



(12) 发明专利申请

(10) 申请公布号 CN 116882372 A

(43) 申请公布日 2023. 10. 13

(21) 申请号 202310876300.5

G06F 16/31 (2019.01)

(22) 申请日 2023.07.17

G06F 16/33 (2019.01)

(71) 申请人 百度时代网络技术(北京)有限公司

地址 100080 北京市海淀区东北旺西路8号
中关村软件园17号楼二层A2

(72) 发明人 黄雪原 王永佳 聂卫国 张玉东
徐宪君

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021

专利代理师 杨静

(51) Int. Cl.

G06F 40/166 (2020.01)

G06F 40/284 (2020.01)

G06F 40/30 (2020.01)

G06F 16/332 (2019.01)

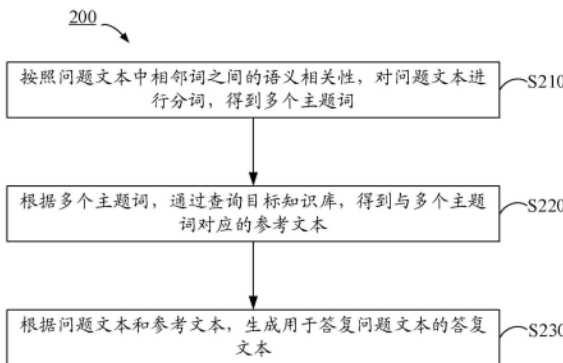
权利要求书4页 说明书13页 附图7页

(54) 发明名称

文本生成方法、装置、电子设备以及存储介
质

(57) 摘要

本公开提供了文本生成方法、装置、电子设备以及存储介质,涉及人工智能技术领域,尤其涉及自然语言处理、深度学习、大语言模型技术领域。该文本生成方法的具体实现方案为:按照问题文本中相邻词之间的语义相关性,对问题文本进行分词,得到多个主题词;根据多个主题词,通过查询目标知识库,得到与多个主题词对应的参考文本,其中,目标知识库是根据目标对象已公开发表的历史观点内容构建的;以及根据问题文本和参考文本,生成用于答复问题文本的答复文本。



1. 一种文本生成方法,包括:

按照问题文本中相邻词之间的语义相关性,对所述问题文本进行分词,得到多个主题词;

根据所述多个主题词,通过查询目标知识库,得到与所述多个主题词对应的参考文本,其中,所述目标知识库是根据目标对象已公开发表的历史观点内容构建的;以及

根据所述问题文本和所述参考文本,生成用于答复所述问题文本的答复文本。

2. 根据权利要求1所述的方法,其中,所述目标知识库包括第一索引知识库和第二索引知识库;所述根据所述多个主题词,通过查询目标知识库,得到与所述多个主题词对应的参考文本,包括:

根据所述多个主题词,通过查询第一索引知识库,得到第一候选文本,其中,所述第一索引知识库包括候选文本的摘要与候选文本的索引关系;

根据所述多个主题词,通过查询第二索引知识库,得到第二候选文本,其中,所述第二索引知识库包括候选文本的摘要的语义特征与候选文本的索引关系;以及

基于所述第一候选文本、所述第二候选文本与所述多个主题词的语义相关性,从所述第一候选文本和所述第二候选文本中得到所述参考文本。

3. 根据权利要求2所述的方法,其中,所述根据所述多个主题词,通过查询第一索引知识库,得到第一候选文本,包括:

将所述多个主题词与所述第一索引知识库的多个候选文本的摘要进行匹配,得到第一匹配概率;

基于所述第一匹配概率,从所述多个候选文本的摘要中得到第一目标摘要;以及

根据所述第一目标摘要,基于候选文本的摘要与候选文本的关联关系,得到所述第一候选文本。

4. 根据权利要求2所述的方法,其中,所述根据所述多个主题词,通过查询第二索引知识库,得到第二候选文本,包括:

提取所述多个主题词的多个语义特征;

将所述多个主题词的多个语义特征与所述第二索引知识库的多个候选文本的摘要的语义特征进行匹配,得到第二匹配概率;

基于所述第二匹配概率,从所述多个候选文本的摘要中得到第二目标摘要;以及

根据所述第二目标摘要,基于候选文本的摘要与候选文本的关联关系,得到所述第二候选文本。

5. 根据权利要求2所述的方法,其中,所述基于所述第一候选文本、所述第二候选文本与所述多个主题词的语义相关性,从所述第一候选文本和所述第二候选文本中得到所述参考文本,包括:

确定所述多个主题词的多个权重;

根据所述第一候选文本的摘要中的主题词、所述第二候选文本的摘要中的主题词、所述问题文本中的主题词和所述多个权重,得到主题词重要度;以及

根据所述主题词重要度,将所述第一候选文本、所述第二候选文本进行排序,得到所述参考文本。

6. 根据权利要求5所述的方法,还包括:

将预定主题词分别与所述第一候选文本的摘要、所述第二候选文本的摘要进行匹配，得到第三匹配概率；以及

根据所述主题词重要度和所述第三匹配概率，将所述第一候选文本、所述第二候选文本进行排序，得到所述参考文本。

7. 根据权利要求6所述的方法，还包括：

提取所述第一候选文本的摘要的第一语义特征和所述多个主题词的第二语义特征；

将所述第一语义特征与所述第二语义特征进行匹配，得到第四匹配概率；以及

根据所述主题词重要度、所述第三匹配概率和所述第四匹配概率，将所述第一候选文本、所述第二候选文本进行排序，得到所述参考文本。

8. 根据权利要求1所述的方法，还包括：

获取所述问题文本的上下文信息；以及

根据所述问题文本的上下文信息，对所述问题文本进行信息补充，生成语义完整的问题文本。

9. 根据权利要求1所述的方法，还包括：

按照所述目标对象已公开发表的历史文本的段落特征，对所述历史文本进行切分，得到多个文本段落；

提取所述多个文本段落的主题内容，得到与所述多个文本段落对应的多个段落摘要；以及

根据所述多个段落摘要和所述多个文本段落，构建所述目标知识库。

10. 根据权利要求9所述的方法，还包括：

提取所述历史文本的主题内容，得到与所述历史文本对应的文本摘要；

根据所述文本摘要，生成与所述文本摘要对应的问题文本；以及

根据所述与所述文本摘要对应的问题文本和所述文本摘要，构建所述目标知识库。

11. 一种文本生成装置，包括：

分词模块，用于按照问题文本中相邻词之间的语义相关性，对所述问题文本进行分词，得到多个主题词；

查询模块，用于根据所述多个主题词，通过查询目标知识库，得到与所述多个主题词对应的参考文本，其中，所述目标知识库是根据目标对象已公开发表的历史观点内容构建的；以及

第一生成模块，用于根据所述问题文本和所述参考文本，生成用于答复所述问题文本的答复文本。

12. 根据权利要求11所述的装置，其中，所述目标知识库包括第一索引知识库和第二索引知识库；所述查询模块包括：

第一查询子模块，用于根据所述多个主题词，通过查询第一索引知识库，得到第一候选文本，其中，所述第一索引知识库包括候选文本的摘要与候选文本的索引关系；

第二查询子模块，用于根据所述多个主题词，通过查询第二索引知识库，得到第二候选文本，其中，所述第二索引知识库包括候选文本的摘要的语义特征与候选文本的索引关系；以及

第一获得子模块，用于基于所述第一候选文本、所述第二候选文本与所述多个主题词

的语义相关性,从所述第一候选文本和所述第二候选文本中得到所述参考文本。

13. 根据权利要求12所述的装置,其中,所述第一查询子模块包括:

第一匹配单元,用于将所述多个主题词与所述第一索引知识库的多个候选文本的摘要进行匹配,得到第一匹配概率;

第一获得单元,用于基于所述第一匹配概率,从所述多个候选文本的摘要中得到第一目标摘要;以及

第二获得单元,用于根据所述第一目标摘要,基于候选文本的摘要与候选文本的关联关系,得到所述第一候选文本。

14. 根据权利要求12所述的装置,其中,所述第二查询子模块包括:

第一提取单元,用于提取所述多个主题词的多个语义特征;

第二匹配单元,用于将所述多个主题词的多个语义特征与所述第二索引知识库的多个候选文本的摘要的语义特征进行匹配,得到第二匹配概率;

第三获得单元,用于基于所述第二匹配概率,从所述多个候选文本的摘要中得到第二目标摘要;以及

第四获得单元,用于根据所述第二目标摘要,基于候选文本的摘要与候选文本的关联关系,得到所述第二候选文本。

15. 根据权利要求12所述的装置,其中,所述第一获得子模块包括:

确定单元,用于确定所述多个主题词的多个权重;

第五获得单元,用于根据所述第一候选文本的摘要中的主题词、所述第二候选文本的摘要中的主题词、所述问题文本中的主题词和所述多个权重,得到主题词重要度;以及

第六获得单元,用于根据所述主题词重要度,将所述第一候选文本、所述第二候选文本进行排序,得到所述参考文本。

16. 根据权利要求15所述的装置,还包括:

