的目标候选问题,进而获取目标候选问题所对应的目标回复文本并作为用于回答用户的第二回复文本。通过实施本发明可以同时兼顾闲聊状态和非闲聊状态下的对话交互,不需要再预先通过手动调节或者选择对应的对话模式,提高了用户使用智能客服对话的便利性。



1. 一种对话交互处理方法,其特征在于,包括:

获取用户输入的文字输入数据;

对所述文字输入数据进行关键词提取,获取关键词数据集;

将所述关键词数据集与预设的语义槽进行比对,根据比对结果,确定用户的聊天意图;其中,所述聊天意图包括:闲聊意图和非闲聊意图;

在所述聊天意图为闲聊意图时,调用预设的闲聊模块,以使所述闲聊模块根据所述文字输入数据,获取用于回答用户的第一回复文本;在所述聊天意图为非闲聊意图时,根据所述文字输入数据以及关键词数据集,从预设的QA知识库中获取候选问题集合,并根据所述文字输入数据以及所述候选问题集合,从所述候选问题集合中获取与所述文字输入数据相似度最大的目标候选问题,获取所述目标候选问题所对应的目标回复文本,将所述目标回复文本作为用于回答用户的第二回复文本。

2. 如权利要求1所述的对话交互处理方法,其特征在于,所述语义槽包括:用于表征业务相关问题主体的故障实体槽位、用于表征问题处理请求关键词的处理请求槽位以及用于表征问题实体对应具体处理点的处理属性槽位。

3. 如权利要求2所述的对话交互处理方法,其特征在于,所述将所述关键词数据集与预设的语义槽进行比对,根据比对结果,确定用户的聊天意图,包括:

将所述关键词数据集中的各个关键词与预设的语义槽进行比对,若各所述关键词均不属于故障实体槽位、处理请求槽位以及处理属性槽位的关键词,或者,各所述关键词中仅有一个关键词属于处理请求槽位的关键词,并且其余关键词均不属于故障实体槽位和处理属性槽位的关键词,则确定用户的聊天意图为闲聊意图;否则,确定用户的聊天意图为非闲聊意图。

4. 如权利要求3所述的对话交互处理方法,其特征在于,所述非闲聊意图包括:强处理意图、处理意图、询问意图、轮询意图以及未知意图;

若所述关键词数据集中有一个属于故障实体槽位的关键词、一个属于处理请求槽位的关键词和一个属于处理属性槽位的关键词,则确定所述非闲聊意图为强处理意图;

若所述关键词数据集中有一个属于处理请求槽位的关键词和有一个属于处理属性槽位的关键词,则确定所述非闲聊意图为处理意图;

若所述关键词数据集中有一个属于故障实体槽位的关键词,或者,有一个属于故障实体槽位的关键词和有一个属于处理请求槽位的关键词,或者,有一个属于故障实体槽位的关键词和有一个属于处理属性槽位的关键词,则确定所述非闲聊意图为询问意图;

若所述关键词数据集中仅有一个关键词属于处理属性槽位的关键词,则确定所述非闲聊意图为轮询意图;

若所述关键词数据集中至少有两个关键词属于故障实体槽位的关键词,和/或,至少有两个关键词属于处理属性槽位的关键词,则确定所述非闲聊意图为未知意图。

5. 如权利要求4所述的对话交互处理方法,其特征在于,所述根据所述文字输入数据以及所述候选问题集合,从所述候选问题集合中获取与所述文字输入数据相似度最大的目标候选问题,包括:

计算所述文字输入数据与所述候选问题集合中各候选问题的语义相似度,获取第一相似度;

计算所述文字输入数据的上下文与所述候选问题集合中各候选问题上下文的语义相似度,获取第二相似度;

计算所述文字输入数据的摘要信息与所述候选问题集合中各候选问题摘要信息的相似度,获取第三相似度;

根据所述第一相似度、第二相似度以及第三相似度,计算所述文字输入数据与所述候选问题集合中各候选问题的目标相似度,将目标相似度最大的候选问题作为目标候选问题。

6. 如权利要求5所述的对话交互处理方法,其特征在于,通过下式计算所述文字输入数据与所述候选问题集合中各候选问题的目标相似度:

$$Score(s_i^{D_i}) = \lambda_1 Sim(x_k^u, s_i^{D_i}) + \lambda_2 Sim(C_k, D_i) + \lambda_3 Sim(S_c, S_{D_i})$$

其中, $Score(s_i^{D_i})$ 为所计算的候选问题与当前输入的相似度; $Sim(x_k^u, s_i^{D_i})$ 表示候选问题与当前输入的第一相似度; $Sim(C_k, D_i)$ 表示候选问题与当前输入的第二相似度; $Sim(S_c, S_{D_i})$ 表示候选问题与当前输入的第三相似度; λ_1 、 λ_2 和 λ_3 分别为预设的加权系数。

7. 如权利要求6所述的对话交互处理方法,其特征在于,所述对所述文字输入数据进行关键词提取,获取关键词数据集,包括:

识别所述文字输入数据中词性为名词或者动词的关键词并进行提取,根据提取的关键词建立关键词数据集。

8. 一种对话交互处理装置,其特征在于,包括:数据获取模块、关键词提取模块、聊天意图模块以及文本获取模块;

所述数据获取模块,用于获取用户输入的文字输入数据;

所述关键词提取模块,用于对所述文字输入数据进行关键词提取,获取关键词数据集;

所述聊天意图模块,用于将所述关键词数据集与预设的语义槽进行比对,根据比对结果,确定用户的聊天意图;其中,所述聊天意图包括:闲聊意图和非闲聊意图;

所述文本获取模块,用于在所述聊天意图为闲聊意图时,调用预设的闲聊模块,以使所述闲聊模块根据所述文字输入数据,获取用于回答用户的第一回复文本;在所述聊天意图为非闲聊意图时,根据所述文字输入数据以及关键词数据集,从预设的QA知识库中获取候选问题集合,并根据所述文字输入数据以及所述候选问题集合,从所述候选问题集合中获取与所述文字输入数据相似度最大的目标候选问题,获取所述目标候选问题所对应的目标回复文本,将所述目标回复文本作为用于回答用户的第二回复文本。

