



(12) 发明专利

(10) 授权公告号 CN 111046158 B

(45) 授权公告日 2020.12.15

(21) 申请号 201911285731.4

(22) 申请日 2019.12.13

(65) 同一申请的已公布的文献号
申请公布号 CN 111046158 A

(43) 申请公布日 2020.04.21

(73) 专利权人 腾讯科技(深圳)有限公司
地址 518000 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72) 发明人 闫昭 张士卫

(74) 专利代理机构 北京派特恩知识产权代理有
限公司 11270

代理人 王姗姗 张颖玲

(51) Int. Cl.

G06F 16/332 (2019.01)

G06F 16/35 (2019.01)

(56) 对比文件

CN 109816111 A, 2019.05.28

CN 110059152 A, 2019.07.26

CN 110442675 A, 2019.11.12

CN 110209782 A, 2019.09.06

CN 110222152 A, 2019.09.10

审查员 李佳玲

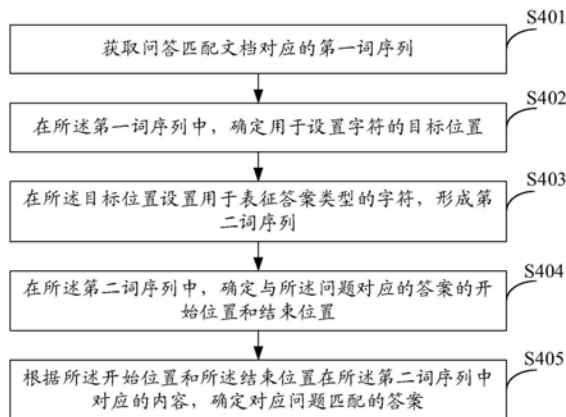
权利要求书3页 说明书22页 附图12页

(54) 发明名称

问答匹配方法及模型训练方法、装置、设备、
存储介质

(57) 摘要

本申请实施例提供一种问答匹配方法及模型训练方法、装置、设备、存储介质,其中,方法包括:获取问答匹配文档对应的第一词序列,所述问答匹配文档包括待解答的问题和用于匹配所述问题答案的文本,所述第一词序列中包括至少两个词,每一词包括一个或多个字;在所述第一词序列中,确定用于设置字符的目标位置;在所述目标位置设置用于表征答案类型的字符,形成第二词序列;在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置;根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。通过本申请,能够支持多答案类型的机器阅读理解任务,从而满足用户对多答案类型的需求,提高用户体验。



1. 一种问答匹配方法,其特征在于,包括:

获取问答匹配文档对应的第一词序列,所述问答匹配文档包括待解答的问题和用于匹配所述问题答案的文本,所述第一词序列中包括至少两个词,每一词包括一个或多个字;

在所述第一词序列中,确定用于设置字符的目标位置;

在所述目标位置设置用于表征答案类型的字符,形成第二词序列;

在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置;

当所述开始位置和所述结束位置均指向所述字符,且所述开始位置和所述结束位置之间满足预设偏移量条件时,根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。

2. 根据权利要求1所述的方法,其特征在于,所述获取问答匹配文档对应的第一词序列,包括:

对所述问题和所述文本依次进行分词处理,得到包括至少一个问题词和至少一个文本词的所述第一词序列,其中所述文本词位于所述问题词之后。

3. 根据权利要求2所述的方法,其特征在于,所述在所述第一词序列中,确定用于设置字符的目标位置,包括:

在所述第一词序列中,将位于所述问题词之前的位置,确定为所述目标位置;或者,

在所述第一词序列中,将位于所述问题词与所述文本词之间的位置,确定为所述目标位置;或者,

在所述第一词序列中,将位于所述文本词之后的位置,确定为所述目标位置。

4. 根据权利要求1所述的方法,其特征在于,所述在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置,包括:

对所述第二词序列中的词和字符分别进行特征提取,对应得到词特征向量和字符特征向量;

根据所述词特征向量和所述字符特征向量,确定所述第二词序列中的词和字符分别作为答案的开始位置和结束位置的概率值;

根据所述概率值,确定所述答案的开始位置和结束位置。

5. 根据权利要求4所述的方法,其特征在于,所述根据所述词特征向量和所述字符特征向量,确定所述第二词序列中的词和字符分别作为答案的开始位置和结束位置的概率值,包括:

分别对所述词特征向量和字符特征向量依次进行线性变换和归一化处理,得到所述第二词序列中的词和字符分别作为所述开始位置和所述结束位置对应的概率值集合;

所述概率值集合中包括至少两个所述概率值,所述概率值是用于表征所述答案的开始位置或结束位置的分布概率。

6. 根据权利要求5所述的方法,其特征在于,所述分别对所述词特征向量和字符特征向量依次进行线性变换和归一化处理,得到所述第二词序列中的词和字符分别作为所述开始位置和所述结束位置对应的概率值集合,包括:

分别对所述词特征向量和字符特征向量,依次进行第一线性变换和第一归一化处理,得到与所述第二词序列中的词和字符对应的第一概率值集合;

所述第一概率值集合中包括至少两个第一概率值,所述第一概率值是用于表征所述答

案的开始位置的分布概率；

分别对所述词特征向量和字符特征向量，依次进行第二线性变换和第二归一化处理，得到与所述第二词序列中的词和字符对应的第二概率值集合；

所述第二概率值集合中包括至少两个第二概率值，所述第二概率值是用于表征所述答案的结束位置的分布概率。

7. 根据权利要求6所述的方法，其特征在于，所述根据所述概率值，确定所述答案的开始位置和结束位置，包括：

将所述第一概率值集合中最大第一概率值对应的词或字符，确定为目标开始内容；将所述目标开始内容在所述第二词序列中所处的位置，确定为所述开始位置；

将所述第二概率值集合中最大第二概率值对应的词或字符，确定为目标结束内容；将所述目标结束内容在所述第二词序列中所处的位置，确定为所述结束位置。

8. 根据权利要求1至7任一项所述的方法，其特征在于，所述方法还包括：

当所述开始位置和所述结束位置均指向所述文本对应的文本词时，将所述开始位置所指向的第一文本词与所述结束位置所指向的第二文本词之间的内容，确定为所述答案；或者，

所述字符对应的答案类型至少包括以下之一：无答案类型、答案为是类型和答案为否类型；对应地，所述方法还包括：当所述开始位置和所述结束位置均指向所述字符时，根据所述字符对应的所述答案类型确定所述答案。

9. 根据权利要求1至7任一项所述的方法，其特征在于，所述方法还包括：

当所述开始位置和所述结束位置均指向所述字符，且所述开始位置位于所述结束位置之后时，对所述开始位置与所述结束位置进行反置处理，以互换所述开始位置和所述结束位置，得到新开始位置和新结束位置；

根据所述新开始位置和所述新结束位置在所述第二词序列中对应的内容，确定对应问题匹配的答案。

10. 根据权利要求1至7任一项所述的方法，其特征在于，所述方法还包括：

采用问答匹配网络模型，对所述第一词序列设置所述字符，得到第二词序列，并在所述第二词序列中，确定与所述问题对应的答案的开始位置和结束位置。

11. 根据权利要求10所述的方法，其特征在于，所述问答匹配网络模型通过以下步骤进行训练：

将样本数据中的待解答的问题、用于匹配所述问题答案的文本和字符，输入至特征提取网络模型中，对应得到与所述问题和所述文本对应的词特征向量、和与所述字符对应的字符特征向量；

将所述词特征向量和所述字符特征向量分别输入至答案位置预测网络模型中，得到与所述问题对应的答案的开始位置和结束位置；

将所述开始位置和所述结束位置输入至预设损失模型中，得到损失结果；

根据所述损失结果，对所述答案位置预测网络模型进行修正，得到所述问答匹配网络模型。

12. 一种问答匹配装置，其特征在于，包括：

获取模块，用于获取问答匹配文档对应的第一词序列，所述问答匹配文档包括待解答

的问题和用于匹配所述问题答案的文本,所述第一词序列中包括至少两个词,每一词包括一个或多个字;

第一确定模块,用于在所述第一词序列中,确定用于设置字符的目标位置;

设置模块,用于在所述目标位置设置用于表征答案类型的字符,形成第二词序列;

第二确定模块,用于在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置;

第三确定模块,用于当所述开始位置和所述结束位置均指向所述字符,且所述开始位置和所述结束位置之间满足预设偏移量条件时,根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。

13. 一种问答匹配设备,其特征在于,包括:

存储器,用于存储可执行指令;处理器,用于执行所述存储器中存储的可执行指令时,实现权利要求1至11任一项所述的方法。

14. 一种存储介质,其特征在于,存储有可执行指令,用于引起处理器执行时,实现权利要求1至11任一项所述的方法。

问答匹配方法及模型训练方法、装置、设备、存储介质

技术领域

[0001] 本申请实施例涉及人工智能技术领域,涉及但不限于一种问答匹配方法及模型训练方法、装置、设备、存储介质。

背景技术

[0002] 随着人工智能技术的发展,机器阅读理解技术也得到了大量的应用,例如,在网页搜索、问答机器人和智能语音助手等多个应用场景下,都会使用到机器阅读理解技术。通常情况下,机器阅读理解时的答案输出包括多种类型,例如,片段型、是否型和无答案类型。

[0003] 目前的机器阅读理解技术通常为抽取式机器阅读理解技术,即在用于匹配问题答案的文本中抽取一部分内容作为问题的答案。目前的机器阅读理解技术仅能支持文本片段型这一种答案类型作为输出,无法应对真实场景下用户对多答案类型的需求,用户体验较差。

发明内容

[0004] 本申请实施例提供一种问答匹配方法及模型训练方法、装置、设备、存储介质,能够支持多答案类型的机器阅读理解任务,从而满足用户对多答案类型的需求,提高用户体验。

[0005] 本申请实施例的技术方案是这样实现的:

[0006] 本申请实施例提供一种问答匹配方法,包括:

[0007] 获取问答匹配文档对应的第一词序列,所述问答匹配文档包括待解答的问题和用于匹配所述问题答案的文本,所述第一词序列中包括至少两个词,每一词包括一个或多个字;

[0008] 在所述第一词序列中,确定用于设置字符的目标位置;

[0009] 在所述目标位置设置用于表征答案类型的字符,形成第二词序列;

[0010] 在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置;

[0011] 根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。

[0012] 本申请实施例提供一种问答匹配网络模型训练方法,包括:

[0013] 将样本数据中的待解答的问题、用于匹配所述问题答案的文本和字符,输入至特征提取网络模型中,对应得到与所述问题和所述文本对应的词特征向量、和与所述字符对应的字符特征向量;

[0014] 将所述词特征向量和所述字符特征向量分别输入至答案位置预测网络模型中,得到与所述问题对应的答案的开始位置和结束位置;

[0015] 将所述开始位置和所述结束位置输入至预设损失模型中,得到损失结果;

[0016] 根据所述损失结果,对所述答案位置预测网络模型进行修正,得到所述问答匹配网络模型。

- [0017] 本申请实施例提供一种问答匹配装置,包括:
- [0018] 获取模块,用于获取问答匹配文档对应的第一词序列,所述问答匹配文档包括待解答的问题和用于匹配所述问题答案的文本,所述第一词序列中包括至少两个词,每一词包括一个或多个字;
- [0019] 第一确定模块,用于在所述第一词序列中,确定用于设置字符的目标位置;
- [0020] 设置模块,用于在所述目标位置设置用于表征答案类型的字符,形成第二词序列;
- [0021] 第二确定模块,用于在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置;
- [0022] 第三确定模块,用于根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。
- [0023] 本申请实施例提供一种问答匹配设备,包括:
- [0024] 存储器,用于存储可执行指令;处理器,用于执行所述存储器中存储的可执行指令时,实现上述的方法。
- [0025] 本申请实施例提供一种存储介质,存储有可执行指令,用于引起处理器执行时,实现上述的方法。
- [0026] 本申请实施例具有以下有益效果:由于在待解答的问题和用于匹配所述问题答案的文本所形成的第一词序列中确定设置字符的目标位置,并在目标位置设置用于表征答案类型的字符,形成第二词序列,并通过第二词序列确定答案的开始位置和结束位置。如此,将多种类型的答案生成方式统一成了预测答案的开始位置和结束位置的形式,从而能够支持多答案类型的机器阅读理解任务,满足用户对多答案类型的需求,提高用户体验。

附图说明

- [0027] 图1A是相关技术中的问答匹配方法的实现流程示意图;
- [0028] 图1B是本申请实施例提供的问答匹配系统的一个可选的架构示意图;
- [0029] 图2A是本申请实施例提供的问答匹配系统应用于区块链系统的一个可选的结构示意图;
- [0030] 图2B是本申请实施例提供的区块结构的一个可选的示意图;
- [0031] 图3是本申请实施例提供的服务器的结构示意图;
- [0032] 图4是本申请实施例提供的问答匹配方法的一个可选的流程示意图;
- [0033] 图5是本申请实施例提供的问答匹配方法的一个可选的流程示意图;
- [0034] 图6A是本申请实施例提供的问答匹配方法的一个可选的流程示意图;
- [0035] 图6B是本申请实施例提供的第二词序列的结构示意图;
- [0036] 图7是本申请实施例提供的问答匹配方法的一个可选的流程示意图;
- [0037] 图8A至8G是本申请实施例中根据开始位置和结束位置确定答案的实现过程示意图;
- [0038] 图9是本申请实施例提供的问答匹配网络模型的训练方法的流程示意图;
- [0039] 图10A是本申请实施例提供的直接搜索产品的界面图;
- [0040] 图10B是本申请实施例提供的搜索结果增强产品的界面图;
- [0041] 图10C是本申请实施例提供的智能助手的界面图;

- [0042] 图10D是本申请实施例提供的智能屏的界面图；
- [0043] 图11是本申请实施例提供的第一类抽取式机器阅读理解方法的实现过程示意图；
- [0044] 图12是本申请实施例提供的第二类抽取式机器阅读理解方法的实现过程示意图；
- [0045] 图13是本申请实施例提供的特殊字符添加位置示意图；
- [0046] 图14是本申请实施例提供的根据开始位置和结束位置确定答案的示意图；
- [0047] 图15是本申请实施例提供的设置位移的示意图；
- [0048] 图16是本申请实施例提供的设置反置的示意图。

具体实施方式

[0049] 为了使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请作进一步地详细描述，所描述的实施例不应视为对本申请的限制，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它实施例，都属于本申请保护的范围。