第三匹配单元,用于将预定主题词分别与所述第一候选文本的摘要、所述第二候选文本的摘要进行匹配,得到第三匹配概率;以及

第七获得单元,用于根据所述主题词重要度和所述第三匹配概率,将所述第一候选文本、所述第二候选文本进行排序,得到所述参考文本。

17. 根据权利要求16所述的装置,还包括:

第二提取单元,用于提取所述第一候选文本的摘要的第一语义特征和所述多个主题词的第二语义特征;

第四匹配单元,用于将所述第一语义特征与所述第二语义特征进行匹配,得到第四匹配概率;以及

第八获得单元,用于根据所述主题词重要度、所述第三匹配概率和所述第四匹配概率,将所述第一候选文本、所述第二候选文本进行排序,得到所述参考文本。

18. 根据权利要求11所述的装置,还包括:

获取模块,用于获取所述问题文本的上下文信息;以及

信息补充模块,用于根据所述问题文本的上下文信息,对所述问题文本进行信息补充,生成语义完整的问题文本。

19. 根据权利要求11所述的装置,还包括:

文本切分模块,用于按照所述目标对象已公开发表的历史文本的段落特征,对所述历史文本进行切分,得到多个文本段落;

第一提取模块,用于提取所述多个文本段落的主题内容,得到与所述多个文本段落对应的多个段落摘要;以及

第一构建模块,用于根据所述多个段落摘要和所述多个文本段落,构建所述目标知识库。

20. 根据权利要求11所述的装置,还包括:

第二提取模块,用于提取所述历史文本的主题内容,得到与所述历史文本对应的文本摘要;

第二生成模块,用于根据所述文本摘要,生成与所述文本摘要对应的问题文本;以及

第二构建模块,用于根据所述与所述文本摘要对应的问题文本和所述文本摘要,构建所述目标知识库。

21. 一种电子设备,包括:

至少一个处理器;以及

与所述至少一个处理器通信连接的存储器;其中,

所述存储器存储有可被所述至少一个处理器执行的指令,所述指令被所述至少一个处理器执行,以使所述至少一个处理器能够执行权利要求1-10中任一项所述的方法。

22. 一种存储有计算机指令的非瞬时计算机可读存储介质,其中,所述计算机指令用于使所述计算机执行根据权利要求1-10中任一项所述的方法。

23. 一种计算机程序产品,包括计算机程序,所述计算机程序在被处理器执行时实现根据权利要求1-10中任一项所述的方法。

文本生成方法、装置、电子设备以及存储介质

技术领域

[0001] 本公开涉及人工智能技术领域,尤其涉及自然语言处理、深度学习、大语言模型技术领域,具体涉及一种文本生成方法、装置、电子设备以及存储介质。

背景技术

[0002] LLM(Large Language Model,大语言模型),是一种基于机器学习和自然语言处理技术的生成式模型,它通过对大量的文本数据进行预训练,来学习服务人类语言理解和生成的能力。

[0003] 针对不同的应用场景,可以通过微调或调整大语言模型的提示输入(Prompt),使得大语言模型具有该应用场景的自然语言理解和生成能力。

发明内容

[0004] 本公开提供了一种文本生成方法、装置、电子设备以及存储介质。

[0005] 根据本公开的一方面,提供了一种文本生成方法,包括:按照问题文本中相邻词之间的语义相关性,对所述问题文本进行分词,得到多个主题词;根据多个主题词,通过查询目标知识库,得到与多个主题词对应的参考文本,其中,目标知识库是根据目标对象已公开发表的历史观点内容构建的;以及根据所述问题文本和所述参考文本,生成用于答复所述问题文本的答复文本。

[0006] 根据本公开的另一方面,提供了一种文本生成装置,包括:分词模块、查询模块和第一生成模块。分词模块,用于按照问题文本中相邻词之间的语义相关性,对所述问题文本进行分词,得到多个主题词。查询模块,用于根据所述多个主题词,通过查询目标知识库,得到与所述多个主题词对应的参考文本,其中,所述目标知识库是根据目标对象已公开发表的历史观点内容构建的。第一生成模块,用于根据所述问题文本和所述参考文本,生成用于答复所述问题文本的答复文本。

[0007] 根据本公开的另一方面,提供了一种电子设备,包括:至少一个处理器;以及与上述至少一个处理器通信连接的存储器;其中,上述存储器存储有可被上述至少一个处理器执行的指令,上述指令被上述至少一个处理器执行,以使上述至少一个处理器能够执行如上的方法。

[0008] 根据本公开的另一方面,提供了一种存储有计算机指令的非瞬时计算机可读存储介质,其中,上述计算机指令用于使上述计算机执行如上的方法。

[0009] 根据本公开的另一方面,提供了一种计算机程序产品,包括计算机程序,上述计算机程序在被处理器执行时实现如上的方法。

[0010] 应当理解,本部分所描述的内容并非旨在标识本公开的实施例的关键或重要特征,也不用于限制本公开的范围。本公开的其它特征将通过以下的说明书而变得容易理解。

附图说明

[0011] 附图用于更好地理解本方案,不构成对本公开的限定。其中:

[0012] 图1示意性示出了根据本公开实施例的可以应用文本生成方法及装置的示例性系统架构;

[0013] 图2示意性示出了根据本公开实施例的文本生成方法的流程图;

[0014] 图3示意性示出了根据本公开实施例的文本生成方法的原理图;

[0015] 图4示意性示出了根据本公开实施例的从目标对象的历史观点内容中进行知识抽取的示意图;

[0016] 图5示意性示出了根据本公开实施例的根据主题词进行知识搜索的示意图;

[0017] 图6示意性示出了根据本公开另一实施例的文本生成方法的示意图;

[0018] 图7示意性示出了根据本公开实施例的文本生成装置的框图;以及

[0019] 图8示意性示出了根据本公开实施例的适于实现文本生成方法的电子设备的框图。

具体实施方式

[0020] 以下结合附图对本公开的示范性实施例做出说明,其中包括本公开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,而不会背离本公开的范围和精神。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。

[0021] 在公共网络平台上,用户可以针对创作者发布的话题在评论区进行提问,当评论区内待回复的提问数量过多时,创作者人工回复评论的效率较低。

[0022] 为了提高回复评论的效率,相关技术中通过调用生成式大语言模型(LLM),利用大语言模型的问答交互能力对评论区的问题进行自动回复。

[0023] 但是,目标基于大语言模型直接生成的回复内容是普适性的,不具有创作者的个性化风格,也不能明确体现出创作者的个人观点,因此,较难满足用户与创作者的互动需求。

[0024] 有鉴于此,本公开实施例提供了一种文本生成方法,通过查询目标知识库,得到与问题相对应的参考文本,再基于问题文本和参考文本,生成答复文本。由于目标知识库是根据创作者已公开发表的历史观点内容构建的,汇集了创作者的个人观点,因此,可以生成具有创作者个性化风格的回复文本,满足用户与创作者的互动需求,提高用户体验。

[0025] 图1示意性示出了根据本公开实施例的可以应用文本生成方法及装置的示例性系统架构。

[0026] 需要注意的是,图1所示仅为可以应用本公开实施例的系统架构的示例,以帮助本领域技术人员理解本公开的技术内容,但并不意味着本公开实施例不可以用于其他设备、系统、环境或场景。例如,在另一实施例中,可以应用文本生成方法及装置的示例性系统架构可以包括终端设备,但终端设备可以无需与服务器进行交互,即可实现本公开实施例提供的文本生成方法及装置。

[0027] 如图1所示,根据该实施例的系统架构100可以包括终端设备101、102、103,网络104和服务器105。网络104用以在终端设备101、102、103和服务器105之间提供通信链路的

介质。网络104可以包括各种连接类型,例如有线和/或无线通信链路等等。

[0028] 用户可以使用终端设备101、102、103通过网络104与服务器105交互,以接收或发送消息等。终端设备101、102、103上可以安装有各种通讯客户端应用,例如知识阅读类应用、网页浏览器应用、搜索类应用、即时通信工具、邮箱客户端和/或社交平台软件等(仅为示例)。

[0029] 终端设备101、102、103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。

[0030] 服务器105可以是提供各种服务的服务器,例如对用户利用终端设备101、102、103所浏览的内容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对接收到的用户请求等数据进行分析等处理,并将处理结果(例如根据用户请求获取或生成的网页、信息、或数据等)反馈给终端设备。

[0031] 需要说明的是,本公开实施例所提供的文本生成方法一般可以由终端设备101、102、或103执行。相应地,本公开实施例所提供的文本生成装置也可以设置于终端设备101、102、或103中。

[0032] 或者,本公开实施例所提供的文本生成方法一般也可以由服务器105执行。相应地,本公开实施例所提供的文本生成装置一般可以设置于服务器105中。本公开实施例所提供的文本生成方法也可以由不同于服务器105且能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群执行。相应地,本公开实施例所提供的文本生成装置也可以设置于不同于服务器105且能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群中。