9. 一种终端设备,其特征在于,包括处理器、存储器以及存储在所述存储器中且被配置为由所述处理器执行的计算机程序,所述处理器执行所述计算机程序时实现如权利要求1至7中任意一项所述的对话交互处理方法。

10. 一种存储介质,其特征在于,所述存储介质包括存储的计算机程序,其中,在所述计算机程序运行时控制所述存储介质所在设备执行如权利要求1至7中任意一项所述的对话交互处理方法。

一种对话交互处理方法、装置、终端设备及存储介质

技术领域

[0001] 本发明涉及智能客服对话领域,尤其涉及一种对话交互处理方法、装置、终端设备及存储介质。

背景技术

[0002] 如今,在面对大群体客户需要沟通时,例如零售行业门店数量较大,动辄上百上千,人力客服资源往往不充足,又或者希望减少客服人力开销的情况下,许多企业都会采用先让客户与自动化的服务系统进行沟通,当自动化系统未能解决问题后,再交由人工服务的方案,以此来提高业务整体的效率;由此,也演变出了各种类型的智能客服,按功能侧重可具体划分为,任务型、问答型和闲聊型。任务型能够通过单轮或多轮对话的方式获取有用信息帮助服务对象完成某些任务,问答型则仅用于回复问题,而闲聊型则能够与服务对象进行互动闲聊。

[0003] 然而目前三种类型智能客服在各自侧重领域可以取得很好的效果,但兼顾多种聊天模式的场景下则有所欠缺,用户需要根据不同的对话需求进行调节或者选择,效率低下。

发明内容

[0004] 本发明提供了一种对话交互处理方法、装置、终端设备及存储介质,可以同时兼顾闲聊状态和非闲聊状态下的对话交互,不需要再预先通过手动调节或者选择对应的对话模式,提高了用户使用智能客服对话的便利性。。

[0005] 本发明提供了一种对话交互处理方法,包括:获取用户输入的文字输入数据;

[0006] 对所述文字输入数据进行关键词提取,获取关键词数据集;

[0007] 将所述关键词数据集与预设的语义槽进行比对,根据比对结果,确定用户的聊天意图;其中,所述聊天意图包括:闲聊意图和非闲聊意图;

[0008] 在所述聊天意图为闲聊意图时,调用预设的闲聊模块,以使所述闲聊模块根据所述文字输入数据,获取用于回答用户的第一回复文本;在所述聊天意图为非闲聊意图时,根据所述文字输入数据以及关键词数据集,从预设的QA知识库中获取候选问题集合,并根据所述文字输入数据以及所述候选问题集合,从所述候选问题集合中获取与所述文字输入数据相似度最大的目标候选问题,获取所述目标候选问题所对应的目标回复文本,将所述目标回复文本作为用于回答用户的第二回复文本。

[0009] 进一步地,所述语义槽包括:用于表征业务相关问题主体的故障实体槽位、用于表征问题处理请求关键词的处理请求槽位以及用于表征问题实体对应具体处理点的处理属性槽位。

[0010] 进一步地,所述将所述关键词数据集与预设的语义槽进行比对,根据比对结果,确定用户的聊天意图,包括:

[0011] 将所述关键词数据集中的各个关键词与预设的语义槽进行比对,若各所述关键词均不属于故障实体槽位、处理请求槽位以及处理属性槽位的关键词,或者,各所述关键词中

仅有一个关键词属于处理请求槽位的关键词,并且其余关键词均不属于故障实体槽位和处理属性槽位的关键词,则确定用户的聊天意图为闲聊意图;否则,确定用户的聊天意图为非闲聊意图。

[0012] 进一步地,所述非闲聊意图包括:强处理意图、处理意图、询问意图、轮询意图以及未知意图;

[0013] 若所述关键词数据集中有一个属于故障实体槽位的关键词、一个属于处理请求槽位的关键词和一个属于处理属性槽位的关键词,则确定所述非闲聊意图为强处理意图;

[0014] 若所述关键词数据集中有一个属于处理请求槽位的关键词和有一个属于处理属性槽位的关键词,则确定所述非闲聊意图为处理意图;

[0015] 若所述关键词数据集中有一个属于故障实体槽位的关键词,或者,有一个属于故障实体槽位的关键词和有一个属于处理请求槽位的关键词,或者,有一个属于故障实体槽位的关键词和有一个属于处理属性槽位的关键词,则确定所述非闲聊意图为询问意图;

[0016] 若所述关键词数据集中仅有一个关键词属于处理属性槽位的关键词,则确定所述非闲聊意图为轮询意图;

[0017] 若所述关键词数据集中至少有两个关键词属于故障实体槽位的关键词,和/或,至少有两个关键词属于处理属性槽位的关键词,则确定所述非闲聊意图为未知意图。

[0018] 进一步地,所述根据所述文字输入数据以及所述候选问题集合,从所述候选问题集合中获取与所述文字输入数据相似度最大的目标候选问题,包括:

[0019] 计算所述文字输入数据与所述候选问题集合中各候选问题的语义相似度,获取第一相似度;

[0020] 计算所述文字输入数据的上下文与所述候选问题集合中各候选问题上下文的语义相似度,获取第二相似度;

[0021] 计算所述文字输入数据的摘要信息与所述候选问题集合中各候选问题摘要信息的相似度,获取第三相似度;

[0022] 根据所述第一相似度、第二相似度以及第三相似度,计算所述文字输入数据与所述候选问题集合中各候选问题的目标相似度,将目标相似度最大的候选问题作为目标候选问题。

[0023] 进一步地,通过下式计算所述文字输入数据与所述候选问题集合中各候选问题的目标相似度:

$$[0024] \quad Score(s_i^{D_i}) = \lambda_1 Sim(x_k^u, s_i^{D_i}) + \lambda_2 Sim(C_k, D_i) + \lambda_3 Sim(S_c, S_{D_i})$$