[0050] 在以下的描述中，涉及到“一些实施例”，其描述了所有可能实施例的子集，但是可以理解，“一些实施例”可以是所有可能实施例的相同子集或不同子集，并且可以在不冲突的情况下相互结合。除非另有定义，本申请实施例所使用的所有的技术和科学术语与属于本申请实施例的技术领域的技术人员通常理解的含义相同。本申请实施例所使用的术语只是为了描述本申请实施例的目的，不是旨在限制本申请。

[0051] 对本申请实施例进行进一步详细说明之前，对本申请实施例中涉及的名词和术语进行说明，本申请实施例中涉及的名词和术语适用于如下的解释。

[0052] 1) 机器阅读理解技术：一种将文本问题和相关文档作为输入，将文本答案作为输出的自动问答技术。

[0053] 2) 抽取式机器阅读理解技术：一种将给定文本问题和相关文档作为输入，从相关文档中选取一段连续的文本作为答案输出的机器阅读理解技术。

[0054] 3) 利用海量文本的语言模型训练方法 (BERT, Bidirectional Encoder Representations from Transformers)：该方法被广泛用于多种自然语言处理任务，如文本分类、文本匹配、机器阅读理解等。

[0055] 4) 神经网络：简称神经网络 (Neural Network, NN)，在机器学习和认知科学领域，是一种模仿生物神经网络结构和功能的数学模型或计算模型，用于对函数进行估计或近似。

[0056] 5) 模型参数：是使用通用变量来建立函数和变量之间关系的一个数量。在人工神经网络中，模型参数通常是实数矩阵。

[0057] 6) 应用程序接口 (API, Application Programming Interface)：是一些预先定义的函数，或指软件系统不同组成部分衔接的约定。目的是提供应用程序与开发人员基于某软件或硬件得以访问一组例程的能力，而又无需访问原码，或理解内部工作机制的细节。

[0058] 7) 软件开发工具包 (SDK, Software Development Kit)：是为特定的软件包、软件框架、硬件平台、操作系统等建立应用软件时的开发工具的集合广义上包括辅助开发某一类软件的相关文档、范例和工具的集合。

[0059] 为了更好地理解本申请实施例中提供的问答匹配方法和问答匹配网络模型训练方法，首先对相关技术中的问答匹配方法进行说明：

[0060] 机器阅读理解技术是一种将文本问题和相关文档作为输入,将文本答案作为输出的自动问答技术。下面给出一个阅读理解以及多种答案类型的样例:

[0061] 相关文档(输入1):“北京故宫是中国明清两代的皇家宫殿,旧称紫禁城,位于北京中轴线的中心,是中国古代宫廷建筑之精华。北京故宫以三大殿为中心,占地面积72万平方米,建筑面积约15万平方米,有大小宫殿七十多座,房屋九千余间。是世界上现存规模最大、保存最为完整的木质结构古建筑之一”。

[0062] 表1 文本问题和输出答案

[0063]	文本问题(输入2)	答案(输出)	答案类型
	故宫的旧称是什么	紫禁城	片段型
[0064]	故宫是最大的木质古建筑之一吗?	是	是否型
	故宫是哪年开始修建的?	不知道	无答案

[0065] 在上述样例中,相关文档和文本问题分别作为输入,答案作为输出。答案的类型包括片段型、是否型和无答案类型,其中,片段型是指由相关文档中的一个连续片段作为问题的答案来回答问题;是否型是指使用是或否来作为问题的答案来回答问题;无答案是指根据给定的相关文档无法回答问题。

[0066] 相关技术中的机器阅读理解技术通常为抽取式机器阅读理解技术,抽取式机器阅读理解技术作为近年来自动问答系统的热点研究方向,可以从给定的文档中提取一个连续的文字片段作为答案回复用户提问。这种技术可以更好地服务于网页搜索、问答机器人、智能语音助手等多个应用场景。

[0067] 相关技术中的抽取式机器阅读理解方法可以大致分为两大类型,1)表示学习+答案起止位置(即开始位置和结束位置)预测;2)预训练语言模型+答案起止位置预测。其中,预测答案的起止位置是抽取式机器阅读理解技术能够生成答案的核心。那么,对于相关技术中的抽取式机器阅读理解方法就存在一下两方面:

[0068] 第一方面,相关技术中的方法均基于一个假设,即给定的文档中一定可以找到相关片段回答用户提问。但是,上述假设在真实场景中通常并不成立。因为,相关文档可能并不能回答用户问题,或者,相关文档包含能够回答问题的信息,但是答案需要用的“是”、“否”等字眼并不出现在相关文档中,无法通过抽取答案片段进行回答。

[0069] 第二方面,相关技术中的方法应对待回答的问题时,主要采用对多种答案类型进行分类的方法。对于片段型与无答案型两种答案类型混合的情况,通常先预测片断答案,然后使用一个分类器对所抽取的片断是否为最终答案进行判断和验证。对于更复杂的片段型、无答案型、是否型混合的多答案类型混合的场景,现有方法可以使用多个分类器对输入进行判断。如图1A所示,是相关技术中的问答匹配方法的实现流程示意图,所述方法包括以下步骤:

[0070] 步骤S101,判断待回答的问题是否可回答。

[0071] 步骤S102,对于可回答的问题,判断答案是片断型还是“是否”型。

[0072] 如果答案是“是否”型,则需要对两种情况进行分类,将答案分成是或否。

[0073] 但是,使用分类器进行分类的方法,通常需要将答案抽取与答案类型判断划分成不同的模块,然后采用级联的方式融合成最终系统,如图1A中判断问题是否可回答、判断答案是片段型还是是否型、以及判断答案是“是”还是“否”这种级联的方式。这种级联的方式通常会产生错误的传导,进而影响最终系统的精准度。继续以图1A为例,如果在判断“可回答”和“不可回答”问题时产生错误,则后续的所有模块都将错误。

[0074] 另外,相关技术中使用分类器进行分类的方法,还需要分别得到不同的答案生成模块,生成答案的形式不统一。采用相关技术方案的系统,片段型答案需要输出答案的起止位置,而是否型和无答案型答案都是输出分类结果。这两种结果的形式是不同的,没有统一。这种不统一会导致训练模型时,需要不同的数据对不同类型输出结果的模块进行单独训练,增加了训练的复杂度和训练数据组织的复杂性。

[0075] 基于相关技术中所存在的上述至少一个问题,本申请实施例提供一种问答匹配方法和问答匹配网络模型训练方法,可以在统一的框架下回答片段型、无答案型、是否型等多答案类型混合的问题。在待解答的问题和用于匹配所述问题答案的文本所形成的第一词序列中确定设置字符的目标位置,并在目标位置设置用于表征答案类型的字符,形成第二词序列,并通过第二词序列所确定的答案的开始位置和结束位置。如此,将多种类型的答案生成方式统一成了预测答案的开始位置和结束位置的形式,从而能够支持多答案类型的机器阅读理解任务,满足用户对多答案类型的需求,提高用户体验。

[0076] 另外,本申请实施例提供的方案还涉及人工智能的网络模型训练技术,比如,可以通过人工智能技术实现训练用于对第一词序列设置字符,得到第二词序列,并在第二词序列中,确定与问题对应的答案的开始位置和结束位置的问答匹配网络模型,将在下文进行说明。

[0077] 这里,需要说明的是,人工智能是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说,人工智能是计算机科学的一个综合技术,它企图了解智能的实质,并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能器的设计原理与实现方法,使机器具有感知、推理与决策的功能。

[0078] 人工智能技术是一门综合学科,涉及领域广泛,既有硬件层面的技术也有软件层面的技术。人工智能基础技术一般包括如传感器、专用人工智能芯片、云计算、分布式存储、大数据处理技术、操作/交互系统、机电一体化等技术。人工智能软件技术主要包括计算机视觉技术、语音处理技术、自然语言处理技术以及机器学习/深度学习等几大方向。

[0079] 下面说明本申请实施例提供的问答匹配设备和问答匹配网络模型训练设备的示例性应用,本申请实施例提供的问答匹配设备和问答匹配网络模型训练设备均可以实施为笔记本电脑,平板电脑,台式计算机,移动设备(例如,移动电话,便携式音乐播放器,个人数字助理,专用消息设备,便携式游戏设备)等各种类型的终端,也可以实施为服务器。下面,将说明问答匹配设备和问答匹配网络模型训练设备实施为服务器时的示例性应用。

[0080] 参见图1B,图1B是本申请实施例提供的问答匹配系统10的一个可选的架构示意图。为实现支撑一个问答匹配应用(例如智能语音助手),问答匹配系统10包括终端(示例性

示出了终端100-1和终端100-2)通过网络200连接所述问答匹配应用的客户端对应的服务器300,终端获取问题,并通过网络200向服务器300发送问题,服务器300搜索用于匹配所述问题答案的文本,并对问题和文本进行分词处理,得到第一词序列;在所述第一词序列中,确定用于设置字符的目标位置;在所述目标位置设置用于表征答案类型的字符,形成第二词序列;在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置;根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案,最后,将答案通过网络200发送给终端。终端在接收到所述答案后,可以在当前界面(示例性示出了当前界面110-1和当前界面110-2)上显示所述答案和所述问题。本申请实施例中,网络200可以是广域网或者局域网,又或者是二者的组合。

[0081] 本申请实施例涉及的问答匹配系统10也可以是区块链系统的分布式系统101,参见图2A,图2A是本申请实施例提供的问答匹配系统10应用于区块链系统的一个可选的结构示意图,其中,所述分布式系统101可以由多个节点102(接入网络中的任意形式的计算设备,如服务器、用户终端)和客户端103形成的分布式节点,节点之间形成组成的点对点(P2P,Peer To Peer)网络,P2P协议是一个运行在传输控制协议(TCP,Transmission Control Protocol)协议之上的应用层协议。在分布式系统中,任何机器如服务器、终端都可以加入而成为节点,节点包括硬件层、中间层、操作系统层和应用层。

[0082] 参见图2A示出的区块链系统中各节点的功能,涉及的功能包括:

[0083] 1) 路由,节点具有的基本功能,用于支持节点之间的通信。

[0084] 节点除具有路由功能外,还可以具有以下功能:

[0085] 2) 应用,用于部署在区块链中,根据实际业务需求而实现特定业务,记录实现功能相关的数据形成记录数据,在记录数据中携带数字签名以表示任务数据的来源,将记录数据发送到区块链系统中的其他节点,供其他节点在验证记录数据来源以及完整性成功时,将记录数据添加到临时区块中。

[0086] 例如,应用实现的业务包括:

[0087] 2.1) 钱包,用于提供进行电子货币的的功能,包括发起交易(即,将当前交易的交易记录发送给区块链系统中的其他节点,其他节点验证成功后,作为承认交易有效的响应,将交易的记录数据存入区块链的临时区块中;当然,钱包还支持查询电子货币地址中剩余的电子货币。

[0088] 2.2) 共享账本,用于提供账目数据的存储、查询和修改等操作的功能,将对账目数据的操作的记录数据发送到区块链系统中的其他节点,其他节点验证有效后,作为承认账目数据有效的响应,将记录数据存入临时区块中,还可以向发起操作的节点发送确认。

[0089] 2.3) 智能合约,计算机化的协议,可以执行某个合约的条款,通过部署在共享账本上的用于在满足一定条件时而执行的代码实现,根据实际的业务需求代码用于完成自动化的交易,例如查询买家所购买商品的物流状态,在买家签收货物后将买家的电子货币转移到商户的地址;当然,智能合约不仅限于执行用于交易的合约,还可以执行对接收的信息进行处理的合约。

[0090] 3) 区块链,包括一系列按照产生的先后时间顺序相互接续的区块(Block),新区块一旦加入到区块链中就不会再被移除,区块中记录了区块链系统中节点提交的记录数据。

[0091] 4) 共识(Consensus),是区块链网络中的一个过程,用于在涉及的多个节点之间对

区块中的交易达成一致,达成一致的区块将被追加到区块链的尾部,实现共识的机制包括工作量证明(PoW,Proof of Work)、权益证明(PoS,Proof of Stake)、股份授权证明(DPoS, Delegated Proof-of-Stake)、消逝时间量证明(PoET,Proof of Elapsed Time)等。

[0092] 参见图2B,图2B是本申请实施例提供的区块结构(Block Structure)的一个可选的示意图,每个区块中包括本区块存储交易记录的哈希值(本区块的哈希值)、以及前一区块的哈希值,各区块通过哈希值连接形成区块链。另外,区块中还可以包括有区块生成时的时间戳等信息。区块链(Blockchain),本质上是一个去中心化的数据库,是一串使用密码学方法相关联产生的数据块,每一个数据块中包含了相关的信息,用于验证其信息的有效性(防伪)和生成下一个区块。

[0093] 参见图3,图3是本申请实施例提供的服务器300的结构示意图,图3所示的服务器300包括:至少一个处理器310、存储器350、至少一个网络接口320和用户接口330。服务器300中的各个组件通过总线系统340耦合在一起。可理解,总线系统340用于实现这些组件之间的连接通信。总线系统340除包括数据总线之外,还包括电源总线、控制总线和状态信号总线。但是为了清楚说明起见,在图3中将各种总线都标为总线系统340。

[0094] 处理器310可以是一种集成电路芯片,具有信号的处理能力,例如通用处理器、数字信号处理器(DSP,Digital Signal Processor),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,其中,通用处理器可以是微处理器或者任何常规的处理器等。

[0095] 用户接口330包括使得能够呈现媒体内容的一个或多个输出装置331,包括一个或多个扬声器和/或一个或多个视觉显示屏。用户接口330还包括一个或多个输入装置332,包括有助于用户输入的用户接口部件,比如键盘、鼠标、麦克风、触屏显示屏、摄像头、其他输入按钮和控件。

[0096] 存储器350可以是可移除的,不可移除的或其组合。示例性的硬件设备包括固态存储器,硬盘驱动器,光盘驱动器等。存储器350可选地包括在物理位置上远离处理器310的一个或多个存储设备。存储器350包括易失性存储器或非易失性存储器,也可包括易失性和非易失性存储器两者。非易失性存储器可以是只读存储器(ROM,Read Only Memory),易失性存储器可以是随机存取存储器(RAM,Random Access Memory)。本申请实施例描述的存储器350旨在包括任意适合类型的存储器。在一些实施例中,存储器350能够存储数据以支持各种操作,这些数据的示例包括程序、模块和数据结构或者其子集或超集,下面示例性说明。

[0097] 操作系统351,包括用于处理各种基本系统服务和执行硬件相关任务的系统程序,例如框架层、核心库层、驱动层等,用于实现各种基础业务以及处理基于硬件的任务;

[0098] 网络通信模块352,用于经由一个或多个(有线或无线)网络接口320到达其他计算设备,示例性的网络接口320包括:蓝牙、无线相容性认证(WiFi)、和通用串行总线(USB, Universal Serial Bus)等;