[0033] 例如,在用户在评论区进行提问时,终端设备101、102、103可以获取评论区的问题文本和该评论区的创作者信息,然后将获取的问题文本和创作者信息发送给服务器105,由服务器105对问题文本进行分词,得到多个主题词;根据创作者信息确定目标知识库;并根据多个主题词,通过查询目标知识库,得到与多个主题词对应的参考文本;以及根据问题文本和参考文本,生成用于答复问题文本的答复文本。或者由能够与终端设备101、102、103和/或服务器105通信的服务器或服务器集群对问题文本进行分析,并最终实现生成具有创作者个人风格的答复文本。

[0034] 应该理解,图1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、网络和服务器的。

[0035] 在本公开的技术方案中,所涉及的用户个人信息的收集、存储、使用、加工、传输、提供、公开和应用等处理,均符合相关法律法规的规定,采取了必要保密措施,且不违背公序良俗。

[0036] 在本公开的技术方案中,在获取或采集用户个人信息之前,均获取了用户的授权或同意。

[0037] 图2示意性示出了根据本公开实施例的文本生成方法的流程图。

[0038] 如图2所示,该方法包括操作S210~S230。

[0039] 在操作S210,按照问题文本中相邻词之间的语义相关性,对问题文本进行分词,得到多个主题词。

[0040] 在操作S220,根据多个主题词,通过查询目标知识库,得到与多个主题词对应的参

考文本。

[0041] 在操作S230,根据问题文本和参考文本,生成用于答复问题文本的答复文本。

[0042] 根据本公开的实施例,问题文本可以是用户在创作者的评论区发起的提问内容。按照问题文本中相邻词之间的语义相关性,对问题文本进行分词,可以实现分词后得到的每一个主题词自身的语义完整性。

[0043] 例如:问题文本可以是“甲省乙区的师范生考试成绩公布时间是什么时候?”按照相邻词之间的语义相关性进行分词后,得到的主题词可以包括:“甲省乙区”、“师范生”、“考试成绩”、“公布时间”。

[0044] 根据本公开的实施例,按照相邻词之间的语义相关性进行分词,可以有效避免无效的检索内容。例如:地名“甲省乙区”被分词成为两个主题词“甲省”、“乙区”,那么在检索结果中可能会存在大量的甲省师范生考试成绩的公布时间。

[0045] 根据本公开的实施例,目标知识库是根据目标对象已公开发表的历史观点内容构建的。目标对象可以是回复文本的创作者,例如:可以是用户进行提问的评论区的博主,也可以是用户问题文本中提到的期望回复该问题的对象。例如:“甲省乙区的师范生报考条件是什么?@A老师”,该问题文本中的A老师可以是目标对象。

[0046] 根据本公开的实施例,目标对象已公开发表的历史观点内容,可以是目标对象通过文本、图像、视频、音频等各种方式公开发表的观点内容。

[0047] 例如:视频内容可以是:“A老师的职业生涯小课堂”,音频内容可以是:“A老师讲职场”,图像、文本内容可以是A老师在自媒体公众号上发表的文章等。

[0048] 需要说明的是,本公开实施例是在获得创作者授权之后,获取该创作者的历史观点内容,以构建目标知识库的,符合相关法律的规定。

[0049] 根据本公开的实施例,根据多个主题词,可以通过查询目标知识库,得到能够用于回复问题的参考文本。由于目标知识库集成了创作者的历史观点内容,因此,参考文本也是具有创作者的个性化观点的内容文本。

[0050] 根据本公开的实施例,可以将问题文本、参考文本以及文本生成指令组装成生成式大模型的提示(prompt)输入大语言模型(LLM),生成用于答复问题文本的答复文本。

[0051] 根据本公开的实施例,通常大语言模型在生成答复文本的过程中,会在答复文本的前后随机添加一些前后缀的关联语句,例如:“回复=”,可以对大语言模型生成的文本进行解析,删除模型随机添加的前后缀的关联语句,得到答复文本。

[0052] 根据本公开的实施例,通过查询目标知识库,得到与问题相对应的参考文本,再基于问题文本和参考文本,生成答复文本。由于目标知识库是根据创作者已公开发表的历史观点内容构建的,汇集了创作者的个人观点,因此,可以生成具有创作者个性化风格的回复文本,满足用户与创作者的互动需求,提高用户体验。

[0053] 在实际应用场景中,评论区中的内容质量参差不齐,例如:可能包含负面情绪的言论,也可能不具有明确的提问意图或提问内容的语义不完整。

[0054] 针对包含负面情绪的言论或不具有明确的提问意图的评论,可以通过构建黑名单词表、无效句式的正则表达式等方式对评论区的内容进行筛选。

[0055] 针对提问内容的语义不完整的评论,可以获取问题文本的上下文信息;以及根据问题文本的上下文信息,对问题文本进行信息补充,生成语义完整的问题文本。

[0056] 例如:在讲解“省考师范生报考条件”的内容下面的评论内容可以是“只招丙地区的?”,则问题文本可以是“只招丙地区的”,显然该问题的语义并不完整,基于当前的问题并不能表达出用户的真实提问意图。因此,可以获取该问题的上下文信息“省考师范生报考条件”,并结合上下文信息“省考师范生报考条件”对问题“只招丙地区的?”进行信息补充,生成语义完整的问题文本“省考师范生只招丙地区的吗?”

[0057] 根据本公开的实施例,可以将问题文本和上下文信息作为预训练的LLM模型的输入,生成语义完整的问题文本。

[0058] 根据本公开的实施例,基于问题文本的上下文对问题文本进行信息补充,可以提高问题文本的语义完整性,以便得到用户真实的提问意图,提高回复文本与问题文本的相关度。

[0059] 图3示意性示出了根据本公开实施例的文本生成方法的原理图。

[0060] 如图3所示,在实施例300中,可以将上下文信息+问题311通过数据增强,得到语义完整的问题312。对语义完整的问题312进行分词,得到主题词313。基于主题词313,通过查询目标知识库325得到与问题最相关的知识326。将语义完整的问题312和与问题最相关的知识326作为提示输入大语言模型,生成答复文本327。

[0061] 目标知识库325是基于授权创作者已公开发表的内容321构建得到的授权创作者已公开发表的内容321可以通过多模态识别得到历史文本322对历史文本进行知识提取,得到段落文本一段落摘要323_1和问答形式的篇章主题一篇章摘要323_2。通过提取检索特征,得到检索特征一知识324,并将提取到的知识和检索特征入库,得到目标知识库325。

[0062] 根据本公开的实施例,目标对象已公开发表的历史观点内容可以是视频、音频、图文等多种格式,因此,可以通过多模态识别技术,例如:OCR(Optical Character Recognition,光学字符识别)、ASR(Automatic Speech Recognition,语音识别),将不同格式的历史观点内容转换为文本格式。

[0063] 由于大语言模型的输入长度限制,因此,引入目标知识库对大语言模型进行知识增强时,需要对创作者的历史观点内容进行有效抽取和压缩,并尽可能的减少抽取和压缩过程对观点知识的损失。

[0064] 例如:按照目标对象已公开发表的历史文本的段落特征,对历史文本进行切分,得到多个文本段落;提取多个文本段落的主题内容,得到与多个文本段落对应的多个段落摘要;以及根据多个段落摘要和多个文本段落,构建目标知识库。

[0065] 根据本公开的实施例,由于在一篇文章中,每一个自然段落的内容通常是针对一个核心主题思想的阐述,因此,可以按照历史文本的原始自然段落对历史文本进行切分,得到多个文本段落。以便于提取该段落的段落摘要。段落摘要可以表征该段落文本的核心主题思想。

[0066] 对于一篇文章而言,还可以以问答形式提取文章的主题思想和文章的摘要内容。

[0067] 例如:提取历史文本的主题内容,得到与历史文本对应的文本摘要;根据文本摘要,生成与文本摘要对应的问题文本;以及根据与文本摘要对应的问题文本和文本摘要,构建目标知识库。

[0068] 图4示意性示出了根据本公开实施例的从目标对象的历史观点内容中进行知识抽取的示意图。

[0069] 如图4所示,在实施例400中,创作者已发表的历史观点内容可以包括图文——“楚人失弓”故事道出了何种境界441和音视频“小A讲故事“楚人失弓”。通过多模态识别,将历史观点内容转化为历史文本443。

[0070] 按照段落进行知识抽取的结果如444所示:段落摘要:儒家:以人文本,修身齐家治国平天下。段落文本:01首先我们来说儒家,孔子在这件事情上,对楚王的评价...儒家讲究一个“修身、齐家、治国、平天下,其中修身是万物之始。

[0071] 按照篇章进行知识抽取的结果如445所示:问:“楚人失弓”的故事体现了儒释道三家的哪些境界?答:这个故事体现了儒释道三家的境界包括:儒家的“修身、齐家、治国、平天下”,佛家的“空”的境界,以及道家的“平等”观念。

[0072] 根据本公开的实施例,按照段落进行知识抽取,可以尽可能减少知识抽取过程中的知识损失,保留较为详尽的创作者观点。按照篇章进行知识抽取,可以尽可能对知识进行压缩,缩小文本长度。

[0073] 由于段落知识抽取和篇章知识抽取两种方式各有利弊,可以将两种知识抽取的结果均保存在创作者的知识库中。可以通过为两种抽取方式得到的知识配置不同的权重,以便于满足不同应用场景对知识检索精度的需求。