[0025] 其中, $Score(s_i^{D_i})$ 为所计算的候选问题与当前输入的相似度; $Sim(x_k^u, s_i^{D_i})$ 表示候选问题与当前输入的第一相似度; $Sim(C_k, D_i)$ 表示候选问题与当前输入的第二相似度; $Sim(S_c, S_{D_i})$ 表示候选问题与当前输入的第三相似度; λ_1 、 λ_2 和 λ_3 分别为预设的加权系数。

[0026] 进一步地,所述对所述文字输入数据进行关键词提取,获取关键词数据集,包括:

[0027] 识别所述文字输入数据中词性为名词或者动词的关键词并进行提取,根据提取的关键词建立关键词数据集。

[0028] 在上述方法项实施例的基础上,本发明对应提供了装置项实施例;

[0029] 本发明提供了一种对话交互处理装置,包括:数据获取模块、关键词提取模块、聊

天意图模块以及文本获取模块；

[0030] 所述数据获取模块,用于获取用户输入的文字输入数据；

[0031] 所述关键词提取模块,用于对所述文字输入数据进行关键词提取,获取关键词数据集；

[0032] 所述聊天意图模块,用于将所述关键词数据集与预设的语义槽进行比对,根据比对结果,确定用户的聊天意图；其中,所述聊天意图包括:闲聊意图和非闲聊意图；

[0033] 所述文本获取模块,用于在所述聊天意图为闲聊意图时,调用预设的闲聊模块,以使所述闲聊模块根据所述文字输入数据,获取用于回答用户的第一回复文本；在所述聊天意图为非闲聊意图时,根据所述文字输入数据以及关键词数据集,从预设的QA知识库中获取候选问题集合,并根据所述文字输入数据以及所述候选问题集合,从所述候选问题集合中获取与所述文字输入数据相似度最大的目标候选问题,获取所述目标候选问题所对应的目标回复文本,将所述目标回复文本作为用于回答用户的第二回复文本。

[0034] 在上述方法项实施例的基础上,本发明对应提供了一终端设备项实施例；

[0035] 本发明提供了一种终端设备,包括处理器、存储器以及存储在所述存储器中且被配置为由所述处理器执行的计算机程序,所述处理器执行所述计算机程序时实现本发明任意一项所述的对话交互处理方法。

[0036] 在上述方法项实施例的基础上,本发明对应提供了一存储介质项实施例；

[0037] 本发明提供了一种存储介质,所述存储介质包括存储的计算机程序,其中,在所述计算机程序运行时控制所述存储介质所在设备执行本发明任意一项所述的对话交互处理方法。

[0038] 本发明的实施例,具有如下有益效果：

[0039] 本发明提供了一种对话交互处理方法、装置、终端设备及存储介质；所述方法,在获取用户输入的文字输入数据后,对所述文字输入数据进行关键词提取,获取关键词数据集；将获取的关键词数据集与预设的语义槽进行比对,进而能够确定用户的聊天意图。若用户的聊天意图为闲聊意图,则直接调用现有预设的闲聊模块,直接输出用于闲聊的第一回复文本；若用户的聊天意图为非闲聊意图,那么需要从预设的QA知识库中获取候选问题集合,计算与所述文字输入数据相似度最大的目标候选问题,进而获取目标候选问题所对应的目标回复文本,并将其作为用于回答用户的第二回复文本。通过实施本发明可以同时兼顾闲聊状态和非闲聊状态下的对话交互,不需要再预先通过手动调节或者选择对应的对话模式,提高了用户使用智能客服对话的便利性。

附图说明

[0040] 图1是本发明一实施例提供的一种对话交互处理方法流程示意图；

[0041] 图2是本发明一实施例提供的另一种对话交互处理方法流程示意图；

[0042] 图3是本发明一实施例提供的有限状态机建模的业务逻辑流程示意图；

[0043] 图4是本发明一实施例提供的多轮对话及关键词提取示意图；

[0044] 图5是本发明一实施例提供的一种对话交互处理装置的结构示意图。

具体实施方式

[0045] 下面将结合本发明中的附图,对本发明中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,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,都属于本发明保护的范围。

[0046] 如图1所示,一实施例提供一种对话交互处理方法,包括:

[0047] 步骤S101:获取用户输入的文字输入数据;

[0048] 步骤S102:对所述文字输入数据进行关键词提取,获取关键词数据集;

[0049] 步骤S103:将所述关键词数据集与预设的语义槽进行比对,根据比对结果,确定用户的聊天意图;其中,所述聊天意图包括:闲聊意图和非闲聊意图;

[0050] 步骤S104:在所述聊天意图为闲聊意图时,调用预设的闲聊模块,以使所述闲聊模块根据所述文字输入数据,获取用于回答用户的第一回复文本;在所述聊天意图为非闲聊意图时,根据所述文字输入数据以及关键词数据集,从预设的QA知识库中获取候选问题集合,并根据所述文字输入数据以及所述候选问题集合,从所述候选问题集合中获取与所述文字输入数据相似度最大的目标候选问题,获取所述目标候选问题所对应的目标回复文本,将所述目标回复文本作为用于回答用户的第二回复文本。

[0051] 对于步骤S101,在一个优选的实施例中,在与用户进行对话交互时,获取用户输入的文字输入数据。

[0052] 对于步骤S102,在一个优选的实施例中对所述文字输入数据进行关键词提取,获取关键词数据集,具体包括:

[0053] 在获取文字输入数据后,识别所述文字输入数据中词性为名词或者动词的关键词,然后将识别到的词性为名词或者动词的关键词提取出来,建立关键词数据集。其中,识别词性为名词或者动词的关键词的方法为现有技术中任意一种方法,本发明不做具体限定。

[0054] 对于步骤S103,在一个优选的实施例中,所述语义槽包括但不限于:用于表征业务相关问题主体的故障实体槽位、用于表征问题处理请求的处理请求槽位以及用于表征问题实体对应具体处理点的处理属性槽位。

[0055] 将所述关键词数据集与预设的语义槽进行比对,根据比对结果,确定用户的聊天意图,具体包括:

[0056] 将关键词数据集中的各个关键词分别与各个预设的语义槽进行比对,若各所述关键词均不属于故障实体槽位、处理请求槽位以及处理属性槽位的关键词,或者,各所述关键词中仅有一个关键词属于处理请求槽位的关键词,且其余关键词均不属于故障实体槽位和处理属性槽位的关键词,此时就能确定用户的聊天意图为闲聊意图。否则,确定用户的聊天意图为非闲聊意图。