[0099] 输入处理模块353,用于对一个或多个来自一个或多个输入装置332之一的一个或多个用户输入或互动进行检测以及翻译所检测的输入或互动。

[0100] 在一些实施例中,本申请实施例提供的装置可以采用软件方式实现,图3示出了存储在存储器350中的一种问答匹配装置354,该问答匹配装置354可以是服务器300中的问答匹配装置,其可以是程序和插件等形式的软件,包括以下软件模块:获取模块3541、第一确

定模块3542、设置模块3543、第二确定模块3544和第三确定模块3545,这些模块是逻辑上的,因此根据所实现的功能可以进行任意的组合或进一步拆分。将在下文中说明各个模块的功能。

[0101] 在另一些实施例中,本申请实施例提供的数据分流网络构建设备也可以实施为服务器,所述服务器包括图3中所示的各个组件。其中,本申请实施例提供的问答匹配网络模型训练装置也可以采用软件方式实现,该问答匹配网络模型训练装置可以是服务器中的问答匹配网络模型训练装置,其可以是程序和插件等形式的软件,包括以下软件模块(图中未示出):第一输入模块、第二输入模块、第三输入模块和修正模块,这些模块是逻辑上的,因此根据所实现的功能可以进行任意的组合或进一步拆分。

[0102] 在又一些实施例中,本申请实施例提供的装置可以采用硬件方式实现,作为示例,本申请实施例提供的装置可以是采用硬件译码处理器形式的处理器,其被编程以执行本申请实施例提供的问答匹配方法和问答匹配网络模型训练方法,例如,硬件译码处理器形式的处理器可以采用一个或多个应用专用集成电路(ASIC,Application Specific Integrated Circuit)、DSP、可编程逻辑器件(PLD,Programmable Logic Device)、复杂可编程逻辑器件(CPLD,Complex Programmable Logic Device)、现场可编程门阵列(FPGA,Field-Programmable Gate Array)或其他电子元件。

[0103] 下面将结合本申请实施例提供的服务器300的示例性应用和实施,说明本申请实施例提供的问答匹配方法。参见图4,图4是本申请实施例提供的问答匹配方法的一个可选的流程示意图,将结合图4示出的步骤进行说明。

[0104] 步骤S401,获取问答匹配文档对应的第一词序列。

[0105] 这里,所述问答匹配文档包括待解答的问题和用于匹配所述问题答案的文本,所述第一词序列中包括至少两个词,每一词包括一个或多个字。

[0106] 本申请实施例中,用户可以通过终端上的客户端输入问题,服务器根据所述问题匹配得到所述文本,或者用户可以通过终端上的客户端输入问题和用于匹配所述问题的文本。当服务器获取到问题和文本后,将问题和文本确定为所述问答匹配文档,并通过分词处理对所述问答匹配文档中的问题和文本依次进行处理,以获取所述第一词序列,其中所述分词处理是指对所述问答匹配文档中的问题和文本依次进行拆分,得到至少一个词,所述词包括一个或多个字,即每一个词可以是一个字,也可以是由多个字组成的词语。所述第一词序列可以包括至少两个词语,也可以包括至少两个字,也可以包括至少一个字和至少一个词语。

[0107] 步骤S402,在所述第一词序列中,确定用于设置字符的目标位置。

[0108] 这里,所述目标位置是所述第一词序列中的词之间的位置,所述目标位置用于设置字符,所述目标位置可以位于所述第一词序列中的任意位置,所述目标位置可以根据实际需要进行确定。例如,可以将第一词序列的序列开始位置确定为所述目标位置,或者将第一词序列的序列结束位置确定为所述目标位置。

[0109] 步骤S403,在目标位置设置用于表征答案类型的字符,形成第二词序列。

[0110] 所述字符用于表征答案类型,所答案类型可以根据实际问答匹配需求进行设置的任意一种答案类型。例如,所述答案类型可以包括无答案类型、答案为是类型和答案与否类型。

[0111] 在所述第二词序列中,可以设置至少一个用于表征答案类型的字符,在同一第二词序列中,如果设置有多个表征不同答案类型的字符,则字符的形式不同,也就是说,同一第二词序列中的不同答案类型对应不同的字符。

[0112] 当同一第二词序列中包括多个不同答案类型的字符时,多个字符可以位于第二词序列中的任意位置,多个字符可以分开设置于第二词序列中,也可以连续设置于第二词序列中,多个字符之间的先后顺序不做限定,可以以任意的排列顺序设置于第二词序列中。

[0113] 步骤S404,在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置。

[0114] 这里,所述开始位置和所述结束位置可以均指向文本对应的词的位置,或者指向所述字符对应的位置。所述开始位置位于所述结束位置之前,所述开始位置和所述结束位置也可以指向同一词或同一字符,所述开始位置也可以位于所述结束位置之后。

[0115] 当所述开始位置和所述结束位置中的任意一个指向问题对应的词的位置时,表明当前所确定的开始位置和结束位置错误,判定此次问答匹配过程无效。

[0116] 步骤S405,根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。

[0117] 这里,将开始位置和结束位置对应的内容确定为答案,至少对应以下场景:所述开始位置和结束位置均指向文本对应的词的位置,则可以将开始位置指向的词与结束位置指向的词之间的文本片段确定为答案;所述开始位置和结束位置均指向字符对应的位置,如果开始位置和结束位置指向同一字符,则将所指向的字符对应的答案类型确定为答案,如果开始位置和结束位置指向不同字符,则将开始位置指向的字符对应的答案类型确定为答案,或者将结束位置指向的字符对应的答案类型确定为答案。

[0118] 本申请实施例提供的问答匹配方法,由于在待解答的问题和用于匹配所述问题答案的文本所形成的第一词序列中确定设置字符的目标位置,并在目标位置设置用于表征答案类型的字符,形成第二词序列,并通过第二词序列确定答案的开始位置和结束位置。如此,将多种类型的答案生成方式统一成了预测答案的开始位置和结束位置的形式,从而能够支持多答案类型的机器阅读理解任务,满足用户对多答案类型的需求,提高用户体验。

[0119] 在一些实施例中,所述问答匹配系统中包括运行有问答匹配应用的客户端和用于实现所述问答匹配方法的服务器,图5是本申请实施例提供的问答匹配方法的一个可选的流程示意图,如图5所示,所述方法包括以下步骤:

[0120] 步骤S501,客户端获取待解答的问题和用于匹配所述问题答案的文本。

[0121] 这里,在用户终端上可以运行有问答匹配应用,所述客户端是所述问答匹配应用的客户端,用户可以通过客户端输入所述问题和所述文本,或者用户可以通过客户端输入所述问题,客户端服务器通过网络搜索出用于匹配所述问题答案的文本,或者客户端服务器从本地搜索用于匹配所述问题答案的文本。

[0122] 步骤S502,客户端将所述问题和所述文本发送给服务器。

[0123] 步骤S503,服务器对所述问题和所述文本依次进行分词处理,得到包括至少一个问题词和至少一个文本词的所述第一词序列。

[0124] 这里,所述问题词可以是一个字也可以是包括多个字的词语,所述文本词可以是一个字也可以是包括多个字的词语。在一些实施例中,所述文本词可以位于所述问题词之

后,当然,所述文本词也可以位于所述问题词之前,或者,所述文本词也可以位于所述问题词之间的任意位置。

[0125] 步骤S504,服务器在所述第一词序列中,将位于所述问题词之前的位置,确定为所述目标位置。

[0126] 步骤S505,服务器在所述第一词序列中,将位于所述问题词与所述文本词之间的位置,确定为所述目标位置。

[0127] 步骤S506,服务器在所述第一词序列中,将位于所述文本词之后的位置,确定为所述目标位置。

[0128] 在上述步骤S504至S506中,分别列举出了确定三种不同目标位置的方法。

[0129] 步骤S507,服务器在所述目标位置设置用于表征答案类型的字符,形成第二词序列。

[0130] 步骤S508,服务器在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置。

[0131] 步骤S509,服务器根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。

[0132] 步骤S510,服务器将所述答案发送给终端。

[0133] 步骤S511,终端在当前界面上显示所述答案。

[0134] 本申请实施例提供的问答匹配方法,服务器获取终端发送的问题和文本,并对问题和文本进行分词处理得到第一词序列,然后在第一词序列中确定设置字符的目标位置,并在目标位置设置用于表征答案类型的字符,形成第二词序列,并通过第二词序列确定答案的开始位置和结束位置。如此,将多种类型的答案生成方式统一成了预测答案的开始位置和结束位置的形式,从而能够支持多答案类型的机器阅读理解任务,满足用户对多答案类型的需求,提高用户体验。

[0135] 在一些实施例中,可以采用训练好的问答匹配网络模型实现本申请实施例的问答匹配方法,也就是说,采用问答匹配网络模型,对所述第一词序列设置所述字符,得到第二词序列,并在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置。其中,所述问答匹配网络模型包括特征提取网络模型和答案位置预测网络模型,所述特征提取网络模型用于对所述问题、所述文本和所述字符对应的第二词序列中的词和字符分别进行特征提取,所述答案位置预测网络模型用于预测所述答案的开始位置和结束位置。

[0136] 基于图4,如图6A所示,是本申请实施例提供的问答匹配方法的一个可选的流程示意图,上述步骤S404中确定开始位置和结束位置,可以通过以下步骤实现:

[0137] 步骤S601,对所述第二词序列中的词和字符分别进行特征提取,对应得到词特征向量和字符特征向量。

[0138] 这里,所述第二词序列中的词可以包括以词语的形式所呈现的词语,还可以包括以单个字的形式所呈现的字。

[0139] 本申请实施例中,可以通过特征提取网络模型对所述第二词序列中的词和字符分别进行特征提取,对应得到词特征向量和字符特征向量。在一些实施例中,可以首先通过查表得到相应词或字符对应的向量表示,然后,将每一词和每一字符的向量表示输入至特征提取网络模型中,通过特征提取网络模型中的特征提取层对每一向量表示进行处理,对应

得到每一词的词特征向量和每一字符的字符特征向量。

[0140] 图6B是本申请实施例提供的第二词序列的结构示意图,如图6B所示,在所述第二词序列中,包括至少一个词D1、D2……Dn和字符T1、T2……Tn。字符T1、T2……Tn可以位于所述第二词序列中的任意位置。

[0141] 步骤S602,根据所述词特征向量和所述字符特征向量,确定所述第二词序列中的词和字符分别作为答案的开始位置和结束位置的概率值。

[0142] 这里,可以通过答案位置预测网络模型确定所述第二词序列中的词和字符分别作为答案的开始位置和结束位置的概率值。

[0143] 在一些实施例中,所述答案位置预测网络模型包括线性变换处理层和归一化处理层,其中,所述线性变换处理层用于对对应的特征向量进行线性变换处理,所述归一化处理层用于对线性变化处理后的结果进行归一化处理。

[0144] 对应地,步骤S602还可以通过以下步骤实现:

[0145] 步骤S6021,分别对所述词特征向量和字符特征向量依次进行线性变换和归一化处理,得到所述第二词序列中的词和字符分别作为所述开始位置和所述结束位置对应的概率值集合。所述概率值集合中包括至少两个所述概率值,所述概率值是用于表征所述答案的开始位置或结束位置的分布概率。

[0146] 步骤S603,根据所述概率值,确定所述答案的开始位置和结束位置。

[0147] 在一些实施例中,通过所述答案位置预测网络模型确定出所述第二词序列中的每一词和每一字符分别作为答案的开始位置和结束位置的概率值,因此,需要对每一词和每一字符进行两次线性变换和归一化处理,一次用于确定每一词和每一字符作为答案的开始位置的概率值,另一次用于确定每一词和每一字符作为答案的结束位置的概率值。

[0148] 基于图6,如图7所示,是本申请实施例提供的问答匹配方法的一个可选的流程示意图,上述步骤S602还可以通过以下步骤实现:

[0149] 步骤S701,分别对所述词特征向量和字符特征向量,依次进行第一线性变换和第一归一化处理,得到与所述第二词序列中的词和字符对应的第一概率值集合。

[0150] 这里,所述第一概率值集合中包括至少两个第一概率值,所述第一概率值是用于表征所述答案的开始位置的分布概率,所述第一概率值集合中的全部第一概率值之和等于1。

[0151] 步骤S702,分别对所述词特征向量和字符特征向量,依次进行第二线性变换和第二归一化处理,得到与所述第二词序列中的词和字符对应的第二概率值集合。

[0152] 所述第二概率值集合中包括至少两个第二概率值,所述第二概率值是用于表征所述答案的结束位置的分布概率,所述第二概率值集合中的全部第一概率值之和等于1。

[0153] 对应地,步骤S603可以通过以下步骤实现:

[0154] 步骤S703,将所述第一概率值集合中最大第一概率值对应的词或字符,确定为目标开始内容。

[0155] 这里,在所述第一概率值集合中确定最大第一概率值,并将最大第一概率值对应的词或字符确定为目标开始内容,当最大第一概率值对应的是第二词序列中的词时,则将该词的内容作为目标开始内容,当最大第一概率值对应的是第二词序列中的字符时,则将该字符对应的答案类型作为目标开始内容。

[0156] 步骤S704,将所述目标开始内容在所述第二词序列中所处的位置,确定为所述开始位置。

[0157] 步骤S705,将所述第二概率值集合中最大第二概率值对应的词或字符,确定为目标结束内容。

[0158] 这里,在所述第二概率值集合中确定最大第二概率值,并将最大第二概率值对应的词或字符确定为目标结束内容,当最大第二概率值对应的是第二词序列中的词时,则将该词的内容作为目标结束内容,当最大第二概率值对应的是第二词序列中的字符时,则将该字符对应的答案类型作为目标结束内容。

[0159] 步骤S706,将所述目标结束内容在所述第二词序列中所处的位置,确定为所述结束位置。

[0160] 本申请实施例中,分别确定第二词序列中的每一词和每一字符作为答案的开始位置和结束位置的概率值,然后根据概率值的大小,将作为答案开始位置的概率值中的最大概率值对应的词或字符作为目标开始内容,将作为答案结束位置的概率值中的最大概率值对应的词或字符作为目标结束内容,如此,根据概率值来确定答案的开始位置和结束位置,能够准确的确定出与问题匹配的答案的开始位置和结束位置。

[0161] 在一些实施例中,在确定出所述开始位置和所述结束位置之后,可以根据所述开始位置和所述结束位置在所述第二词序列中所指向的内容,确定与所述问题对应的答案,下面举例说明上述实施例中的步骤S405根据所述开始位置和所述结束位置确定答案的一些可选确定方式:

[0162] 方式一:所述开始位置和所述结束位置均指向所述文本对应的文本词。

[0163] 步骤S4051,当所述开始位置和所述结束位置均指向所述文本对应的文本词时,将所述开始位置所指向的第一文本词与所述结束位置所指向的第二文本词之间的内容,确定为所述答案。

[0164] 如图8A所示,是本申请实施例中根据开始位置和结束位置确定答案的一种可选实现过程示意图,其中,问题对应的问题词包括Q1、Q2和Q3,文本对应的文本词包括D1、D2和D3。

[0165] 当根据所述概率值确定出的开始位置指向D1,结束位置指向D2时,则将D1和D2两个词连接起来形成了文本片段作为问题的答案。

[0166] 方式二:所述开始位置和所述结束位置均指向所述字符。其中,所述字符对应的答案类型至少包括以下之一:无答案类型、答案为是类型和答案为否类型。

[0167] 步骤S4052,当所述开始位置和所述结束位置均指向所述字符时,根据所述字符对应的所述答案类型确定所述答案。

[0168] 如图8B所示,是本申请实施例中根据开始位置和结束位置确定答案的一种可选实现过程示意图,其中,问题对应的问题词包括Q1、Q2和Q3,文本对应的文本词包括D1、D2和D3,在问题词和文本词之间有无答案类型字符NA、答案为是字符YES和答案为否字符NO。

[0169] 当根据所述概率值确定出的开始位置指向YES字符,结束位置也指向YES字符时,则确定问题的答案为是;或者,在一些实施例中,当所述开始位置指向NO字符,结束位置也指向NO字符时(图中未示出),则确定问题的答案为否;或者,在一些实施例中,当所述开始位置指向NA字符,结束位置也指向NA字符时(图中未示出),则确定问题在该文本中无对应

答案。

[0170] 方式三:当所述开始位置和所述结束位置均指向所述字符时,表明所述问题的答案不是片段型的,因此,为所述开始位置和所述结束位置设置偏移量,所述偏移量是指所述结束位置与所述开始位置之间相隔的词的数量。

[0171] 步骤S4053,当所述开始位置和所述结束位置均指向所述字符,且所述开始位置和所述结束位置之间满足预设偏移量条件时,根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。

[0172] 如图8C所示,是本申请实施例中根据开始位置和结束位置确定答案的一种可选实现过程示意图,其中,问题对应的问题词包括Q1、Q2和Q3,文本对应的文本词包括D1、D2和D3,在问题词和文本词之间有无答案类型字符NA、答案为是字符YES和答案为否字符NO;在问题词Q3与无答案类型字符NA之间还可以添加分隔字符SEP,在答案为否字符NO与文本词D1之间也可以添加分隔字符SEP。

[0173] 本申请实施例中,可以预设开始位置与结束位置之间的预设偏移量条件,例如,预设偏移量条件可以是偏移量为1,那么,当所述开始位置指向NA字符,结束位置指向YES字符时,则确定问题在该文本中无对应答案;或者,在一些实施例中,当所述开始位置指向YES字符,结束位置指向NO字符时(图中未示出),则确定问题的答案为是;或者,在一些实施例中,当所述开始位置指向NO字符,结束位置指向SEP字符时(图中未示出),则确定问题的答案为否。

[0174] 方式四:当所述问题的答案不是片段型的,且结束位置位于开始位置之前时,则可以通过以下步骤确定问题的答案:

[0175] 步骤S4054,当所述开始位置和所述结束位置均指向所述字符,且所述开始位置位于所述结束位置之后时,对所述开始位置与所述结束位置进行反置处理,以互换所述开始位置和所述结束位置,得到新开始位置和新结束位置。

[0176] 步骤S4055,根据所述新开始位置和所述新结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。

[0177] 如图8D所示,是本申请实施例中根据开始位置和结束位置确定答案的一种可选实现过程示意图,其中,问题对应的问题词包括Q1、Q2和Q3,文本对应的文本词包括D1、D2和D3,在问题词和文本词之间有无答案类型字符NA、答案为是字符YES和答案为否字符NO;在问题词Q3与无答案类型字符NA之间还可以添加分隔字符SEP,在答案为否字符NO与文本词D1之间也可以添加分隔字符SEP。

[0178] 这里,如果根据所述概率值确定出的开始位置位于结束位置之后,例如,当所述开始位置指向YES字符,结束位置指向NA字符时,则确定问题在该文本中无对应答案;或者,在一些实施例中,当所述开始位置指向NO字符,结束位置指向YES字符时(图中未示出),则确定问题的答案为是;或者,在一些实施例中,当所述开始位置指向SEP字符,结束位置指向NO字符时(图中未示出),则确定问题的答案为否。

[0179] 在一些实施例中,还提供一种分析系统所确定出的开始位置和结束位置出错的方法,当开始位置和结束位置预测出错时,禁止根据开始位置和结束位置确定问题对应的答案,可以重新对问答匹配网络模型进行训练,或者,采用问答匹配网络模型重新计算概率值,并重新确定新的开始位置和新的结束位置。

[0180] 这里,提供开始位置和结束位置预测出错的几种情况:

[0181] 情况一:当所述开始位置和所述结束位置中至少一个指向所述问题对应的问题词时,表明开始位置和结束位置预测出错。

[0182] 如图8E所示,当所述开始位置指向所述问题对应的问题词Q1,结束位置指向文本词D1时,表明开始位置和结束位置预测出错。

[0183] 情况二:当所述开始位置和所述结束位置均指向所述字符,且所述开始位置和所述结束位置之间不满足预设偏移量条件时,表明开始位置和结束位置预测出错。

[0184] 如图8F所示,如果预设偏移量条件是1,当所述开始位置指向NA字符,结束位置指向N0字符时,表明开始位置和结束位置预测出错。

[0185] 情况三:当所述开始位置和所述结束位置均指向分隔字符SEP时,表明开始位置和结束位置预测出错。

[0186] 如图8G所示,当所述开始位置指向第一个SEP字符,结束位置指向第二个SEP字符时,表明开始位置和结束位置预测出错。

[0187] 当然,在本申请实施例中,开始位置和结束位置可以指向的内容具有多种组合方式,为了节约篇幅,本申请实施例不能一一列举,对于开始位置和结束位置的其他指向情况,可以根据实际情况决定是否能够确定出有效的答案,本申请实施例不再赘述。

[0188] 在一些实施例中,可以采用问答匹配网络模型,对所述第一词序列设置所述字符,得到第二词序列,并在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置,从而可以根据所述开始位置和所述结束位置确定所述问题的答案。

[0189] 这里,提供一种问答匹配网络模型的训练方法,如图9所示,是本申请实施例提供的问答匹配网络模型的训练方法的流程示意图,所述方法包括:

[0190] 步骤S901,将样本数据中的待解答的问题、用于匹配所述问题答案的文本和字符,输入至特征提取网络模型中,对应得到与所述问题和所述文本对应的词特征向量、和与所述字符对应的字符特征向量。

[0191] 这里,所述样本数据包括待解答的问题、用于匹配所述问题答案的文本和用于表征答案类型的至少一个字符。所述特征提取网络模型包括特征提取层,所述特征提取层用于对所述问题、所述文本和所述字符对应的至少两个词和字符分别进行特征提取,对应得到词特征向量和字符特征向量。

[0192] 步骤S902,将所述词特征向量和所述字符特征向量分别输入至答案位置预测网络模型中,得到与所述问题对应的答案的开始位置和结束位置。

[0193] 这里,所述答案位置预测网络模型用于预测所述答案的开始位置和结束位置,所述答案位置预测网络模型包括线性变换处理层和归一化处理层,其中,所述线性变换处理层用于对对应的特征向量进行线性变换处理,所述归一化处理层用于对线性变化处理后的结果进行归一化处理。

[0194] 步骤S903,将所述开始位置和所述结束位置输入至预设损失模型中,得到损失结果。

[0195] 这里,所述预设损失模型用于将所述开始位置和结束位置与预设的开始位置和预设的结束位置进行比较,得到所述损失结果。所述预设损失模型中包括损失函数,通过所述损失函数可以计算所述开始位置与预设的开始位置之间的距离,以及计算所述结束位置与

预设的结束位置之间的距离,将所述距离确定为所述损失结果。

[0196] 步骤S904,根据所述损失结果,对所述答案位置预测网络模型进行修正,得到所述问答匹配网络模型。

[0197] 这里,当所述距离大于预设阈值时,则所述损失结果表明当前的问答匹配网络模型中的特征提取网络模型不能准确的对词和字符进行特征提取,或者答案位置预测网络模型不能对答案的开始位置和结束位置进行准确预测。因此,需要对当前的问答匹配网络模型进行修正,则可以根据所述距离,对所述特征提取网络模型和所述答案位置预测网络模型进行修正,直至所述问答匹配网络模型输出的开始位置和结束位置与预设的开始位置和预设的结束位置之间的距离满足预设条件时,将对应的问答匹配网络模型确定为训练好的问答匹配网络模型。

[0198] 本申请实施例中,所述预设阈值可以为零,也就是说,当所述距离等于零时,则表明当前的问答匹配网络模型为训练好的问答匹配网络模型。

[0199] 本申请实施例提供的模型训练方法,由于将样本数据中的待解答的问题、用于匹配所述问题答案的文本和字符,输入至特征提取网络模型中,对应得到与所述问题和所述文本对应的词特征向量、和与所述字符对应的字符特征向量;以及,将所述词特征向量和所述字符特征向量分别输入至答案位置预测网络模型中,得到与所述问题对应的答案的开始位置和结束位置,并根据损失函数对开始位置和结束位置进行判断,从而能够根据损失结果对所述特征提取网络模型和所述答案位置预测网络模型进行修正,所得到的问答匹配网络模型能够准确的预测不同答案类型的答案的开始位置和结束位置,从而使得采用该问答匹配网络模型进行问答匹配时,能够得到与问题匹配的不同答案类型的准确答案,提高用户体验。

[0200] 下面,将说明本申请实施例在一个实际的应用场景中的示例性应用。

[0201] 本申请实施例提供一种问答匹配方法和问答匹配网络模型训练方法,该方法的核心是:1)将无答案和是否型答案作为特殊字符,以特定的组合方式,合并入原有输入;2)在预测答案时,对于片段型、无答案型和是否型三类问题,均采用统一的形式,即预测答案的开始位置和结束位置;3)在预测无答案型和是否型问题时,可配置位移或反置情况,以提高预测准确度。

[0202] 本申请实施例所提出的方法,通过有机的结合上述三个技术核心点,可以更好的应对多答案类型的机器阅读理解任务,进而可以部署更符合实际情况的相关服务,构建相关产品。本申请实施例的方法统一了多种答案类型问题的答案输出形式,均使用预测答案的开始位置和结束位置。采用本申请实施例的方法构建的系统,不会出现多模块级联的情况,因此不会产生级联错误传导的情况。实验结果也证明,本申请实施例提出的方法有着更优的系统准确度。另外,由于本申请实施例的方法统一了答案输出形式,因此在模型训练时可以使用统一形式的数据进行训练,训练数据的构建和训练的难度均可以得到明显的降低。

[0203] 下面是使用本申请实施例的方法与相关技术中的方法,分别在进行问答匹配时的实验结果对比:

[0204] 在英文公开数据集SQuAD 2.0上的测试中,采用相关技术中的BERT模型进行问答匹配时的平衡F分数(即精确率和召回率的调和平均数,F1值)为84.31,而采用本申请实施

例提供的方法进行问答匹配时的F1值为85.07。

[0205] 在中文公开数据集CJRC 2019上的测试中,采用相关技术中的BERT模型进行问答匹配时的F1值为79.20,而采用本申请实施例提供的方法进行问答匹配时的F1值为82.4。

[0206] 本申请实施例可以通过多种智能问答产品的形式向用户提供服务,或者以云服务的形式服务企业或者个人开发者客户。其中,可以至少应用于终端产品和云服务产品。

[0207] 对于终端产品,包括但不限于以下产品形式:

[0208] 1) 智能搜索展示。

[0209] 本申请实施例可以作为搜索产品的一部分,特别优化问答类(Query)的用户体验。实际变现形式既可以是直接搜索,也可以是对已有搜索结果的增强。至于搜索产品则可以是网页搜索、APP内搜索或者某个产品或者应用的单独功能,可以不做限制,均是本申请实施例所适用的场景。

[0210] 如图10A所示,是本申请实施例提供的直接搜索产品的界面图,通过在输入框1001输入问题,并点击搜索,则可以直接显示搜索结果。进一步地,在图10A的基础上,图10B是本申请实施例提供的搜索结果增强产品的界面图,如图10B所示,可以在搜索的同时,将获取搜索结果的搜索文本内容1002显示在当前界面上。

[0211] 2) 智能语音助手

[0212] 本申请实施例也可以作为智能语音助手的一个技能嵌入于手机或者智能屏等设备,提供相应的问答或者智能搜索等服务。

[0213] 如图10C所示,是本申请实施例提供的智能助手的界面图,用户可以语音输入问题“日本的首都在哪里”,则系统可以输出答案“东京是日本的首都”,以及,用户输入文本问题“东京是日本最大的城市吗”,系统输出答案“是”。

[0214] 需要说明的是,在上述问答匹配时,先检索“东京”的相关文档,然后基于本申请实施例所提供的方法在文档中找到答案并反馈。

[0215] 如图10D所示,是本申请实施例提供的智能屏的界面图,所述智能屏终端上可以运行有A智能语音助手,用户可以对着智能屏终端进行语音提问,例如,可以提问“丽江有什么好玩的地方”,则系统可以给出答案“丽江的著名景点是老君山,虎跳峡,丽江古城,白沙壁画,束河古镇,玉龙雪山”;用户还可以对着智能屏终端进行语音提问“丽江的南面在哪里”,则系统由于无法匹配到对应的答案,则给出回答“抱歉,我还没掌握该项知识”。

[0216] 需要说明的是,在上述问答匹配时,先检索“丽江”的相关文档,然后基于本申请实施例所提供的方法在文档中找到答案并反馈。

[0217] 对于云服务产品,本申请实施例可以作为云服务提供给企业客户,向企业提供抽取式机器阅读理解的API,该API还可以集成至某个SDK当中。

[0218] 下面提供一个API的例子作为参考,如下表2所示,为API的输入参数:

[0219] 表2 输入参数

[0220]

参数名称	必选	类型	参数描述
Action	是	String	公共参数, 本接口取值: ReadingComprehension
Version	是	String	公共参数, 本接口取值: 2019-08-17
Region	是	String	公共参数, 详见产品支持的地域列表
DocumentText	是	String	相关文档文本
QuestionText	是	String	问题文本
Language	是	String	目标语言
ProjectId	是	Integer	项目 ID