[0074] 根据本公开的实施例,目标知识库可以包括第一索引知识库和第二索引知识库。第一索引知识库可以包括候选文本的摘要与候选文本的索引关系,例如:第一索引知识库可以是倒排索引库,索引式可以是候选文本的摘要与候选文本的标识编码的映射。第二索引知识库可以包括候选文本的摘要的语义特征与候选文本的索引关系。例如:第二索引知识库可以是语义向量库,索引式可以是候选文本的摘要的语义特征向量与候选文本的标识编码的映射。

[0075] 根据本公开的实施例,根据多个主题词,通过查询目标知识库,得到与多个主题词对应的参考文本,可以包括如下操作:根据多个主题词,通过查询第一索引知识库,得到第一候选文本;根据多个主题词,通过查询第二索引知识库,得到第二候选文本;基于第一候选文本、第二候选文本与多个主题词的语义相关性,从第一候选文本和第二候选文本中得到参考文本。

[0076] 根据本公开的实施例,基于第一索引知识库的检索,可以通过设定召回条件筛选检索结果。基于第二索引知识库的检索,可以通过对检索结果进行正向排序,再进行特征筛选。

[0077] 根据本公开的实施例,第一索引知识库的召回条件可以是与主题词的匹配概率。

[0078] 例如:将多个主题词与第一索引知识库的多个候选文本的摘要进行匹配,得到第一匹配概率;基于第一匹配概率,从多个候选文本的摘要中得到第一目标摘要;以及根据第一目标摘要,基于候选文本的摘要与候选文本的关联关系,得到第一候选文本。

[0079] 例如:多个主题词可以是“甲地乙区”、“师范生”、“考试时间”、“公布”,对于候选文本的摘要中仅包括上述一个主题词的,第一匹配概率可以是25%,对于候选文本的摘要中包括上述两个主题词的,第一匹配概率可以是50%,以此类推。可以通过设定第一匹配概率阈值,例如:75%,将第一匹配概率大于75%的候选文本的摘要作为第一目标摘要。再根据摘要与文本的索引关系,得到第一候选文本。

[0080] 根据本公开的实施例,基于主题词与候选文本摘要的匹配对候选文本进行筛选,

检索结果覆盖的范围较为全面,漏检率较低。

[0081] 参考文本与问题的语义相关性是决定最终生成的答复文本与问题的语义相关性的关键因此,因此,可以基于第二索引知识库,按照主题词的语义特征进行检索。

[0082] 例如:可以提取多个主题词的多个语义特征;将多个主题词的多个语义特征与第二索引知识库的多个候选文本的摘要的语义特征进行匹配,得到第二匹配概率;基于第二匹配概率,从多个候选文本的摘要中得到第二目标摘要;以及根据第二目标摘要,基于候选文本的摘要与候选文本的关联关系,得到第二候选文本。

[0083] 根据本公开的实施例,可以将多个主题词的多个语义特征进行选择性地拼接,例如:共有5个主题词,随机选择3个主题词的语义特征进行拼接,则主题词占比率可以是60%。基于拼接后的语义特征与候选文本的摘要的语义特征进行相似度计算,得到的相似度。再根据主题词占比率和相似度的乘积,得到第二匹配概率。可以通过设定第二匹配概率阈值,将大于第二匹配概率阈值的候选文本的摘要确定为第二目标摘要。再根据摘要的语义特征与文本的索引关系,得到第二候选文本。

[0084] 根据本公开的实施例,基于主题词的语义特征与候选文本摘要的匹配对候选文本进行筛选,检索到的第二候选文本与问题文本的语义相关性较高。

[0085] 基于第一索引知识库的检索结果范围较全面,基于第二索引知识库的检索结果与问题文本的语义相关性较高,因此,可以将两个检索结果进行混合排序,以提高检索精度。

[0086] 另外,在实际应用场景中,不同的主题词对该问题的重要性可以是不同的,可以基于主题词对问题的重要性,配置该主题词的权重。

[0087] 例如:确定多个主题词的多个权重;根据第一候选文本的摘要中的主题词、第二候选文本的摘要中的主题词、问题文本中的主题词和多个权重,得到主题词重要度;根据主题词重要度,将第一候选文本、第二候选文本进行排序,得到参考文本。

[0088] 根据本公开的实施例,主题词的匹配概率是从主题词的重要性的维度表征检索结果的命中率。

[0089] 例如:可以按照式(1)和(2)计算主题词重要度:

[0090] $Cqr =$ 候选文本摘要和问题文本均命中的主题词的权重之和/问题文本中主题词的权重之和;(1)

[0091] $Ctr =$ 候选文本摘要和问题文本均命中的主题词的权重之和/候选文本摘要中主题词的权重之和。(2)

[0092] 其中, Cqr 表示第一命中主题词的重要度; Ctr 表示第二命中主题词的重要度。

[0093] 根据本公开的实施例,可以对第一命中主题词的重要度和第二命中主题词的重要度进行加权求和,得到主题词重要度。

[0094] 根据本公开的实施例,可以按照每一个候选文本的主题词重要度,对第一候选文本、第二候选文本进行正向排序,得到排序结果。参考文本可以是排序结果中排在前100位的文本。

[0095] 根据本公开的实施例,按照主题词重要度,针对性地从候选文本中筛选出较为重要的主题词对应的文本,提高检索精度。

[0096] 为了进一步从候选文本中筛选出与问题文本最相关的参考文本,还可以基于对候选文本进行不可省词识别,对候选文本进行筛选。不可省词可以是指在答复文本中不能缺

少的词。

[0097] 例如:将预定主题词分别与第一候选文本的摘要、第二候选文本的摘要进行匹配,得到第三匹配概率;以及根据主题词重要度和第三匹配概率,将第一候选文本、第二候选文本进行排序,得到参考文本。

[0098] 根据本公开的实施例,预定主题词可以是候选文本中的不可省词,例如:人名、地名,针对不同的应用场景,可以设定不同的预定主题词。第三匹配概率可以表征候选文本的摘要中是否包括不可省词,当候选文本的摘要包括不可省词时,第三匹配概率为1,否则为0。

[0099] 例如:不可省词可以是“甲省”,候选文本 T_1 的摘要可以是“全国高考成绩公布时间是XX月XX日”;候选文本 T_2 的摘要可以是“甲省高考成绩公布时间是YY月YY日”。由于不可省词是“甲省”,候选文本 T_1 的摘要中不包括该词,候选文本 T_2 的摘要中包括该词,因此,可以将候选文本 T_2 确定为参考文本。

[0100] 根据本公开的实施例,可以为主题词重要度和第三匹配概率配置不同的权重,基于主题词重要度的权重、第三匹配概率的权重,对主题词重要度和第三匹配概率进行加权求和,并基于加权求和的结果对第一候选文本、第二候选文本进行排序,得到排序结果。参考文本可以是排序结果中排在前1000位的文本。

[0101] 根据本公开的实施例,通过增加不可省词的识别,可以从候选文本中筛选出与问题文本更加相关的参考文本,从而生成有价值的答复文本。

[0102] 由于第一索引知识库是基于主题词与候选文本摘要的匹配进行检索的,因此,第一候选文本中可能存在与问题的语义相关度较低的文本。为了排除第一候选文本中与问题的语义相关度较低的文本,可以采用如下操作:提取第一候选文本的摘要的第一语义特征和多个主题词的第二语义特征;将第一语义特征与第二语义特征进行匹配,得到第四匹配概率;以及根据主题词重要度、第三匹配概率和第四匹配概率,将第一候选文本、第二候选文本进行排序,得到参考文本。

[0103] 根据本公开的实施例,第二语义特征可以是问题文本中得到的全部主题词的语义特征。

[0104] 根据本公开的实施例,第四匹配概率可以是第一语义特征和第二语义特征的相似度。相似度算法可以是余弦相似度、欧式距离等可以应用于计算文本相似度的任何算法,本公开实施例对此不作具体限定。

[0105] 根据本公开的实施例,可以为主题词重要度、第三匹配概率、第四匹配概率配置不同的权重,基于主题词重要度的权重、第三匹配概率的权重和第四匹配概率的权重,对主题词重要度、第三匹配概率和第四匹配概率进行加权求和,并基于加权求和的结果对第一候选文本、第二候选文本进行排序,得到排序结果。参考文本可以是排序结果中排在前3000位的文本。

[0106] 根据本公开的实施例,基于候选文本的主题词重要度、不可省词识别和语义相关度的综合评定结果,可以得到与问题文本最相关的参考文本,以便提高答复文本的生成质量。

[0107] 图5示意性示出了根据本公开实施例的根据主题词进行知识搜索的示意图。

[0108] 如图5所示,在实施例500中,分别根据主题词521对term倒排库522、ANN语义向量

库523进行检索,其中,ANN语义向量库523的检索结果需要进入正排库524进行进一步的信息筛选,最终得到合并候选文本525。根据主题词521和合并候选文本525进行分析,得到候选文本与问题的相关性526。按照候选文本与问题的相关性526,从合并候选文本525中筛选得到与问题最相关的知识527。