[0057] 判断用户的聊天意图为闲聊意图,主要是基于关键词与预设的语义槽位的比对结果,如果所有的关键词都不属于预设的槽位,或者只有一个关键词属于处理请求槽位,且其余关键词都不属于故障实体和处理属性槽位,说明用户的聊天内容没有涉及到预设的业务相关问题,或者仅涉及处理请求而没有涉及具体的问题实体和属性;这种情况下,用户的聊天意图更倾向于闲聊,因此判断用户的聊天意图为闲聊意图。

[0058] 在一个优选的实施例中,所述非闲聊意图包括:强处理意图、处理意图、询问意图、轮询意图以及未知意图;

[0059] 若所述关键词数据集中有一个属于故障实体槽位的关键词、一个属于处理请求槽位的关键词和一个属于处理属性槽位的关键词,则确定所述非闲聊意图为强处理意图;

[0060] 若所述关键词数据集中有一个属于处理请求槽位的关键词和有一个属于处理属性槽位的关键词,则确定所述非闲聊意图为处理意图;

[0061] 若所述关键词数据集中有一个属于故障实体槽位的关键词,或者,有一个属于故障实体槽位的关键词和有一个属于处理请求槽位的关键词,或者,有一个属于故障实体槽位的关键词和有一个属于处理属性槽位的关键词,则确定所述非闲聊意图为询问意图;

[0062] 若所述关键词数据集中仅有一个关键词属于处理属性槽位的关键词,则确定所述非闲聊意图为轮询意图;

[0063] 若所述关键词数据集中至少有两个关键词属于故障实体槽位的关键词,和/或,至少有两个关键词属于处理属性槽位的关键词,则确定所述非闲聊意图为未知意图。

[0064] 对于步骤S104,在一个优选的实施例中,在确定聊天意图为闲聊意图时,那么直接可以调用现有预设的闲聊模块,使所述闲聊模块根据所述文字输入数据检索对应的文本进行回答。由于在闲聊意图下与用户进行对话的方法属于现有技术中的常规的手段,因此本申请不对所述闲聊模块做具体限定,属于现有技术中任意一种能够实现相关功能的闲聊模块。

[0065] 在确定聊天意图为非闲聊意图时,根据所述文字输入数据以及关键词数据集,对预设的QA知识库进行检索,获取与所述文字输入数据以及关键词数据集具有关联关系的候选问题集合,接着再根据文字输入数据以及所述候选问题集合,从所述候选问题集合中获取与文字输入数据相似度最大的目标候选问题;由于目标候选问题所对应的目标回复文本均为预先存储在QA知识库的数据,因此在确定目标候选问题后,可以直接获取对应的目标回复文本,然后将目标回复文本作为用于回答用户的第二回复文本。

[0066] 在一个优选的实施例中,根据所述文字输入数据以及所述候选问题集合,从所述候选问题集合中获取与所述文字输入数据相似度最大的目标候选问题,具体包括:

[0067] 首先计算所述文字输入数据与所述候选问题集合中各候选问题的语义相似度,获取第一相似度;

[0068] 接着计算所述文字输入数据的上下文与所述候选问题集合中各候选问题上下文的语义相似度,获取第二相似度;

[0069] 再计算所述文字输入数据的摘要信息与所述候选问题集合中各候选问题摘要信息的相似度,获取第三相似度;

[0070] 最后根据所述第一相似度、第二相似度以及第三相似度,计算所述文字输入数据与所述候选问题集合中各候选问题的目标相似度,将目标相似度最大的候选问题作为目标候选问题。

[0071] 可选地,计算第二相似度的方法为:通过循环神经网络获取当前文字输入数据的上下文的语境分布式表达,作为第一语境分布式表达;通过循环神经网络分别获取候选问题集合中各候选问题的上下文的语境分布式表达,作为第二语境分布式表达;最后分别计算第一语境分布式表达和第二语境分布式表达的相似度,作为对应的候选问题与当前输入

的第二相似度。

[0072] 计算第三相似度的方法为:对当前文字输入数据(也可以是多轮对话数据)进行摘要抽取,获得当前多轮对话的摘要信息;分别对候选多轮对话集合中各候选多轮对话进行摘要抽取,获得每个候选多轮对话的摘要信息,分别计算当前多轮对话的摘要信息与所述候选多轮对话集合中各候选多轮对话的摘要信息的相似度,作为对应的候选问题与当前输入的第三相似度。

[0073] 在一个优选的实施例中,通过通过加权求和的方式,计算所述文字输入数据与所述候选问题集合中各候选问题的目标相似度,如下式表示:

$$[0074] \quad Score(s_i^{D_i}) = \lambda_1 Sim(x_k^u, s_i^{D_i}) + \lambda_2 Sim(C_k, D_i) + \lambda_3 Sim(S_c, S_{D_i})$$

[0075] 其中, $Score(s_i^{D_i})$ 为所计算的候选问题与当前输入的相似度; $Sim(x_k^u, s_i^{D_i})$ 表示候选问题与当前输入的第一相似度; $Sim(C_k, D_i)$ 表示候选问题与当前输入的第二相似度; $Sim(S_c, S_{D_i})$ 表示候选问题与当前输入的第三相似度; λ_1 、 λ_2 和 λ_3 分别为预设的加权系数。

[0076] 在一个可选的实施例中,如图2所示,对话交互处理方法包括以下步骤:

[0077] S1:接收用户输入的文字输入;

[0078] S2:识别用户意图并解析槽位;

[0079] S3:按照FSM业务逻辑用对话轨迹追踪回溯和槽位约束执行对话逻辑处理,了解用户是否有聊天意图;

[0080] S4:若是聊天意图为闲聊意图,调用预设的闲聊模块根据文字输入进行查询处理,获取第一回复文本;