[0221] 如上表2所示,所有参数均为字符串(String)类型,其中,输入参数可以至少包括:功能参数(Action)、版本参数(Version)、地域参数(Region)、相关文档文本(DocumentText)、问题文本(QuestionText)、目标语言(Language)和项目ID(ProjectId)。其中,

[0222] 功能参数:为公共参数,接口取值为ReadingComprehension;

[0223] 版本参数:为公共参数,接口取值可以是2019-08-17;

[0224] 地域参数:为公共参数,参数取值详见产品支持的地域列表;

[0225] 相关文档文本:文本可以统一使用8位元(utf-8,Universal Character Set/Unicode Transformation Format)格式编码,非utf-8格式编码字符会处理失败,请传入有效文本、超文本标记语言(HTML,Hyper Text Markup Language)标记等非常规文本会处理失败,且单次请求的文本长度需要低于2000;

[0226] 问题文本:格式可以参考上述DocumentText;

[0227] 目标语言:参照支持语言列表。其中支持语言列表可以包括:zh表示简体中文;zh-TW表示繁体中文;en表示英语……

[0228] 项目ID:可以根据控制台-账号中心-项目管理中的配置填写,如无配置请填写默认项目ID为0。

[0229] 如下表3所示,为API的输出参数:

[0230] 表3 输出参数

	参数名称	类型	描述
	AnswerText	String	答案文本
[0231]	LongAnswerText	String	答案所在句子
	AnswerType	String	答案的类型
	RequestId	String	唯一请求 ID

[0232] 如上表3所示,所有参数均为字符串(String)类型,其中,输出参数包括:答案文本(AnswerText)、答案所在句子(LongAnswerText)、答案的类型(AnswerType)和唯一请求ID(RequestId)。其中,答案的类型包括无答案、是否答案、文本答案;唯一请求ID在每次请求时都会返回,在定位问题时需要提供该次请求的唯一请求ID。

[0233] 下面继续说明本申请实施例提供的问答匹配方法,本本申请实施例的方法的核心有三点,包括:1) 在特定的位置,增加三类特殊字符,以分别表示“无法回答”、“是”、“否”三种特殊类型的答案。2) 基于1),将多种类型的答案生成,统一为片段型答案的抽取。即,通过预测答案的开始位置和结束位置,预测答案的类型与答案文字。3) 基于2),对增加的三个特殊字符(“无法回答”、“是”、“否”)的开始位置和结束位置进行位移或反置等操作,以提高最终答案的准确率。

[0234] 通过综合利用以上三点,本申请实施例的方法将多种类型的答案生成方式统一成了预测答案的开始位置和结束位置的形式。采用该方法构建的系统,由于不同答案类型问题的预测方法得到了统一,因此可以避免多模块级联的情况,不会产生级联错误传导的情况。

[0235] 下面,首先介绍抽取式机器阅读理解的答案生成方法,以便于更好的理解本申请实施例的方法。

[0236] 预测答案的起止位置是抽取式机器阅读理解技术能够生成答案的核心。抽取式机器阅读理解方法可以大致分为两类方法,第一类,表示学习+答案起止位置预测;第二类,预训练语言模型+答案起止位置预测。

[0237] 对于上述第一类方法,答案预测方法如图11所示:首先,对于问题Q和相关文档D中的每个词(或字) $Tokq_1, Tokq_2 \dots Tokq_n$ 和 $Tokd_1, Tokd_2 \dots Tokd_n$,查表得到对应的向量表示 $Eq_1, Eq_2 \dots Eq_n$ 和 $Ed_1, Ed_2 \dots Ed_n$,然后通过用于表示学习的神经网络1101,得到答案中每个词(或字)的特征向量表达 $Td_1, Td_2 \dots Td_n$ 。之后,上述向量表达通过另一个全连接神经网络1102(即答案位置预测神经网络)通过线性变换得到每个位置作为答案开始的分数 P_1 (如图11中的 $P_{11}, P_{12} \dots P_{1n}$)和每个位置作为答案开始的分数 P_2 (如图11中的 $P_{21}, P_{22} \dots P_{2n}$),再分别通过归一化指数函数(Softmax函数),分别生成答案开始位置的概率分布和答案终止位置的概率分布。其中每个虚线框中的所有数值相加为1。最后,分别找到两行中值最大位置作为答案的起止位置,就可以得到最终的答案片段。

[0238] 对于上述第二类方法,答案预测方法如图12所示:首先,将问题Q和文本D的每个词(或字)按照图中顺序相连,并在开始、中间、结尾处拼接“[CLS]”和“[SEP]”两种特殊字符,

其中[CLS]表示开始字符,[SEP]表示分隔字符。接着,通过用于表示学习的神经网络1201(通常为一种深度神经网络),得到问题Q的每个词(或字)的向量表示 $Tq_1, Tq_2 \cdots Tq_n$ 和文本窗口D中每个词(或字)的向量表示 $Td_1, Td_2 \cdots Td_n$,以及特殊字符对应的向量表示 $T[cls], T[sep]$ 。然后,通过另一个全连接神经网络1202预测得到答案开始和终止的概率分布。最后,分别找到两行中值最大位置作为答案的起止位置,就可以得到最终的答案片段。特别的,如果预测的开始位置和结束位置在所添加的[CLS]位置时,则表示无答案。

[0239] 为了将多种类型的答案生成方式统一成了预测答案的开始位置和结束位置的形式,本申请实施例还设计了在特定的位置增加三类特殊字符的方法。其中,三类特殊字符是“[NA]”、“[YES]”和“[NO]”,分别表示“无法回答”(No Answer)、“是”(Yes)、“否”(No)三种特殊类型的答案。值得说明的是,以上三个字符表示的是三个不同于字词和已有的[CLS]与[SEP]的字符,与拼写无关。

[0240] 在上述实施例中,介绍了基于表示学习的神经网络方法的输入拼接方法,即依次将[CLS]、问题中的各个词(字)、[SEP]、文档中的各个词(字)、[SEP]拼接起来。在此基础上,添加[NA]、[YES]、[NO]三个特殊的字符和[SEP]字符的序列。除[SEP]必须在其他三个字符后,其他三个特殊字符的顺序并不限定。

[0241] 这里,可以有以下三个不同的添加位置,如图13所示,三个位置分别是:在[CLS]后,或者说问题第一个字之前;在第一个[SEP]后,或者说在文档第一个字之前;在第二个[SEP]后,或者说在原序列的最后。在使用时,三个位置需要选择其中一个使用,不能同时使用。经过实践,添加在第一种位置,即在[CLS]后时,通常可以取得最好的预测精度。

[0242] 本申请实施例中,可以将多种类型的答案生成方式统一成预测答案的开始与终止位置的形式。位置预测模型(通常是一个神经网络)对拼接后的每一个词(或字)表示的位置生成两个概率,一个表示该位置作为答案开始位置的概率,另一个表示该位置作为答案终止位置的概率。

[0243] 如图14所示,从所有的位置中,选取答案开始位置最大概率和答案结束位置最大概率两个位置。如果预测的开始位置和结束位置均指向所添加的[NA]位置(或者[YES],或者[NO])时,则表示答案为“无答案”(或者“是”,或者“否”)。如果开始和终止位置指向表示词(或字)的某个位置,则从开始位置到终止位置的连续的词(字)组成答案片段。

[0244] 在一些实施例,对于无答案类型和是否答案类型的问题,还可以通过设置开始位置和结束位置的位移和反置,可以有效地提高系统预测的精确度。

[0245] 其中,位移:表示在预测时要求答案的开始位置与结束位置有一个词(字)的偏移量。而不是像上述实施例当中所说的,开始位置与结束位置指向同一个词(字)。如图15所示,可以设置位移的偏移量为一个词(字),则仅当开始位置在“[NA]”且终止位置在“[YES]”时(如图15中虚线框圈出的字符),表示所预测的答案为“无答案”。同理,当开始位置在“[YES]”且终止位置在“[NO]”时(图中未示出),表示所预测的答案为“是”;当开始位置在“[NO]”且终止位置在“[sep]”时(图中未示出),表示所预测的答案为“否”。

[0246] 反置:表示当出现开始位置与结束位置的前后顺序反转时,预测结果为无答案或者是否答案。正常情况下,开始位置通常在终止位置的前面,而反置情况,就是专门设置在出现反常规的位置情况时,预测特定类型的答案。如图16所示,仅当开始位置在“[YES]”且终止位置在“[NA]”时(如图16中虚线框圈出的字符),则可以将开始位置和结束位置反置,

进而根据反置后的结果确定答案类型,其中图16中表示所预测的答案为“无答案”。同理,当开始位置在“[NO]”且终止位置在“[YES]”时(图中未示出),表示所预测的答案为“是”;当开始位置在“[SEP]”且终止位置在“[NO]”时(图中未示出),表示所预测的答案为“否”。

[0247] 需要说明的是,位移或反置可根据实际情况进行设置,且仅可使用其中一个。当然,也可以不选用其中任何一个,本申请实施例不做限定。

[0248] 本申请实施例提出的方法改进的是抽取式阅读理解中答案起止位置预测部分。目前的主流方法,包括但不限于基于BERT、ELMo、GPT、GPT2.0 XL-Net等预训练语言模型的方法,以及基于BiDAF、RNet、QANet等表示学习方法,都可以通过使用本申请实施例提出的方法对原有方法进行改造升级。

[0249] 本申请实施例的一个重点是特定符号插入的位置,当然,除了本申请实施例所列举的插入位置,还可以将特定符号插入其他位置,也可以起到类似的效果。对于特定问题或者特定场景,位置的选择会影响最终的结果。

[0250] 本申请实施例支持改变特定符号的排列顺序,并非一定按照本文所述的顺序[NA]、[YES]、[NO]、[SEP]。另外,如果不使用特定符号,而使用指定的词(字)也是可以的,比如“无”、“是”、“否”、[SEP]。

[0251] 下面继续说明本申请实施例提供的问答匹配装置354实施为软件模块的示例性结构,在一些实施例中,如图3所示,存储在存储器350的问答匹配装置354中的软件模块可以是服务器300中的问答匹配装置,包括:

[0252] 获取模块3541,用于获取问答匹配文档对应的第一词序列,所述问答匹配文档包括待解答的问题和用于匹配所述问题答案的文本,所述第一词序列中包括至少两个词,每一词包括一个或多个字;

[0253] 第一确定模块3542,用于在所述第一词序列中,确定用于设置字符的目标位置;

[0254] 设置模块3543,用于在所述目标位置设置用于表征答案类型的字符,形成第二词序列;

[0255] 第二确定模块3544,用于在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置;

[0256] 第三确定模块3545,用于根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。

[0257] 在一些实施例中,所述获取模块还用于:对所述问题和所述文本依次进行分词处理,得到包括至少一个问题词和至少一个文本词的所述第一词序列,其中所述文本词位于所述问题词之后。

[0258] 在一些实施例中,所述第一确定模块还用于:在所述第一词序列中,将位于所述问题词之前的位置,确定为所述目标位置;或者,在所述第一词序列中,将位于所述问题词与所述文本词之间的位置,确定为所述目标位置;或者,在所述第一词序列中,将位于所述文本词之后的位置,确定为所述目标位置。

[0259] 在一些实施例中,所述第二确定模块还用于:对所述第二词序列中的词和字符分别进行特征提取,对应得到词特征向量和字符特征向量;根据所述词特征向量和所述字符特征向量,确定所述第二词序列中的词和字符分别作为答案的开始位置和结束位置的概率值;根据所述概率值,确定所述答案的开始位置和结束位置。

[0260] 在一些实施例中,所述第二确定模块还用于:分别对所述词特征向量和字符特征向量依次进行线性变换和归一化处理,得到所述第二词序列中的词和字符分别作为所述开始位置和所述结束位置对应的概率值集合;所述概率值集合中包括至少两个所述概率值,所述概率值是用于表征所述答案的开始位置或结束位置的分布概率。

[0261] 在一些实施例中,所述第二确定模块还用于:分别对所述词特征向量和字符特征向量,依次进行第一线性变换和第一归一化处理,得到与所述第二词序列中的词和字符对应的第一概率值集合;所述第一概率值集合中包括至少两个第一概率值,所述第一概率值是用于表征所述答案的开始位置的分布概率;分别对所述词特征向量和字符特征向量,依次进行第二线性变换和第二归一化处理,得到与所述第二词序列中的词和字符对应的第二概率值集合;所述第二概率值集合中包括至少两个第二概率值,所述第二概率值是用于表征所述答案的结束位置的分布概率。

[0262] 在一些实施例中,所述第二确定模块还用于:将所述第一概率值集合中最大第一概率值对应的词或字符,确定为目标开始内容;将所述目标开始内容在所述第二词序列中所处的位置,确定为所述开始位置;将所述第二概率值集合中最大第二概率值对应的词或字符,确定为目标结束内容;将所述目标结束内容在所述第二词序列中所处的位置,确定为所述结束位置。

[0263] 在一些实施例中,所述装置还包括:第四确定模块,用于当所述开始位置和所述结束位置均指向所述文本对应的文本词时,将所述开始位置所指向的第一文本词与所述结束位置所指向的第二文本词之间的内容,确定为所述答案;或者,所述字符对应的答案类型至少包括以下之一:无答案类型、答案为是类型和答案为否类型;对应地,第四确定模块还用于:当所述开始位置和所述结束位置均指向所述字符时,根据所述字符对应的所述答案类型确定所述答案。

[0264] 在一些实施例中,所述第三确定模块还用于:当所述开始位置和所述结束位置均指向所述字符,且所述开始位置和所述结束位置之间满足预设偏移量条件时,根据所述开始位置和所述结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。

[0265] 在一些实施例中,所述第三确定模块还用于:当所述开始位置和所述结束位置均指向所述字符,且所述开始位置位于所述结束位置之后时,对所述开始位置与所述结束位置进行反置处理,以互换所述开始位置和所述结束位置,得到新开始位置和新结束位置;根据所述新开始位置和所述新结束位置在所述第二词序列中对应的内容,确定对应问题匹配的答案。

[0266] 在一些实施例中,所述装置还包括:控制模块,用于控制采用问答匹配网络模型,对所述第一词序列设置所述字符,得到第二词序列,并在所述第二词序列中,确定与所述问题对应的答案的开始位置和结束位置。

[0267] 需要说明的是,本申请实施例装置的描述,与上述方法实施例的描述是类似的,具有同方法实施例相似的有益效果,因此不做赘述。对于本装置实施例中未披露的技术细节,请参照本申请方法实施例的描述而理解。

[0268] 本申请实施例提供一种存储有可执行指令的存储介质,其中存储有可执行指令,当可执行指令被处理器执行时,将引起处理器执行本申请实施例提供的方法,例如,如图4示出的方法。