[0109] 在实施例500中还包括知识入库的过程,即授权创作者已公开发表的内容5201通过ASR、OCR等多模态算法的识别得到历史文本5202。通过对历史文本5202进行切分后,输入LLM模型,得到段落文本一段落摘要5203_1和问答形式的篇章主题一篇章摘要5203_2。基于粗粒度切词、词性、term重要性(词重要性)对段落文本一段落摘要5203_1和问答形式的篇章主题一篇章摘要5203_2处理之后,存储在Term倒排库522中。基于粗粒度切词、词性、term重要性(词重要性)、摘要的语义特征对段落文本一段落摘要5203_1和问答形式的篇章主题一篇章摘要5203_2处理之后,存储在ANN语义向量库523中。

[0110] 图6示意性示出了根据本公开另一实施例的文本生成方法的示意图。

[0111] 如图6所示,在实施例600中,基于上下文信息“甲地B校公费生的报考条件是什么?”和问题“我也在关注这个呢,都说只招甲地的?”601,利用LLM进行筛选提炼和信息补充,得到语义完整的问题:“请问甲地B校公费生的报考条件是什么?是否只招甲地的?”602。

[0112] 基于创作者信息“A老师生涯课堂uid”检索A老师的知识库,得到参考文本{历史问答:问:甲地B校公费生的报考条件是什么?答:甲地B校的报考条件包括分数、体检和单科成绩,……一般都在xx分以上的考生才能报考。按照招生计划,没有对生源所在地的限制,……}603。将语义完整的问题“请问甲地B校公费生的报考条件是什么?是否只招甲地的?”和参考文本{历史问答:问:甲地B校公费生的报考条件是什么?答:甲地B校的报考条件包括分数、体检和单科成绩,……一般都在xx分以上的考生才能报考。按照招生计划,没有对生源所在地的限制,……}603组装成Prompt,作为提示604输入LLM模型,得到答复文本“甲地B校公费生的报考条件中,没有对生源所在地的限制”605。

[0113] 图7示意性示出了根据本公开实施例的文本生成装置的框图。

[0114] 如图7所示,该装置700可以包括分词模块710、查询模块720和第一生成模块730。

[0115] 分词模块710,用于按照问题文本中相邻词之间的语义相关性,对问题文本进行分词,得到多个主题词;

[0116] 查询模块720,用于根据多个主题词,通过查询目标知识库,得到与多个主题词对应的参考文本,其中,目标知识库是根据目标对象已公开发表的历史观点内容构建的;以及

[0117] 第一生成模块730,用于根据问题文本和参考文本,生成用于答复问题文本的答复文本。

[0118] 根据本公开的实施例,目标知识库包括第一索引知识库和第二索引知识库;查询模块可以包括:第一查询子模块、第二查询子模块和第一获得子模块。

[0119] 第一查询子模块,用于根据多个主题词,通过查询第一索引知识库,得到第一候选文本,其中,第一索引知识库包括候选文本的摘要与候选文本的索引关系。

[0120] 第二查询子模块,用于根据多个主题词,通过查询第二索引知识库,得到第二候选文本,其中,第二索引知识库包括候选文本的摘要的语义特征与候选文本的索引关系。

[0121] 第一获得子模块,用于基于第一候选文本、第二候选文本与多个主题词的语义相关性,从第一候选文本和第二候选文本中得到参考文本。

[0122] 根据本公开的实施例,第一查询子模块可以包括:第一匹配单元、第一获得单元和第二获得单元。

[0123] 第一匹配单元,用于将多个主题词与第一索引知识库的多个候选文本的摘要进行匹配,得到第一匹配概率。

[0124] 第一获得单元,用于基于第一匹配概率,从多个候选文本的摘要中得到第一目标摘要。

[0125] 第二获得单元,用于根据第一目标摘要,基于候选文本的摘要与候选文本的关联关系,得到第一候选文本。

[0126] 根据本公开的实施例,第二查询子模块可以包括:第一提取单元、第二匹配单元、第三获得单元和第四获得单元。

[0127] 第一提取单元,用于提取多个主题词的多个语义特征。

[0128] 第二匹配单元,用于将多个主题词的多个语义特征与第二索引知识库的多个候选文本的摘要的语义特征进行匹配,得到第二匹配概率。

[0129] 第三获得单元,用于基于第二匹配概率,从多个候选文本的摘要中得到第二目标摘要。

[0130] 第四获得单元,用于根据第二目标摘要,基于候选文本的摘要与候选文本的关联关系,得到第二候选文本。

[0131] 根据本公开的实施例,第一获得子模块可以包括:确定单元、第五获得单元和第六获得单元。

[0132] 确定单元,用于确定多个主题词的多个权重。

[0133] 第五获得单元,用于根据第一候选文本的摘要中的主题词、第二候选文本的摘要中的主题词、问题文本中的主题词和多个权重,得到主题词重要度。

[0134] 第六获得单元,用于根据主题词重要度,将第一候选文本、第二候选文本进行排序,得到参考文本。

[0135] 根据本公开的实施例,上述第一获得子模块还包括:第三匹配单元和第七获得单元。第三匹配单元,用于将预定主题词分别与第一候选文本的摘要、第二候选文本的摘要进行匹配,得到第三匹配概率。第七获得单元,用于根据主题词重要度和第三匹配概率,将第一候选文本、第二候选文本进行排序,得到参考文本。

[0136] 根据本公开的实施例,上述第一获得子模块还包括:第二提取单元、第四匹配单元和第八获得单元。第二提取单元,用于提取第一候选文本的摘要的第一语义特征和多个主题词的第二语义特征。第四匹配单元,用于将第一语义特征与第二语义特征进行匹配,得到第四匹配概率。第八获得单元,用于根据主题词重要度、第三匹配概率和第四匹配概率,将第一候选文本、第二候选文本进行排序,得到参考文本。

[0137] 根据本公开的实施例,上述装置还包括:文本切分模块、第一提取模块和第一构建模块。文本切分模块,用于按照目标对象已公开发表的历史文本的段落特征,对历史文本进行切分,得到多个文本段落。第一提取模块,用于提取多个文本段落的主题内容,得到与多个文本段落对应的多个段落摘要。第一构建模块,用于根据多个段落摘要和多个文本段落,构建目标知识库。

[0138] 根据本公开的实施例,上述装置还包括:第二提取模块、第二生成模块和第二构建

模块。第二提取模块,用于提取历史文本的主题内容,得到与历史文本对应的文本摘要。第二生成模块,用于根据文本摘要,生成与文本摘要对应的问题文本。第二构建模块,用于根据与文本摘要对应的问题文本和文本摘要,构建目标知识库。

[0139] 根据本公开的实施例,本公开还提供了一种电子设备、一种可读存储介质和一种计算机程序产品。

[0140] 根据本公开的实施例,一种电子设备,包括:至少一个处理器;以及与至少一个处理器通信连接的存储器;其中,存储器存储有可被至少一个处理器执行的指令,指令被至少一个处理器执行,以使至少一个处理器能够执行如上的方法。

[0141] 根据本公开的实施例,一种存储有计算机指令的非瞬时计算机可读存储介质,其中,计算机指令用于使计算机执行如上的方法。

[0142] 根据本公开的实施例,一种计算机程序产品,包括计算机程序,计算机程序在被处理器执行时实现如上的方法。

[0143] 图8示出了可以用来实施本公开的实施例的示例电子设备800的示意性框图。电子设备旨在表示各种形式的数字计算机,诸如,膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置,诸如,个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例,并且不意在限制本文中描述的和/或者要求的本公开的实现。

[0144] 如图8所示,设备800包括计算单元801,其可以根据存储在只读存储器(ROM)802中的计算机程序或者从存储单元808加载到随机访问存储器(RAM)803中的计算机程序,来执行各种适当的动作和处理。在RAM 803中,还可存储设备800操作所需的各种程序和数据。计算单元801、ROM 802以及RAM 803通过总线804彼此相连。输入/输出(I/O)接口805也连接至总线804。

[0145] 设备800中的多个部件连接至I/O接口805,包括:输入单元806,例如键盘、鼠标等;输出单元807,例如各种类型的显示器、扬声器等;存储单元808,例如磁盘、光盘等;以及通信单元809,例如网卡、调制解调器、无线通信收发机等。通信单元809允许设备800通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

[0146] 计算单元801可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元801的一些示例包括但不限于中央处理单元(CPU)、图形处理单元(GPU)、各种专用的人工智能(AI)计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器(DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元801执行上文所描述的各个方法和处理,例如文本生成方法。例如,在一些实施例中,文本生成方法可被实现为计算机软件程序,其被有形地包含于机器可读介质,例如存储单元808。在一些实施例中,计算机程序的部分或者全部可以经由ROM 802和/或通信单元809而被载入和/或安装到设备800上。当计算机程序加载到RAM 803并由计算单元801执行时,可以执行上文描述的文本生成方法的一个或多个步骤。备选地,在其他实施例中,计算单元801可以通过其他任何适当的方式(例如,借助于固件)而被配置为执行文本生成方法。