[0081] S5:若聊天意图为非闲聊意图,利用QA知识库推理的多轮多话回复选择方法对用户输入的文字进行查询处理,获取第二回复文本。

[0082] 在一个可选的实施例中,对于步骤S2,识别用户意图并解析槽位,判断所述用户的聊天意图,具体包括:

[0083] S21:创建语义槽,所述语义槽包括故障实体A、处理请求B以及处理属性C三个槽位。具体的,故障实体A指业务相关的问题主体,如“我的电脑屏幕坏了,需要进行修理怎么办?”,其中“电脑”即故障实体。处理请求B指能够代表问题处理请求的关键词,如“修理”、“怎么办”、“?”。而处理属性C则代表问题实体对应的具体处理点,如“屏幕”就是“电脑”处理属性。

[0084] S22:对所述文字输入进行语义提取,若从文字输入中不能提取到属于所述语义槽中任何一个槽位的关键词,或者,从文字输入中只提取到一个属于处理请求的关键词,则判定所述文字输入的意图为闲聊意图。具体来说,若文字输入的语义中仅存在属于处理请求的关键词,或者无可用语义的关键词,即可判定该文字输入为闲聊意图。否则,判定所述文字输入的意图为非闲聊意图。

[0085] 本实施例中,所述非闲聊意图包括强处理意图、处理意图、询问意图、轮询意图以及未知意图。其中,若从所述文字输入中提取到一个属于故障实体的关键词、一个属于处理请求的关键词和一个属于处理属性的关键词,判定所述文字输入的意图为强处理意图。

[0086] 若从所述文字输入中提取到一个属于处理请求的关键词和一个属于处理属性的关键词,判定所述文字输入的意图为处理意图。

[0087] 若从所述文字输入中提取到一个属于故障实体的关键词,或者,提取到一个属于故障实体的关键词和一个属于处理请求的关键词,或者,提取到一个属于故障实体的关键词和一个属于处理属性的关键词,判定所述文字输入的意图为询问意图。

[0088] 若从所述文字输入中只提取到一个属于处理属性的关键词,判定所述文字输入的意图为轮询意图。

[0089] 若从所述文字输入中提取到多个属于故障实体的关键词,和/或,多个属于处理属性的关键词,判定所述文字输入的意图为未知意图。

[0090] 在一个可选的实施例中,对于步骤S3,如图3所示:按照FSM业务逻辑用对话轨迹追踪回溯和槽位约束执行对话逻辑处理,了解用户是否有聊天意图,具体包括:

[0091] S31:利用识别的用户意图和/或解析槽位,更新状态追踪的概率分布;

[0092] S32:FSM决定对话状态是否转移,并更新对话轨迹和即时约束;

[0093] S33:检查当前对话是否存在槽位否定或约束修改;

[0094] S34:若是,根据回溯列表定位回溯位置,FSM决定转移方向,并根据转移后的状态和FSM的节点位置执行业务逻辑,继续步骤S5;

[0095] S35:若不是,根据当前的状态和FSM的节点位置执行业务逻辑,继续步骤S4。其中,可以首先在步骤S31利用步骤S22中识别的用户意图和/或解析的槽位,更新状态分布。然后在步骤S32中,有限状态机根据当前状态分布和即时约束列表,决定对话状态是否转移,并更新对话轨迹和即时约束。在步骤S33中,检查当前对话是否存在槽位否定或约束修改。如果存在槽位否定或约束修改,则在步骤S34中根据回溯列表定位回溯位置,有限状态机根据更新后的状态决定转移方向,并根据转移后的状态和有限状态机的节点位置执行有限状态机建模的业务逻辑;如果既不存在槽位否定,也不存在约束修改,则在步骤S35中根据当前的状态和有限状态机的节点位置执行有限状态机建模的业务逻辑。FSM表示有限状态机。即首先,通过识别用户意图和槽位信息来更新状态分布,然后根据当前状态分布和即时约束列表来决定对话状态是否转移,并更新对话轨迹和即时约束。接着,检查当前对话是否存在槽位否定或约束修改,如果存在则根据回溯列表定位回溯位置,并根据更新后的状态决定转移方向;如果不存在,则根据当前的状态和有限状态机的节点位置执行有限状态机建模的业务逻辑,这个处理流程用于实现对话系统的动态调整和流程控制。

[0096] 上述FSM节点位置:有限状态机(FSM)是一种计算模型,用于描述对象的状态和状态的转换,在对话系统中,FSM节点位置可以理解为当前对话所处的状态或阶段。约束修改:约束是指对某些行为或条件的限制或规定,在对话系统中,约束可能涉及到对话的逻辑、流程、规则等方面的限制,如果存在槽位否定或约束修改,这可能意味着对话的状态或流程需要调整或回溯。定位回溯位置:回溯是一种算法,属于现有技术,用于在搜索过程中回到先前状态并尝试其他路径,在对话系统中,回溯位置可以理解为需要返回的对话状态或步骤。有限状态机:有限状态机是一种计算模型,用于描述系统或对象的状态以及状态之间的转换,在对话系统中,有限状态机可以用来建模对话的状态和流程。

[0097] 此外,使用有限状态机通过设置以下基础状态节点进行业务逻辑建模:用户节点(UserNode):该节点负责接收解析用户询问(query)意图,只有用户询问与该节点负责的解析的意图一致,才会激活该节点,同时该节点兼有对话槽位收集功能,在该节点可以指定收集多个节点,支持槽位的澄清、修改,用户可以按照任意顺序告知槽位信息。机器人节点

(BotNode):该节点负责将对话管理器(DM)的执行结果告知用户,同时兼有对话逻辑引导的功能,通过发出预置的引导语,引导用户按照期望的逻辑进行对话。功能节点(FunctionNode,或FUNCNode):该节点可以执行预置的功能性脚本或调用第三方应用程序接口(API),使用对话过程中收集到的信息执行一些与对话无关的功能。切换节点(SwitchNode):该节点带有分支转发功能,一般用于根据对话中收集到的槽位、全局变量、函数执行结果甚至意图,进行对话逻辑分支的切换。槽位节点(SlotNode):该节点通常用于机器端主导的槽位收集,在该节点可以指定收集多个节点,支持槽位的澄清、修改,用户可以按照任意顺序告知槽位信息。主题节点(SubNode):该节点封装了调用指定主题的功能,它负责管理该主题的触发和结果返回。它实现了主题之间有依赖的嵌套。它负责对话逻辑的复用,可以将通用的对话逻辑封装成一个独立的主题,任何使用该对话逻辑的主题都可以通过主题节点复用该对话逻辑。