[0269] 在一些实施例中,存储介质可以是铁电存储器 (FRAM, Ferromagnetic Random Access Memory)、只读存储器 (ROM, Read Only Memory)、可编程只读存储器 (PROM, Programmable Read Only Memory)、可擦除可编程只读存储器 (EPROM, Erasable Programmable Read Only Memory)、带电可擦可编程只读存储器 (EEPROM, Electrically Erasable Programmable Read Only Memory)、闪存、磁表面存储器、光盘、或光盘只读存储器 (CD-ROM, Compact Disk-Read Only Memory) 等存储器;也可以是包括上述存储器之一或任意组合的各种设备。

[0270] 在一些实施例中,可执行指令可以采用程序、软件、软件模块、脚本或代码的形式,按任意形式的编程语言 (包括编译或解释语言,或者声明性或过程性语言) 来编写,并且其可按任意形式部署,包括被部署为独立的程序或者被部署为模块、组件、子例程或者适合在计算环境中使用的其它单元。

[0271] 作为示例,可执行指令可以但不一定对应于文件系统中的文件,可以可被存储在保存其它程序或数据的文件的一部分,例如,存储在超文本标记语言 (HTML, Hyper Text Markup Language) 文档中的一个或多个脚本中,存储在专用于所讨论的程序的单个文件中,或者,存储在多个协同文件 (例如,存储一个或多个模块、子程序或代码部分的文件) 中。作为示例,可执行指令可被部署为在一个计算设备上执行,或者在位于一个地点的多个计算设备上执行,又或者,在分布在多个地点且通过通信网络互连的多个计算设备上执行。

[0272] 以上所述,仅为本申请的实施例而已,并非用于限定本申请的保护范围。凡在本申请的精神和范围之内所作的任何修改、等同替换和改进等,均包含在本申请的保护范围之内。