[0147] 本文中以上描述的系统和技术和各种实施方式可以在数字电子电路系统、集成电路系统、场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、芯片上系统

的系统(SOC)、复杂可编程逻辑设备(CPLD)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括:实施在一个或者多个计算机程序中,该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该可编程处理器可以是专用或者通用可编程处理器,可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令,并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

[0148] 用于实施本公开的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器或控制器,使得程序代码当由处理器或控制器执行时使流程图和/或框图中所规定的功能/操作被实施。程序代码可以完全在机器上执行、部分地在机器上执行,作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务服务器上执行。

[0149] 在本公开的上下文中,机器可读介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、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或快闪存储器)、光纤、便捷式紧凑盘只读存储器(CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

[0150] 为了提供与用户的交互,可以在计算机上实施此处描述的系统和技术,该计算机具有:用于向用户显示信息的显示装置(例如,CRT(阴极射线管)或者LCD(液晶显示器)监视器);以及键盘和指向装置(例如,鼠标或者轨迹球),用户可以通过该键盘和该指向装置来将输入提供给计算机。其它种类的装置还可以用于提供与用户的交互;例如,提供给用户的反馈可以是任何形式的传感反馈(例如,视觉反馈、听觉反馈、或者触觉反馈);并且可以用任何形式(包括声输入、语音输入或者、触觉输入)来接收来自用户的输入。

[0151] 可以将此处描述的系统和技术实施在包括后台部件的计算系统(例如,作为数据服务器)、或者包括中间件部件的计算系统(例如,应用服务器)、或者包括前端部件的计算系统(例如,具有图形用户界面或者网络浏览器的用户计算机,用户可以通过该图形用户界面或者该网络浏览器来与此处描述的系统和技术实施方式交互)、或者包括这种后台部件、中间件部件、或者前端部件的任何组合的计算系统中。可以通过任何形式或者介质的数字数据通信(例如,通信网络)来将系统的部件相互连接。通信网络的示例包括:局域网(LAN)、广域网(WAN)和互联网。

[0152] 计算机系统可以包括客户端和服务端。客户端和服务端一般远离彼此并且通常通过通信网络进行交互。通过在相应的计算机上运行并且彼此具有客户端-服务器关系的计算机程序来产生客户端和服务端的关系。服务器可以是云服务器,也可以是分布式系统的服务器,或者是结合了区块链的服务器。

[0153] 应该理解,可以使用上面所示的各种形式的流程,重新排序、增加或删除步骤。例如,本发公开中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行,只要能够实现本公开公开的技术方案所期望的结果,本文在此不进行限制。

[0154] 上述具体实施方式,并不构成对本公开保护范围的限制。本领域技术人员应该明白的是,根据设计要求和因素,可以进行各种修改、组合、子组合和替代。任何在本公开的精神和原则之内所作的修改、等同替换和改进等,均应包含在本公开保护范围之内。