[0098] 在一个可选的实施例中,如图4所示,步骤S5包括以下步骤:

[0099] 对当前多轮对话进行关键词提取,获取第一关键词组合;

[0100] 以第一关键词组合、当前输入作为问题,从QA知识库中检索出候选问题集合,并获取各候选问题相应的上下文,构建候选多轮对话集合;QA知识库基于预先获取的多轮对话数据集进行构建;

[0101] 构建候选多轮对话集合,具体包括:根据第一关键词组合对QA知识库进行检索,获得包含第一关键词组合中关键词的多轮对话,作为第一多轮对话集合;根据当前输入对QA知识库进行检索,得到与当前输入最相似的N个候选句子及包含这个句子的多轮对话,作为第二多轮对话集合;对第一多轮对话集合和第二多轮对话集合取交集,得到候选问题集合和候选多轮对话集。

[0102] QA知识库的构建方法包括:对预先获取的多轮对话数据集进行分词、关键词抽取;使用关键词对多轮对话数据集建立索引,存储关键词及包含关键词的多轮对话的ID;对每一个多轮对话,将其中的每一句话建立倒排索引,存储包含关键词的句子、这个句子所属多轮对话的ID以及这个句子在多轮对话中的位置。

[0103] 在一个可选的实施例中,自然语言对话交互处理方法工作原理包括:先根据用户的文字输入判断所述用户的聊天意图,根据聊天意图分配相应查询模块进行查询处理,获得相应的回复文本,尽可能满足用户需求,高效地完成问答任务。本发明的方法不需要大量数据结合深度学习进行关系抽取、命名实体识别、意图识别以及自然语言生成的训练就能够使用,而且面向的是一些通用的问答。本方法面向小样本数据,其核心逻辑主要靠制定规则实现,而并非通过深度学习的方式实现。实施本发明方法的交互机器人适用于零售行业,能够灵活扩展,只需要少量数据即可灵活高效回复,并能够自由添加特定场景。基于QA知识库的回复选择方法能够有效利用知识库中的上下文信息以及当前多轮对话的上下文信息进行推以很大程度上提高回复语句的相关性和合理性,提高了正确选择回复语句的鲁棒性,从而大幅提升用户的对话体验。

[0104] 在上述方法项实施例的基础上,本发明对应提供了装置项实施例。

[0105] 如图5所示,本发明一实施例提供了一种对话交互处理装置,包括:数据获取模块、关键词提取模块、聊天意图模块以及文本获取模块;

[0106] 所述数据获取模块,用于获取用户输入的文字输入数据;

[0107] 所述关键词提取模块,用于对所述文字输入数据进行关键词提取,获取关键词数据集;

[0108] 所述聊天意图模块,用于将所述关键词数据集与预设的语义槽进行比对,根据比对结果,确定用户的聊天意图;其中,所述聊天意图包括:闲聊意图和非闲聊意图;

[0109] 所述文本获取模块,用于在所述聊天意图为闲聊意图时,调用预设的闲聊模块,以使所述闲聊模块根据所述文字输入数据,获取用于回答用户的第一回复文本;在所述聊天意图为非闲聊意图时,根据所述文字输入数据以及关键词数据集,从预设的QA知识库中获取候选问题集合,并根据所述文字输入数据以及所述候选问题集合,从所述候选问题集合中获取与所述文字输入数据相似度最大的目标候选问题,获取所述目标候选问题所对应的目标回复文本,将所述目标回复文本作为用于回答用户的第二回复文本。

[0110] 需说明的是,以上所描述的装置实施例仅仅是示意性的,其中所述作为分离部件说明的模块可以是或者也可以不是物理上分开的,作为模块显示的部件可以是或者也可以不是物理模块,即可以位于一个地方,或者也可以分布到多个网络模块上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。另外,本发明提供的装置实施例附图中,模块之间的连接关系表示它们之间具有通信连接,具体可以实现为一条或多条通信总线或信号线。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0111] 所述领域的技术人员可以清楚地了解到,为的方便和简洁,上述描述的装置的具体工作过程,可参考前述方法实施例中对应的过程,在此不再赘述。

[0112] 在上述方法项实施例的基础上,本发明对应提供了终端设备项实施例。

[0113] 本发明另一实施例提供了一种终端设备,包括处理器、存储器以及存储在所述存储器中且被配置为由所述处理器执行的计算机程序;所述处理器执行所述计算机程序时,实现本发明任意一实施例的对话交互处理方法。

[0114] 示例性的,在这一实施例中所述计算机程序可以被分割成一个或多个模块,所述一个或者多个模块被存储在所述存储器中,并由所述处理器执行,以完成本发明。所述一个或多个模块元可以是能够完成特定功能的一系列计算机程序指令段,该指令段用于描述所述计算机程序在所述设备中的执行过程;

[0115] 所述终端设备可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述设备可包括,但不仅限于,处理器、存储器;

[0116] 所称处理器可以是中央处理模块(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等,所述处理器是所述设备的控制中心,利用各种接口和线路连接整个设备的各个部分;

[0117] 所述存储器可用于存储所述计算机程序和/或模块,所述处理器通过运行或执行存储在所述存储器内的计算机程序和/或模块,以及调用存储在存储器内的数据,实现所述设备的各种功能。所述存储器可主要包括存储程序区和存储数据区,其中,存储程序区可存

储操作系统、至少一个功能所需的应用程序等;此外,存储器可以包括高速随机存取存储器,还可以包括非易失性存储器,例如硬盘、内存、插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)、至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

[0118] 在上述方法项实施例的基础上,本发明对应提供了存储介质项实施例。

[0119] 本发明另一实施例提供了一种存储介质,所述存储介质包括存储的计算机程序,其中,在所述计算机程序运行时控制所述存储介质所在的设备执行本发明任意一实施例的对话交互处理方法。