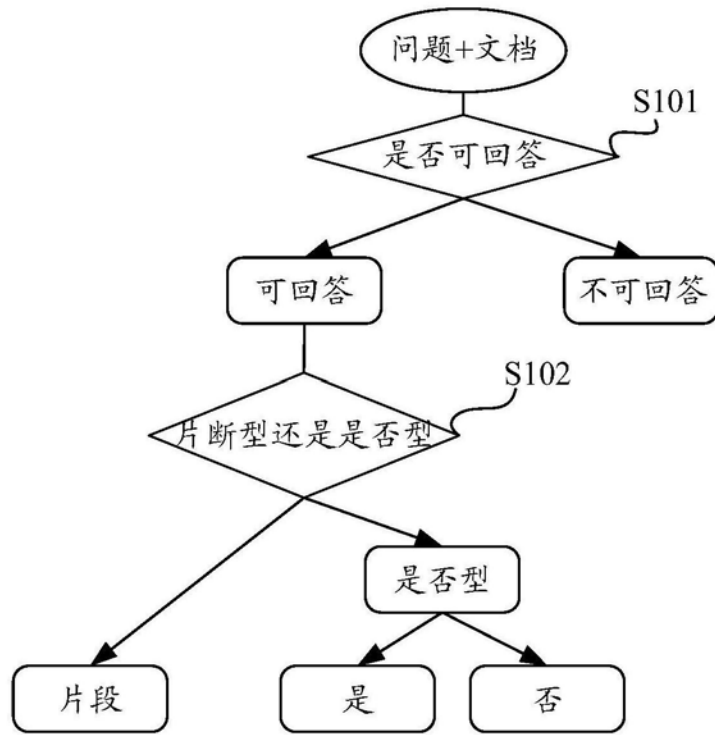


图1A

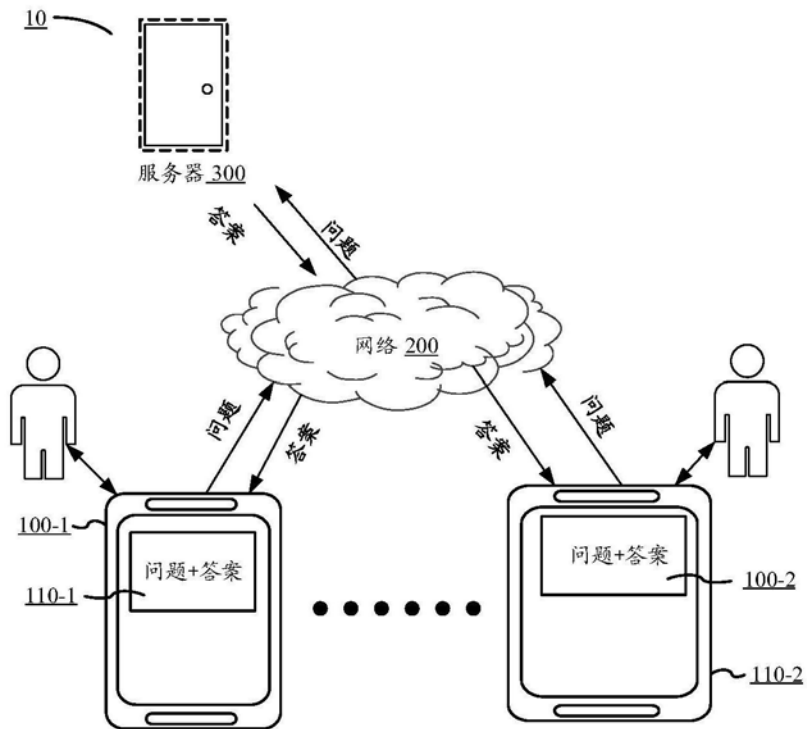


图1B

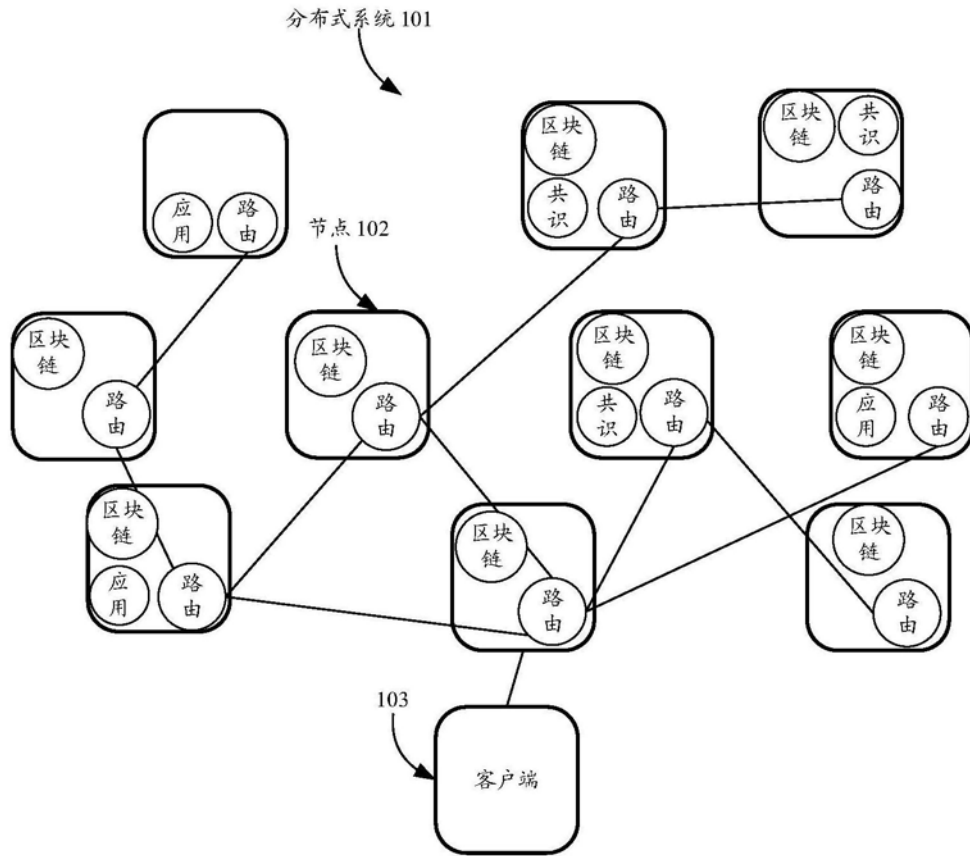


图2A

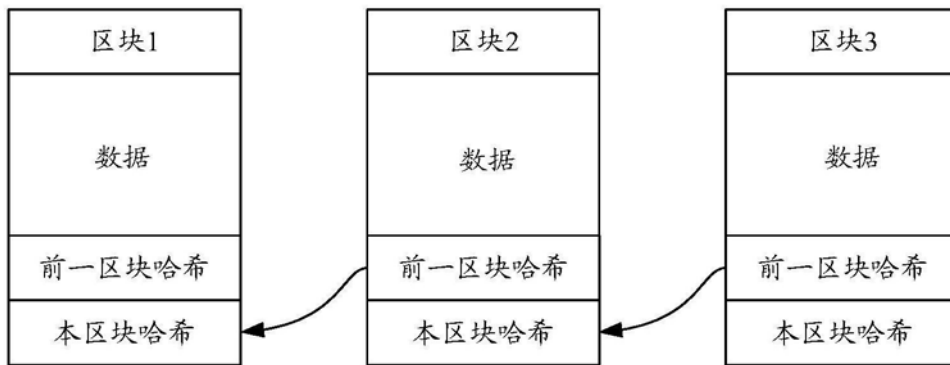


图2B

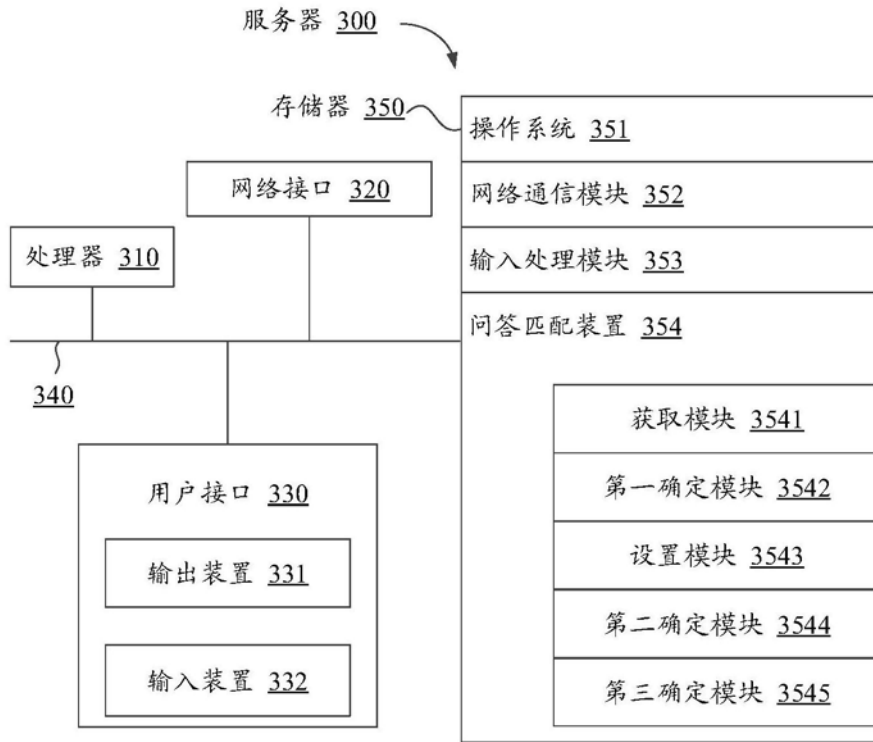


图3

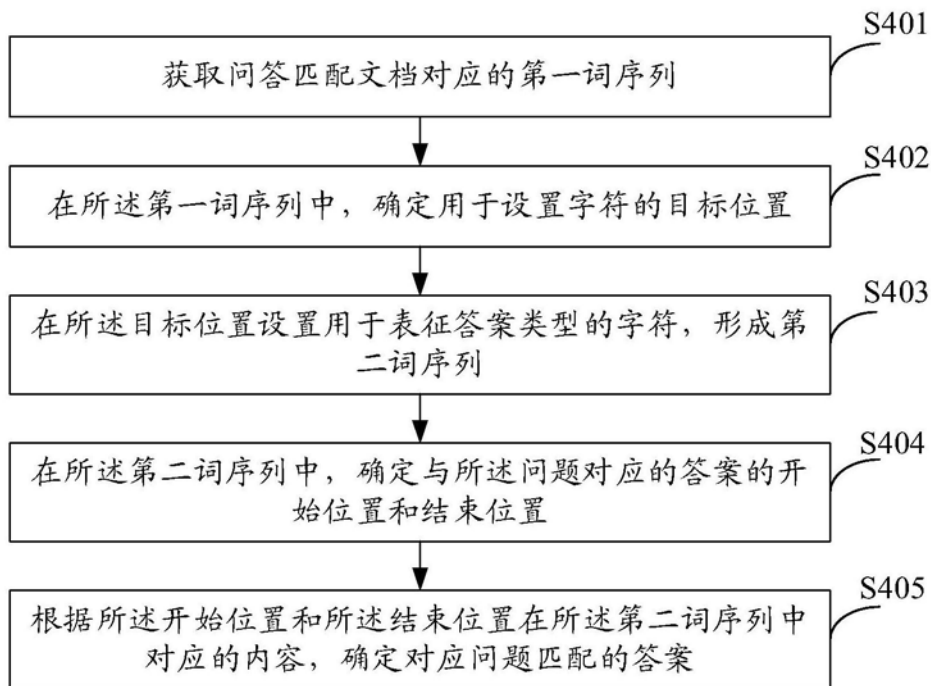


图4

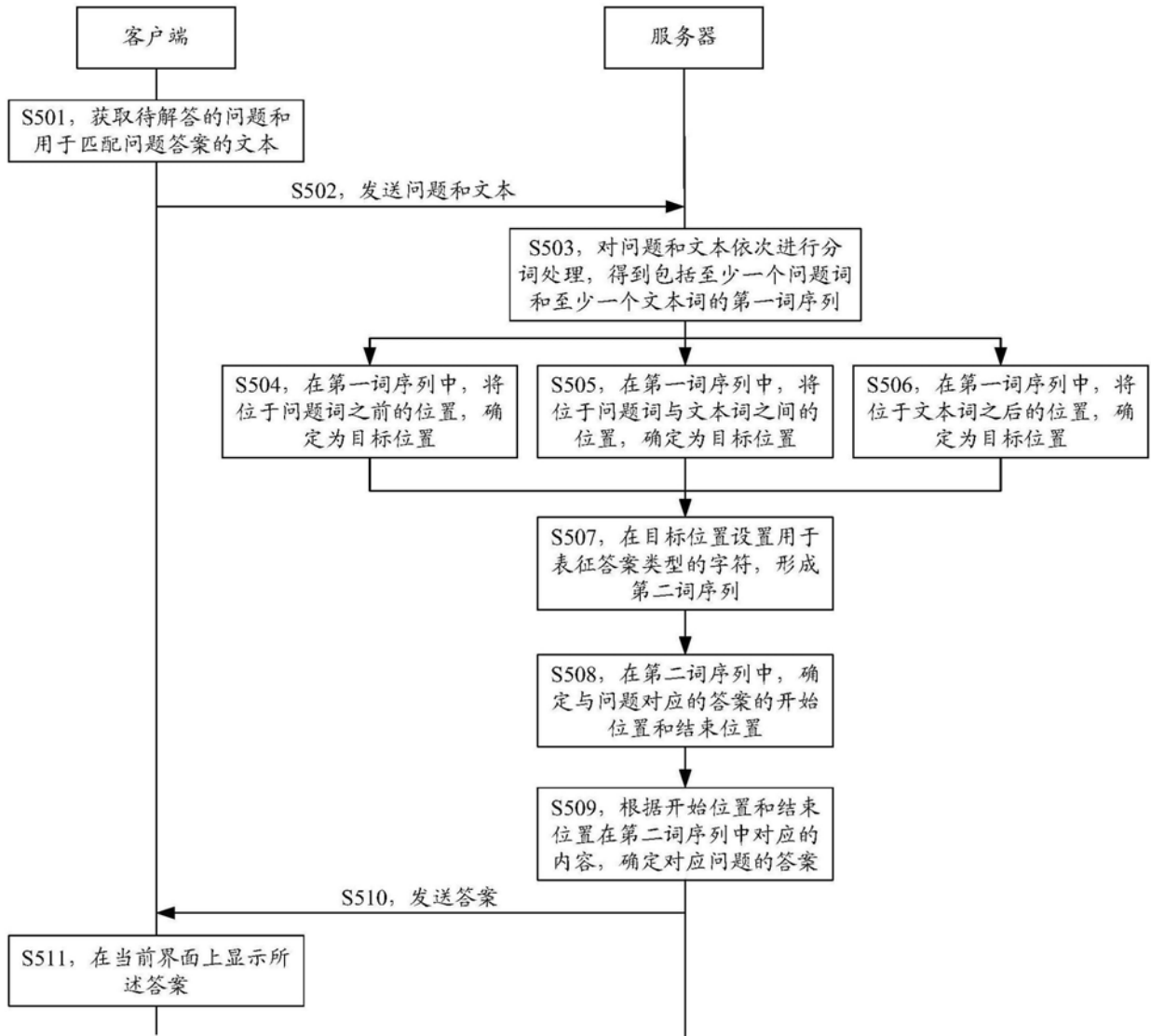


图5

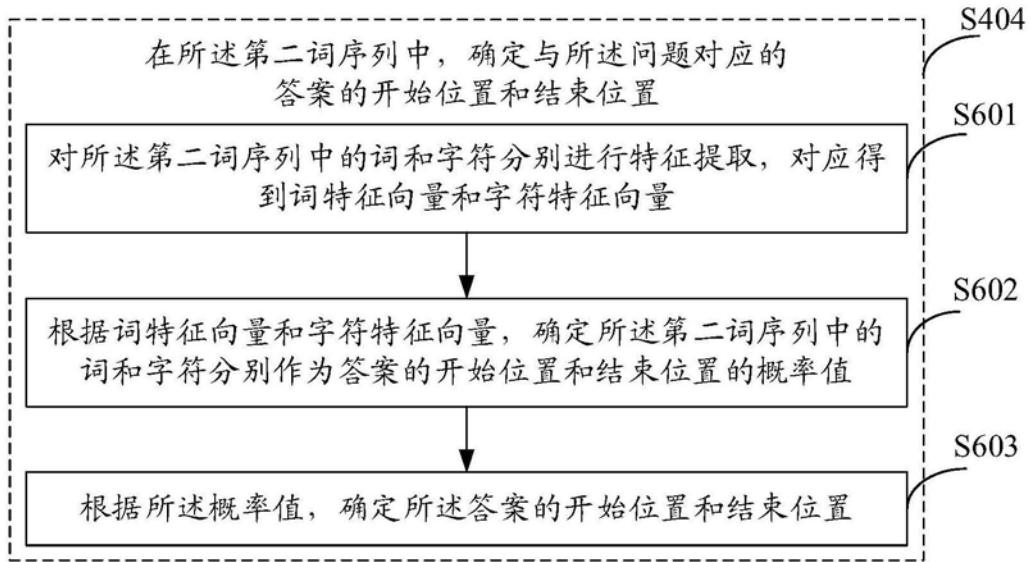


图6A

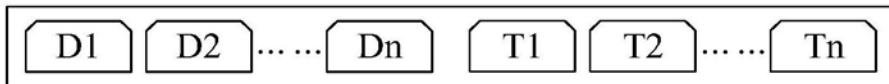


图6B

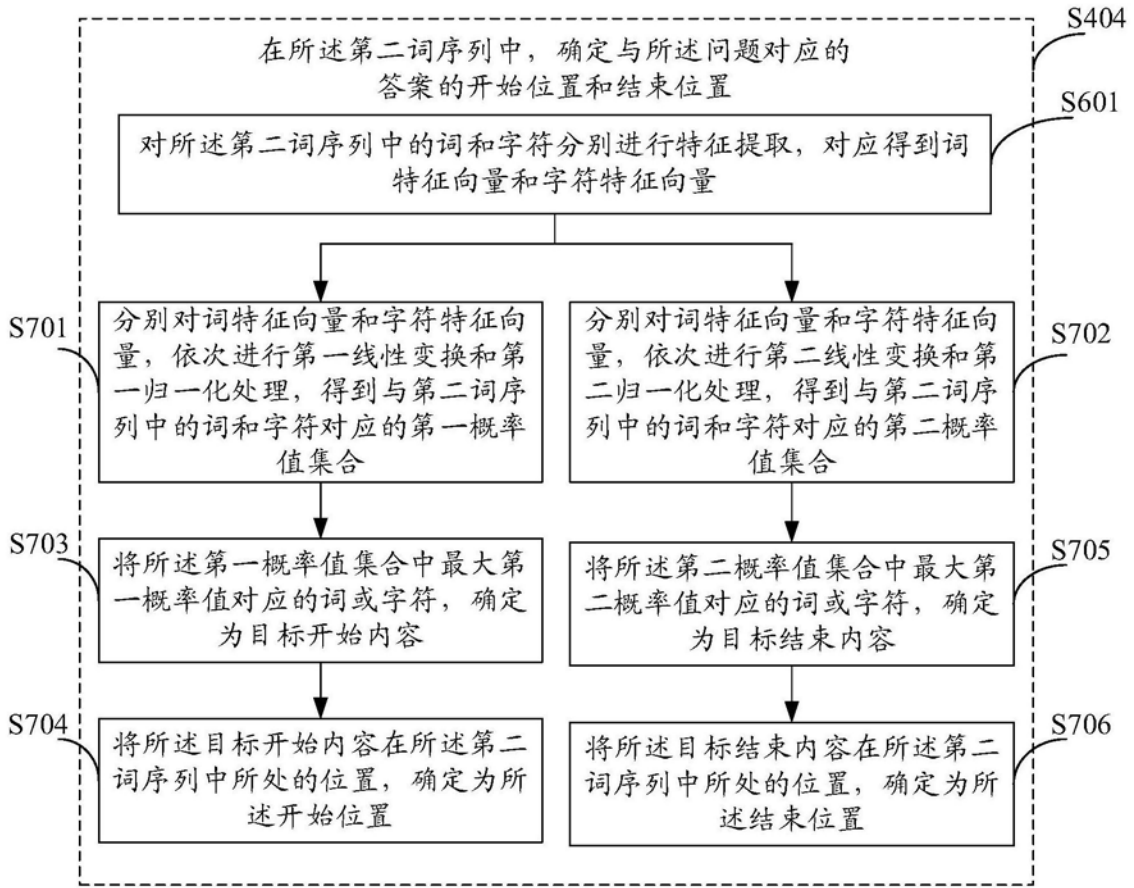


图7

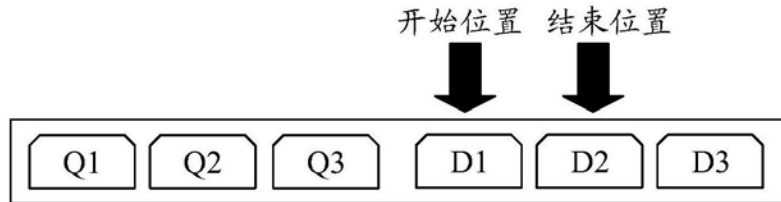


图8A

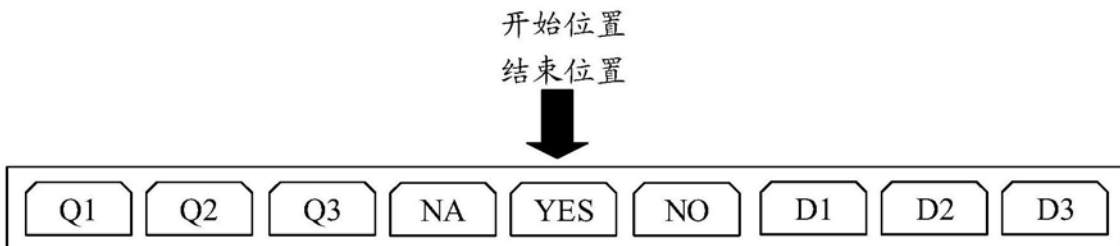


图8B

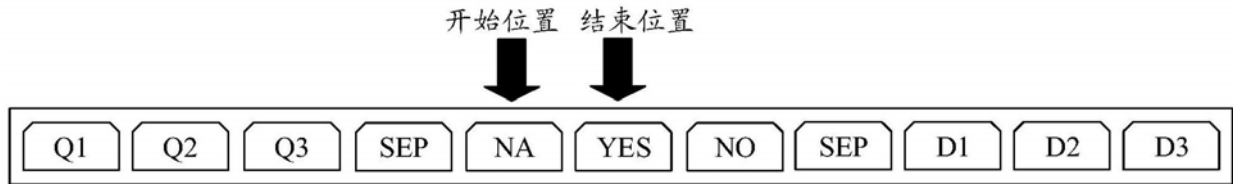


图8C

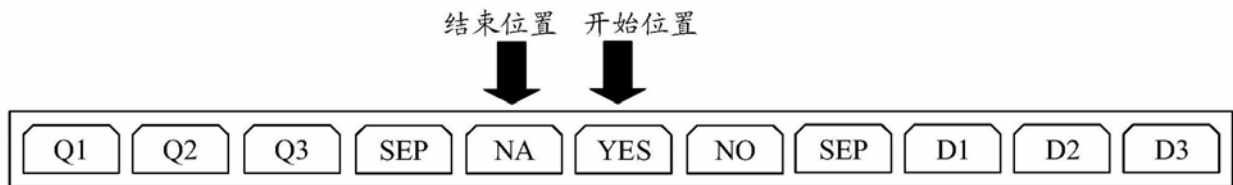


图8D

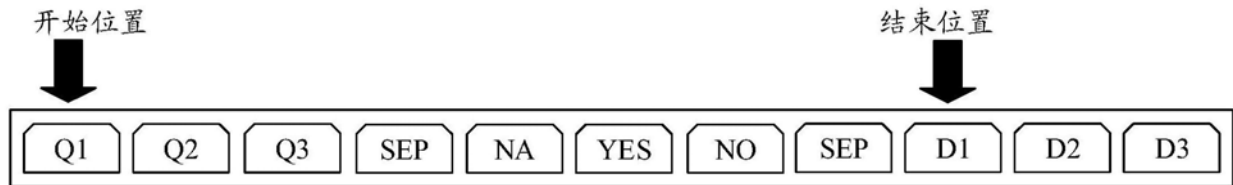


图8E

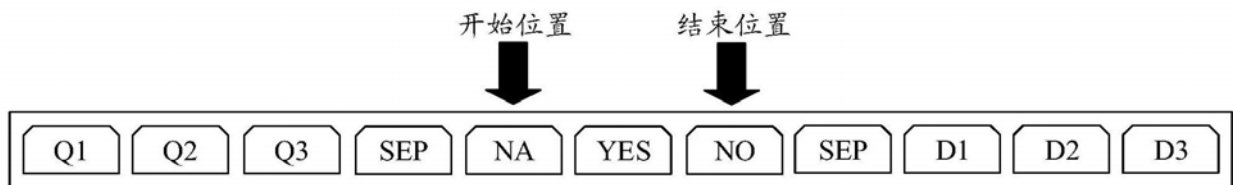


图8F

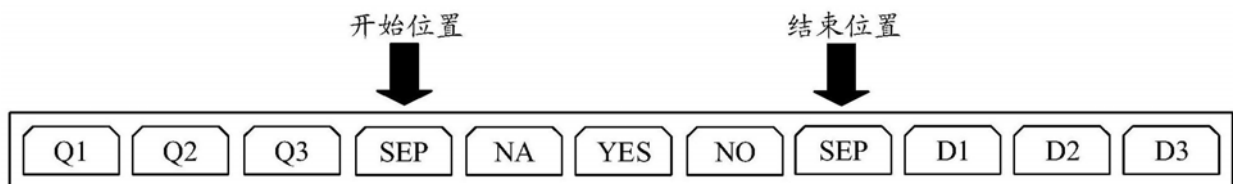


图8G

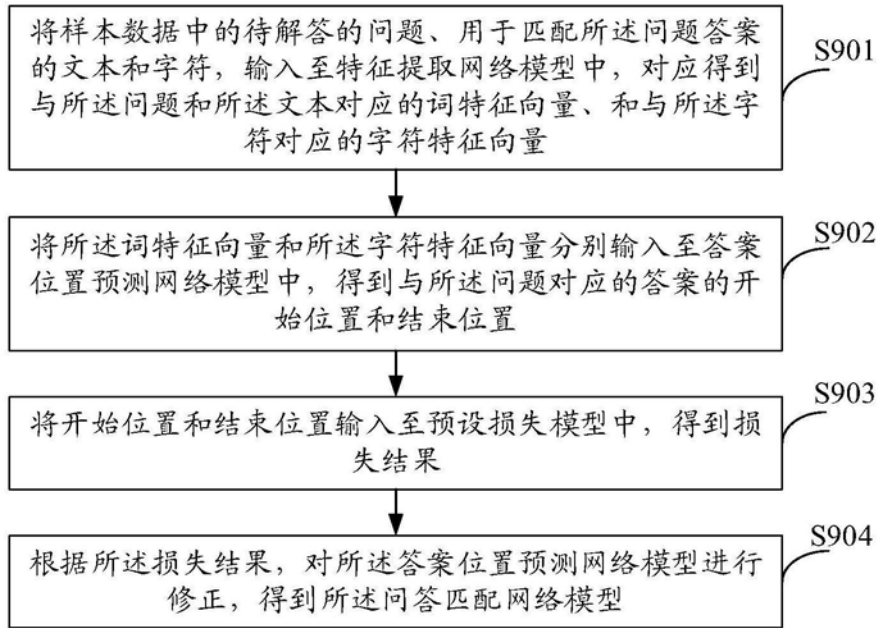


图9

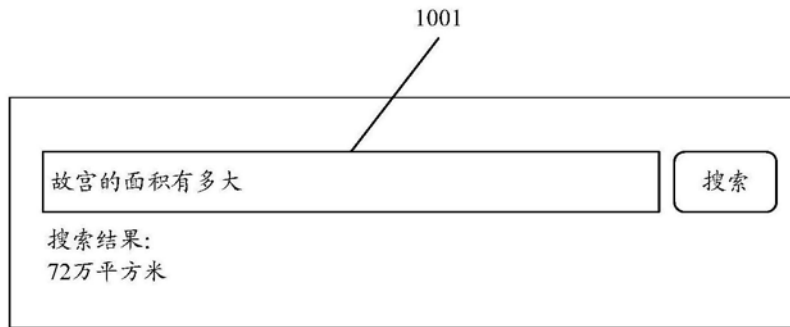
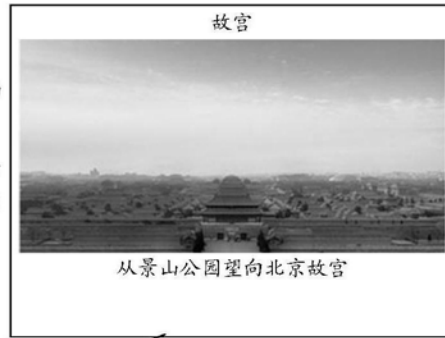