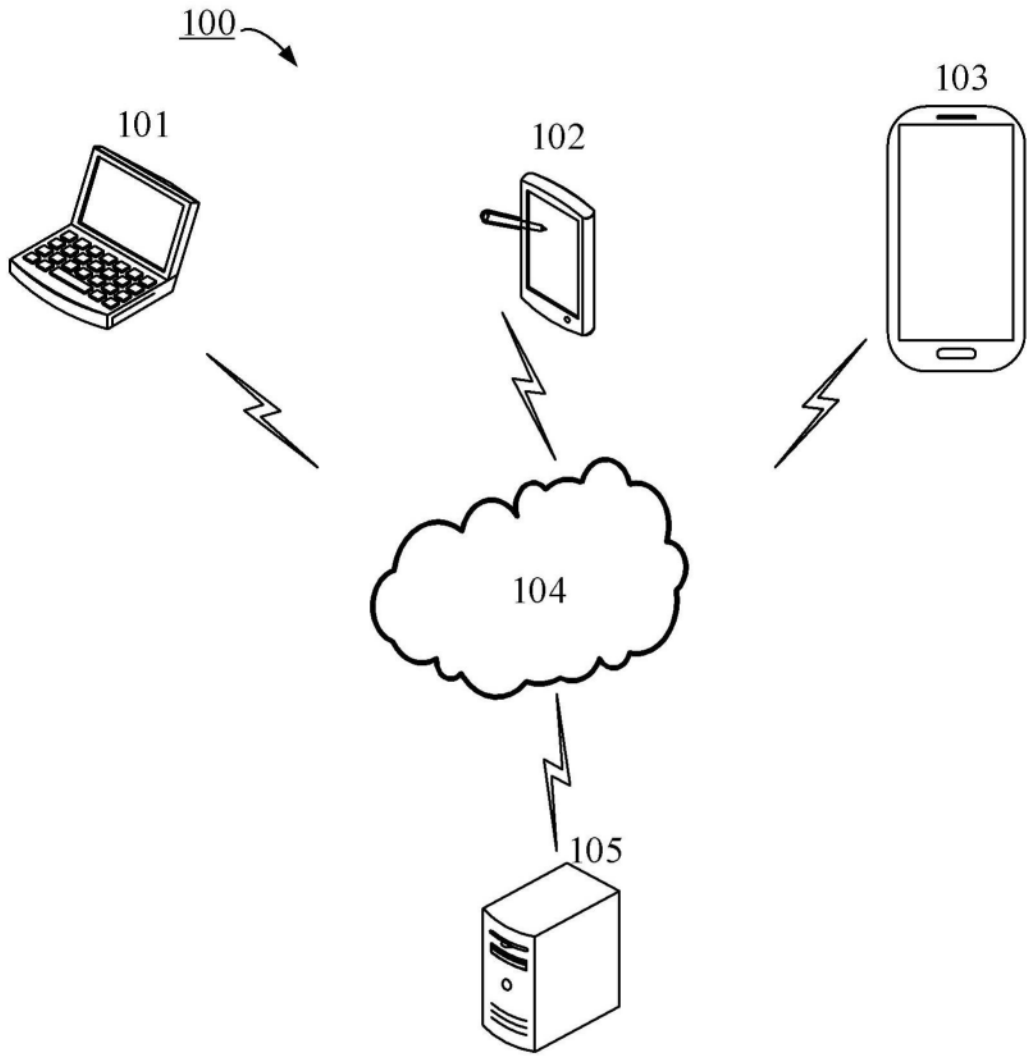


图1

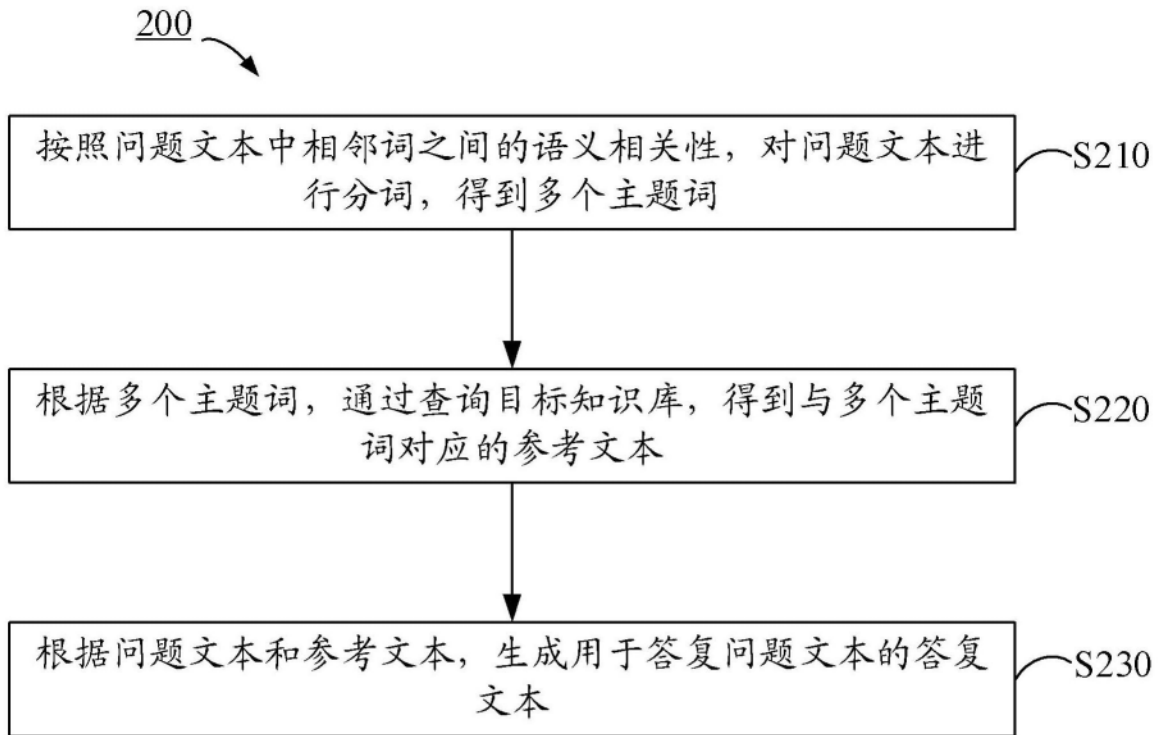


图2

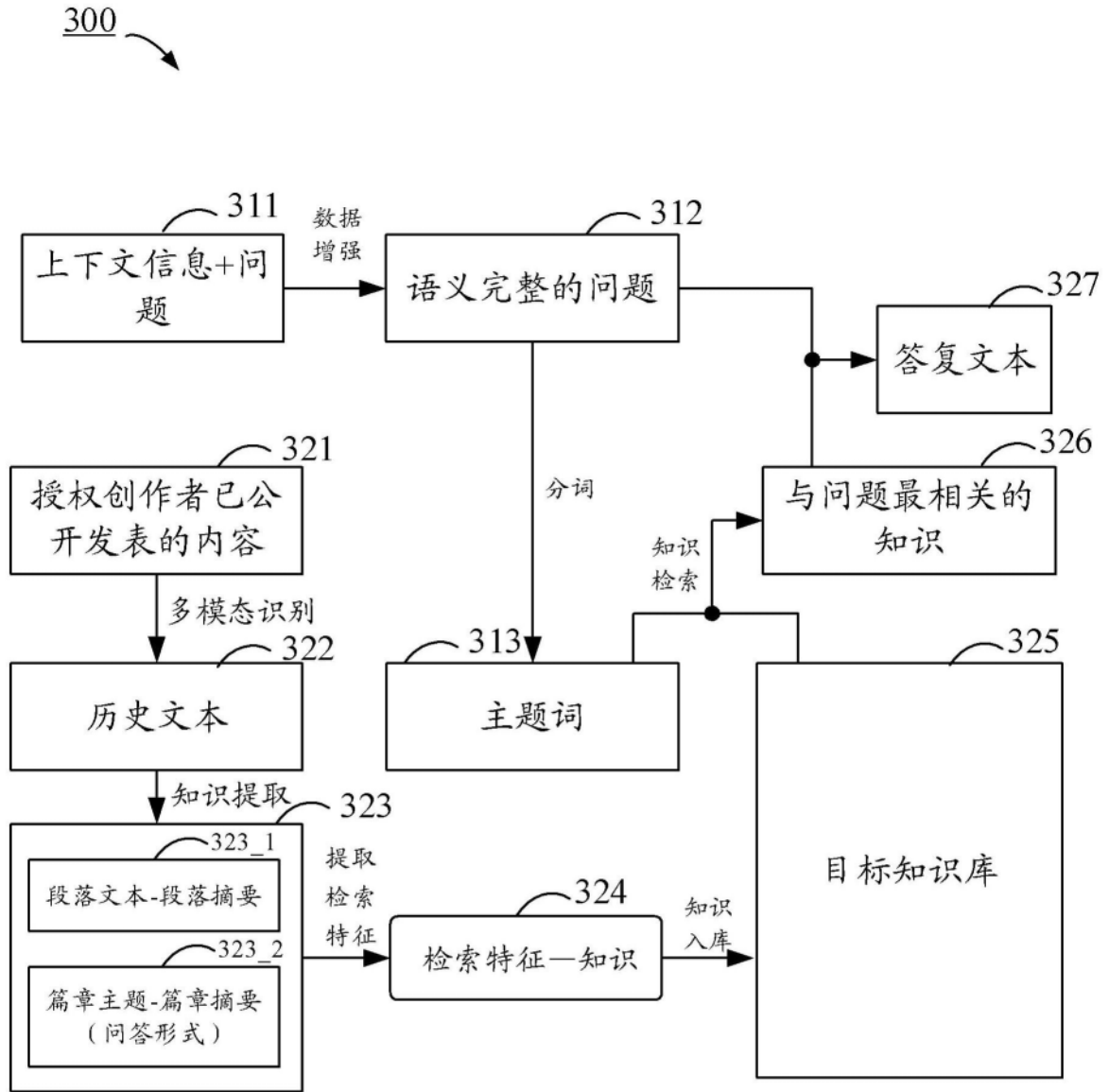


图3

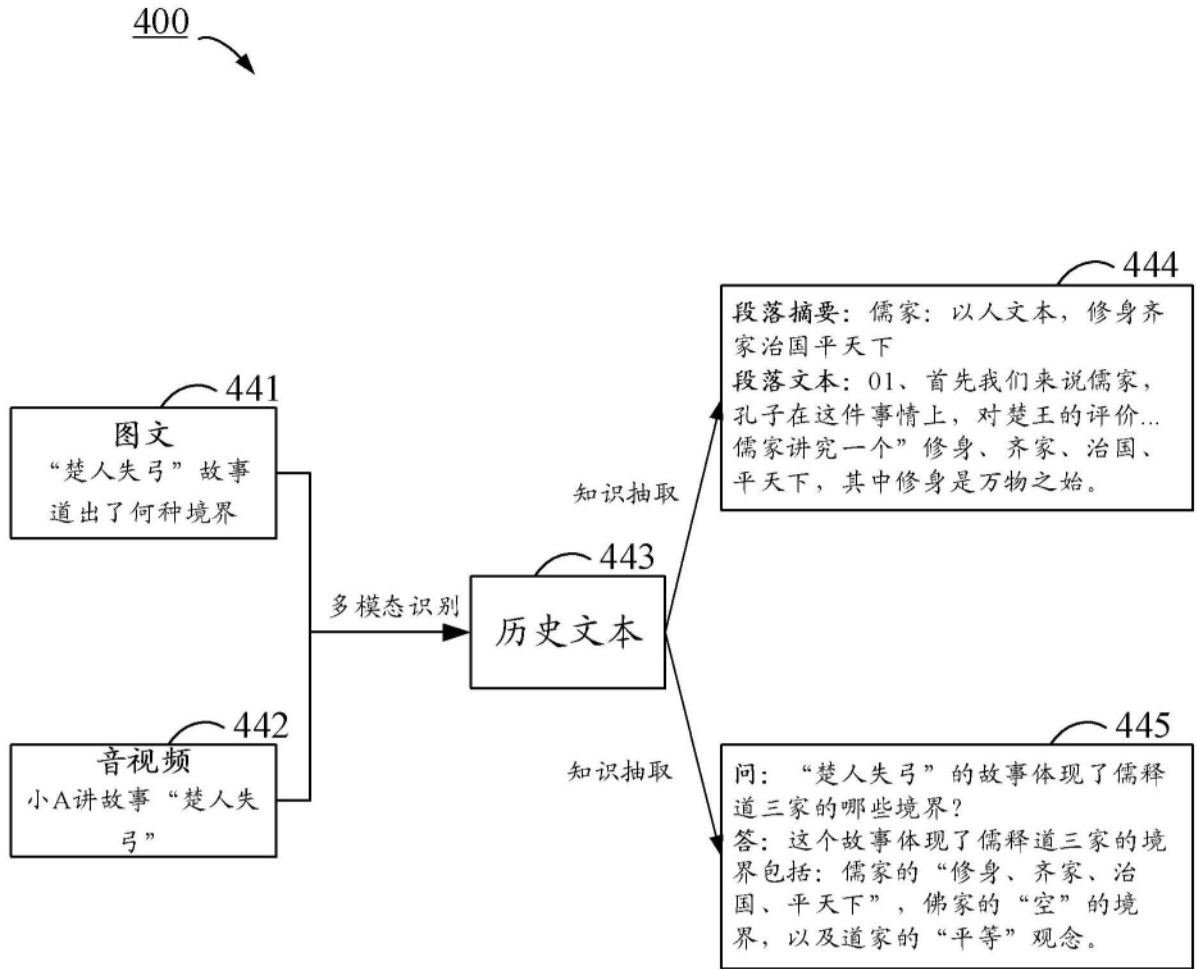


图4