[0120] 在这一实施例中,上述存储介质为计算机可读存储介质,所述计算机程序包括计算机程序代码,所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括:能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、电载波信号、电信信号以及软件分发介质等。

[0121] 通过实施本发明上述实施例具有如下效果:

[0122] 1.本发明的方法不需要大量数据结合深度学习进行关系抽取、命名实体识别、意图识别以及自然语言生成的训练就能够使用,而且面向的是一些通用的问答。本方法面向小样本数据,其核心逻辑主要靠制定规则实现,而并非通过深度学习的方式实现。

[0123] 2.实施本发明方法的交互机器人适用于零售行业,能够灵活扩展,只需要少量数据即可灵活高效回复,并能够自由添加特定场景。

[0124] 3.基于QA知识库的回复选择方法能够有效利用知识库中的上下文信息以及当前多轮对话的上下文信息进行推理,可以很大程度上提高回复语句的相关性和合理性,提高了正确选择回复语句的鲁棒性,从而大幅提升用户的对话体验。

[0125] 4.可以同时兼顾闲聊状态和非闲聊状态下的对话交互,不需要再预先通过手动调节或者选择对应的对话模式,提高了用户使用智能客服对话的便利性。

[0126] 上所述是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也视为本发明的保护范围。

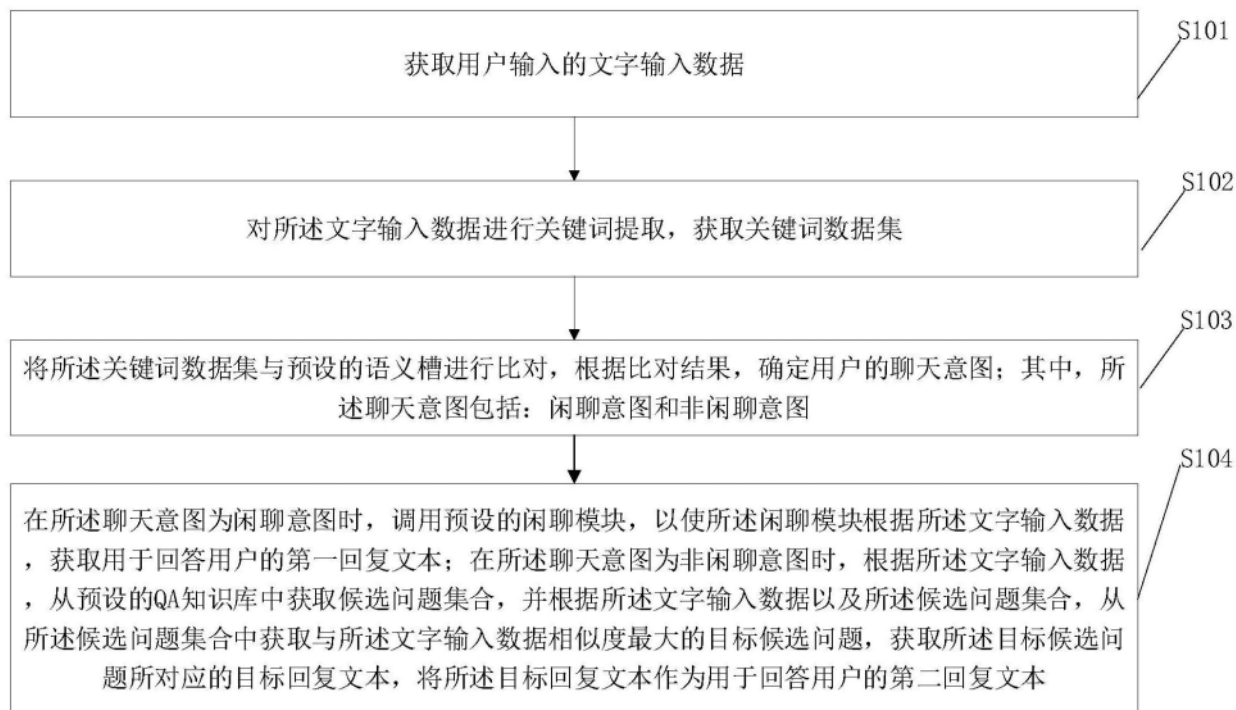


图1

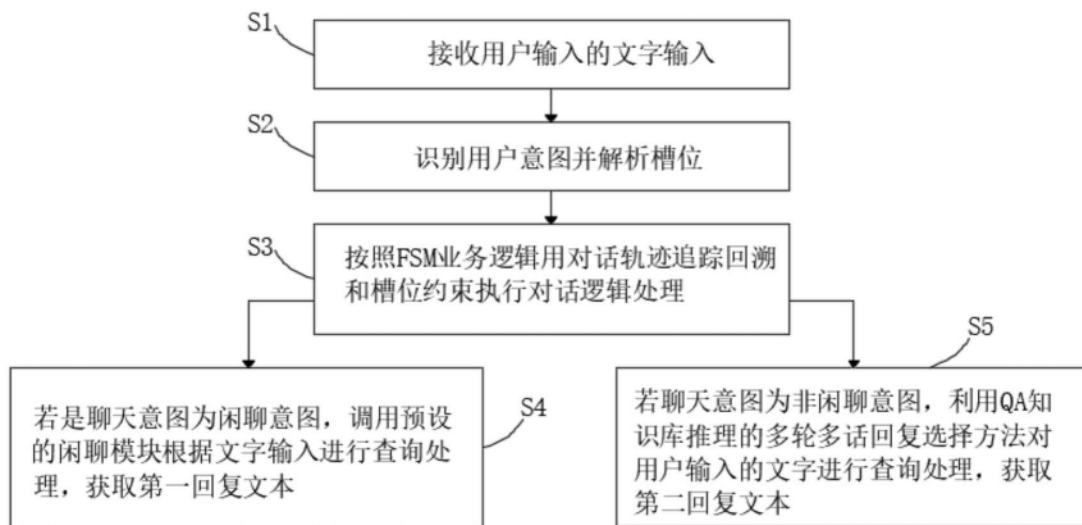


图2

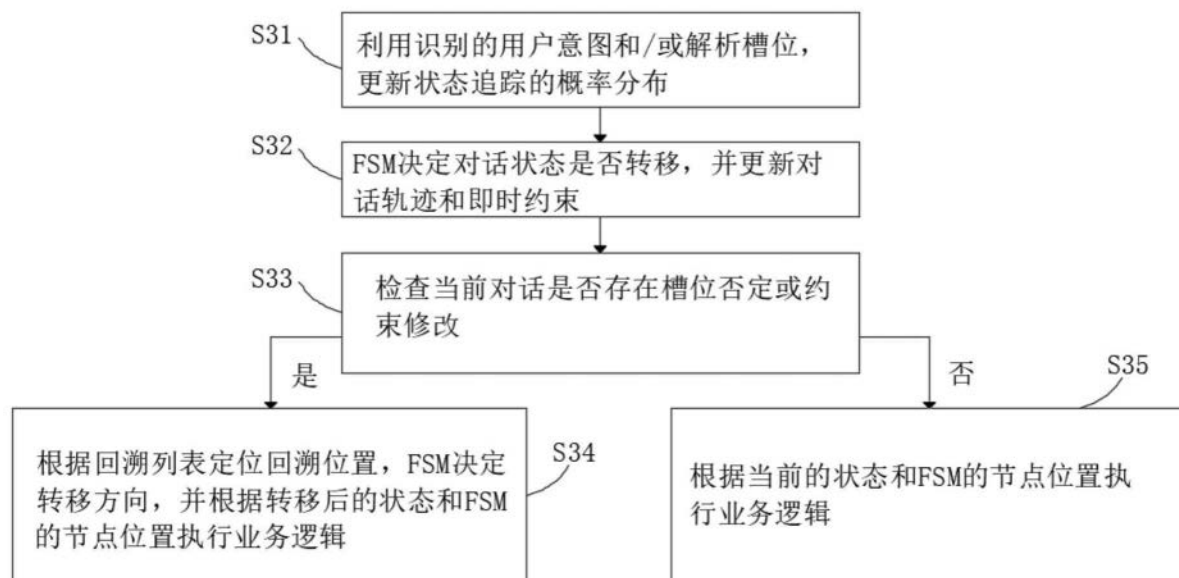


图3

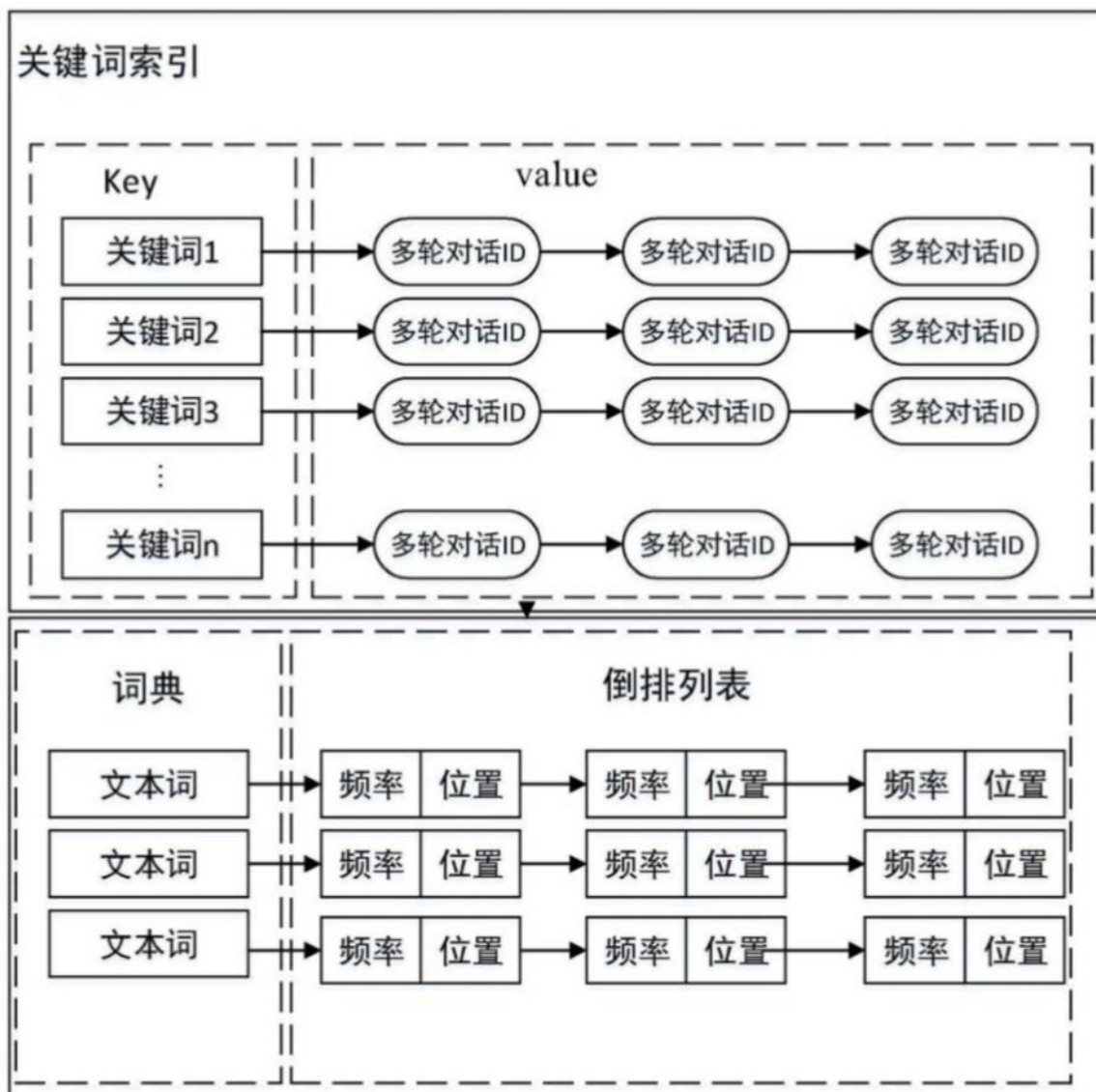


图4



图5