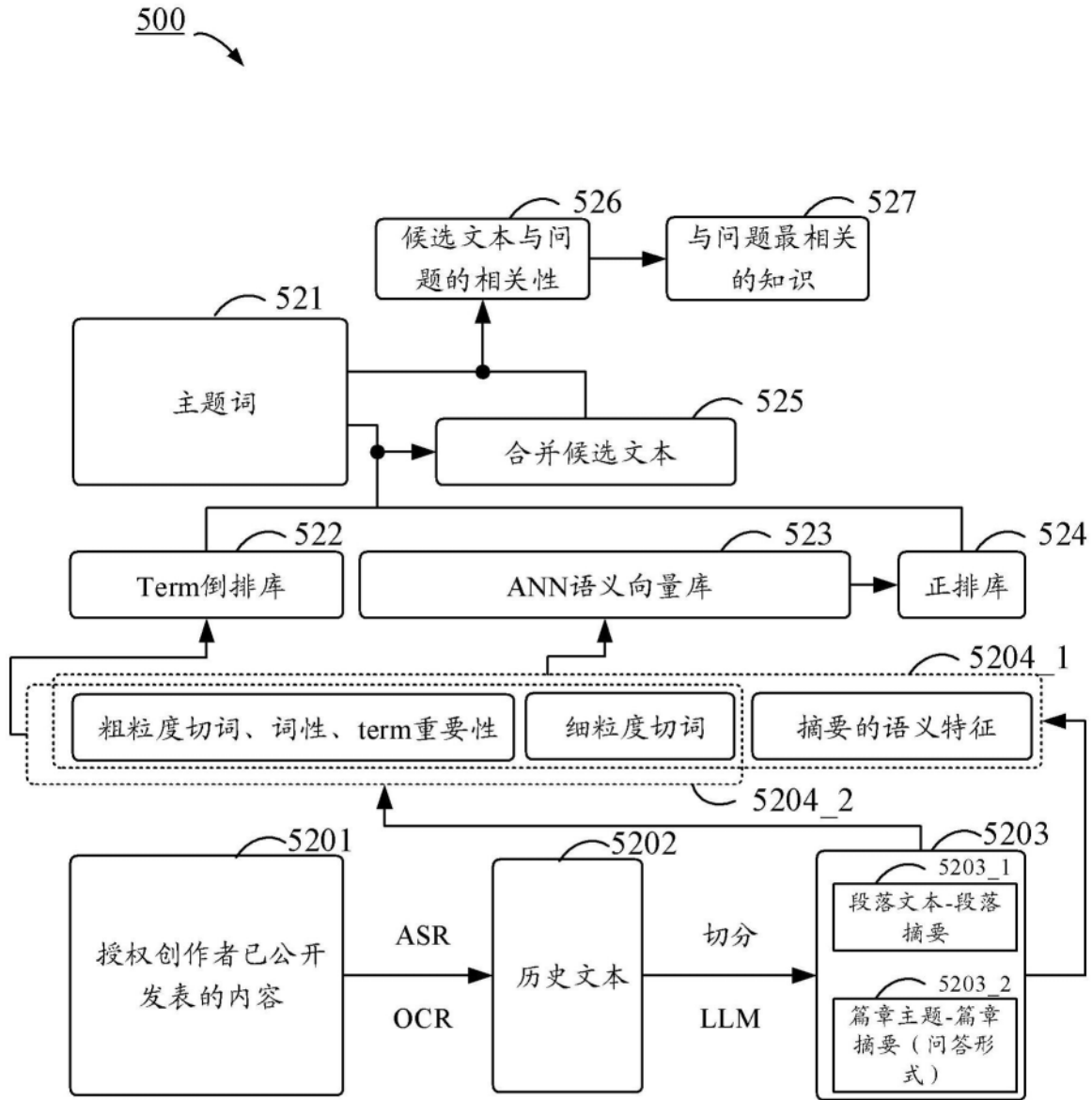


图5

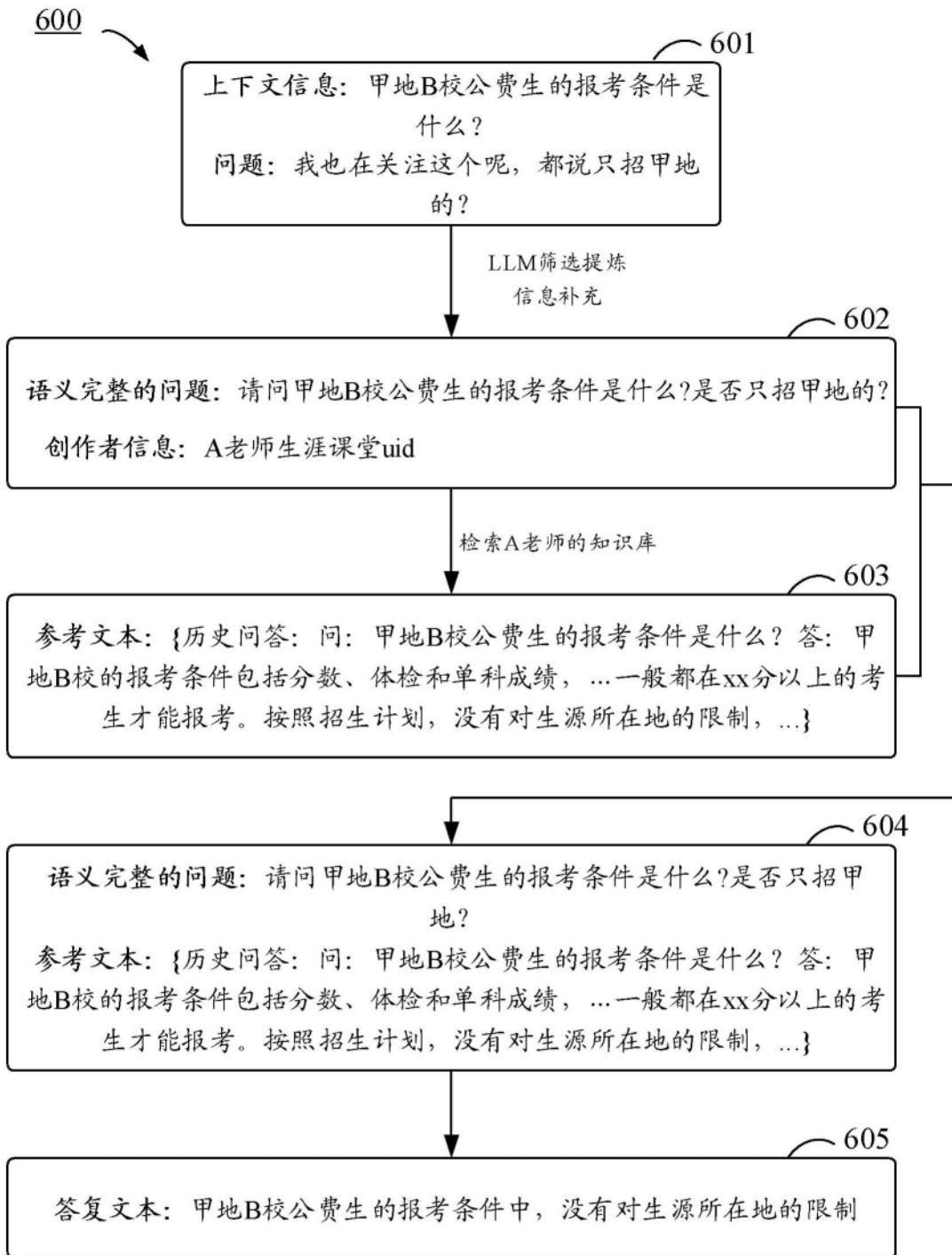


图6

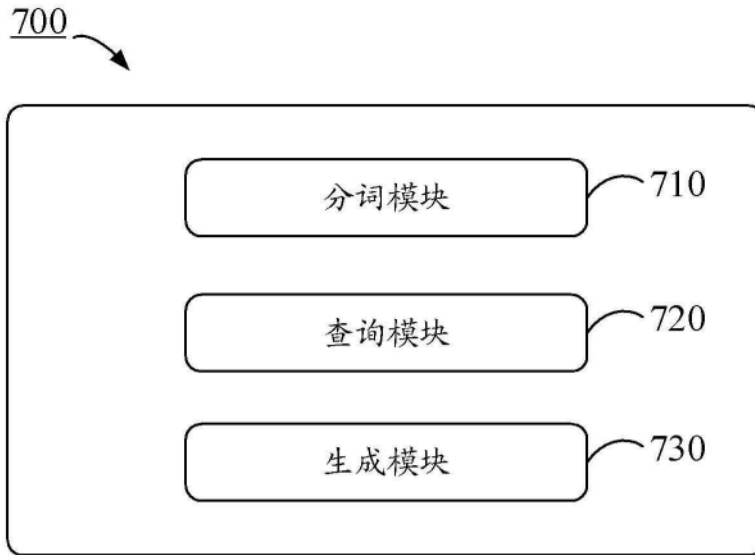


图7

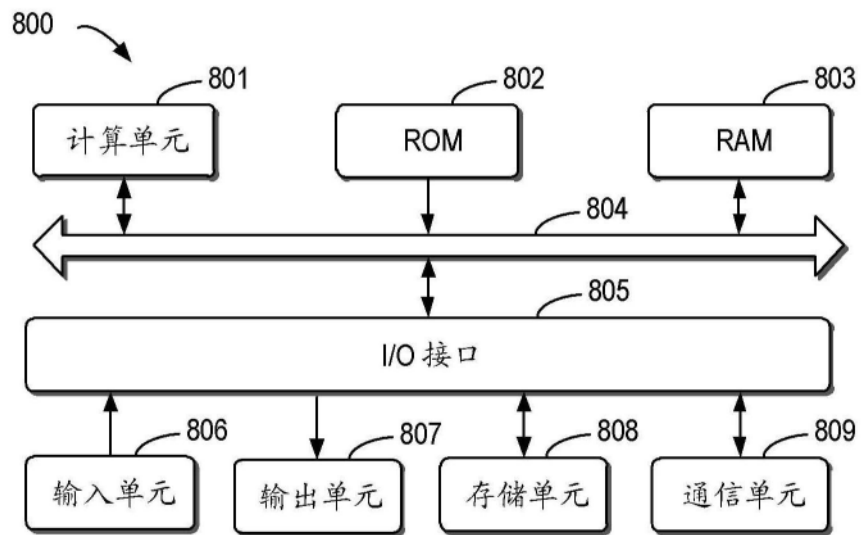


图8