图10A

故宫的面积有多大 搜索

搜索结果：
72万平方米

北京故宫,即紫禁城,是明清两朝廿四位皇帝的皇宫。故宫始建于明成祖永乐四年(1406年),永乐十八年(1420年)落成;位于北京中轴线的中心,占地面积72万平方米,建筑面积约15万平方米,为世界上现存规模最大的宫殿型建筑。北京故宫是第一批全国重点文物保护单位、第一批国家5A级旅游景区,1987年入选《世界文化遗产》名录。故宫现为故宫博物院,藏品主要以明、清两代宫廷收藏为基础;是国家一级博物馆,与俄罗斯埃米塔什博物馆、法国卢浮宫、美国大都会博物馆、英国大英博物馆并称为世界五大博物馆。
明初定都于南京应天府。建文年间,燕王朱棣自北平起



1002

图10B

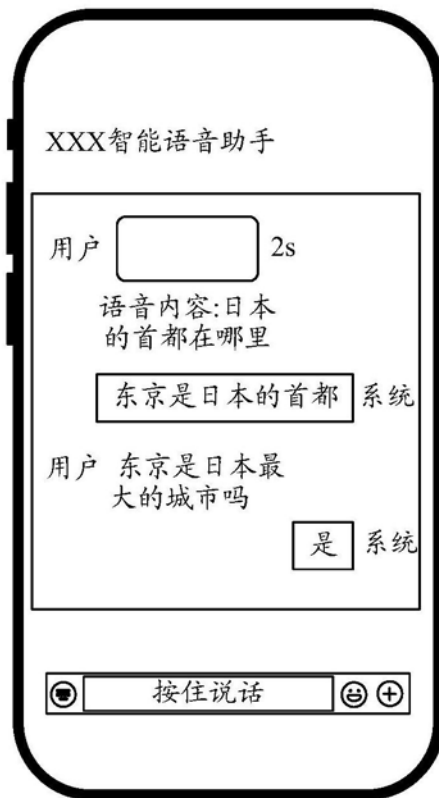


图10C

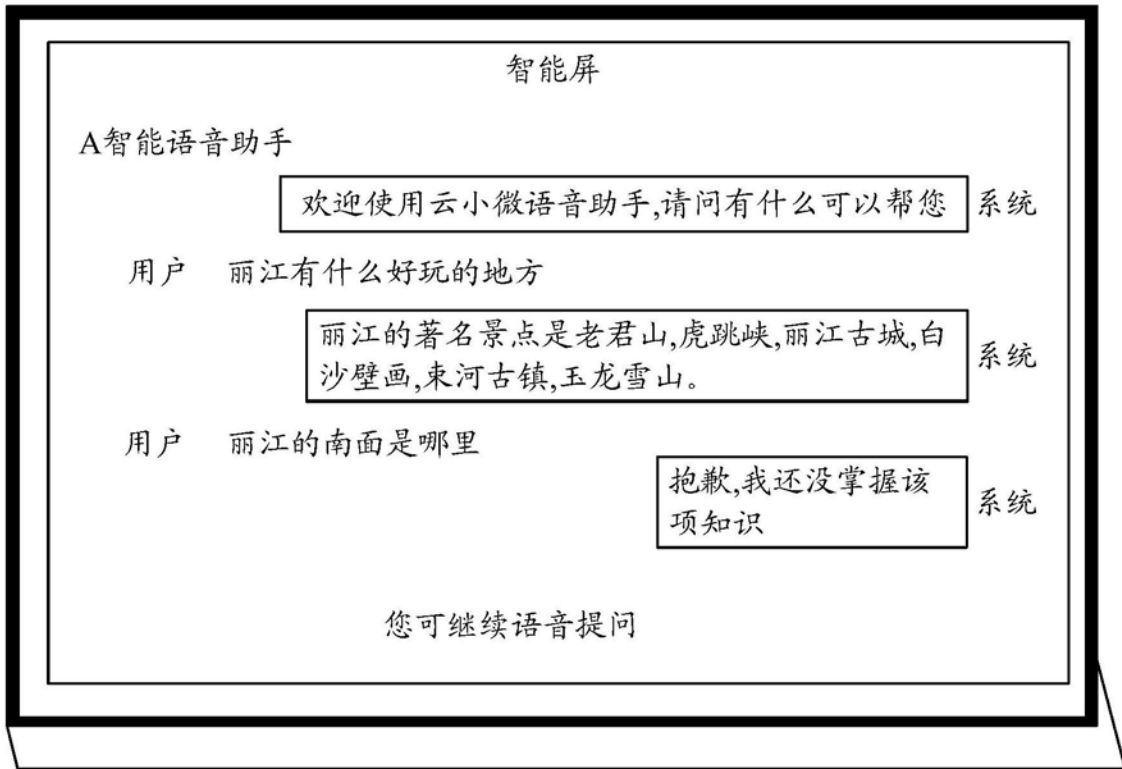


图10D

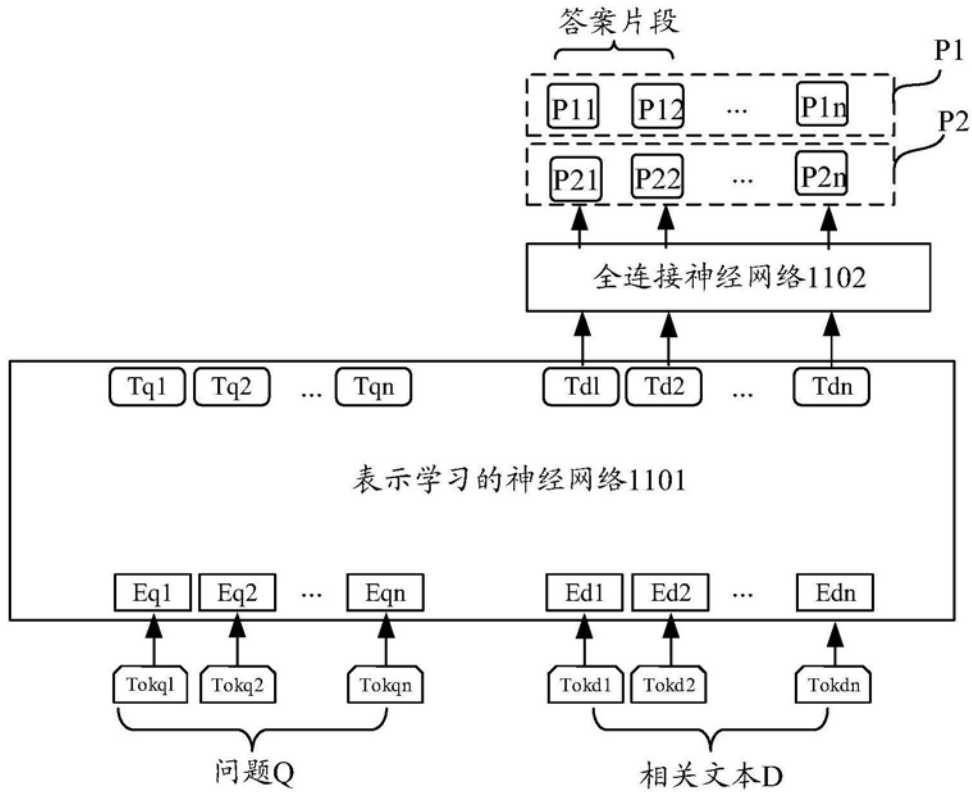


图11

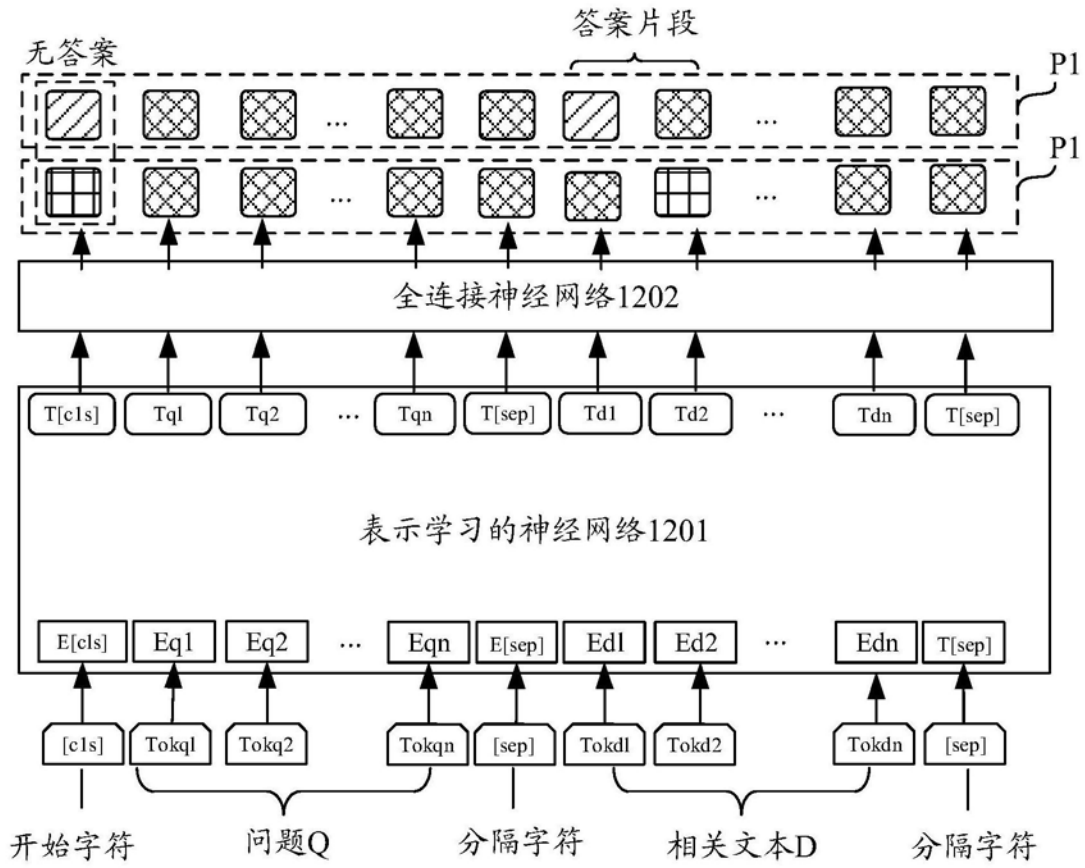


图12

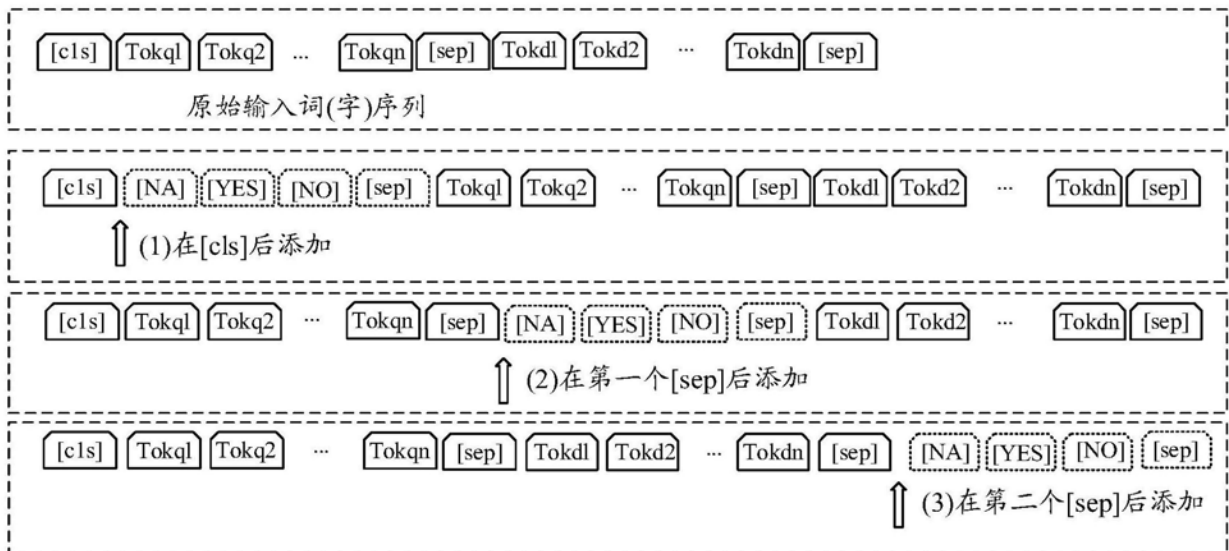


图13

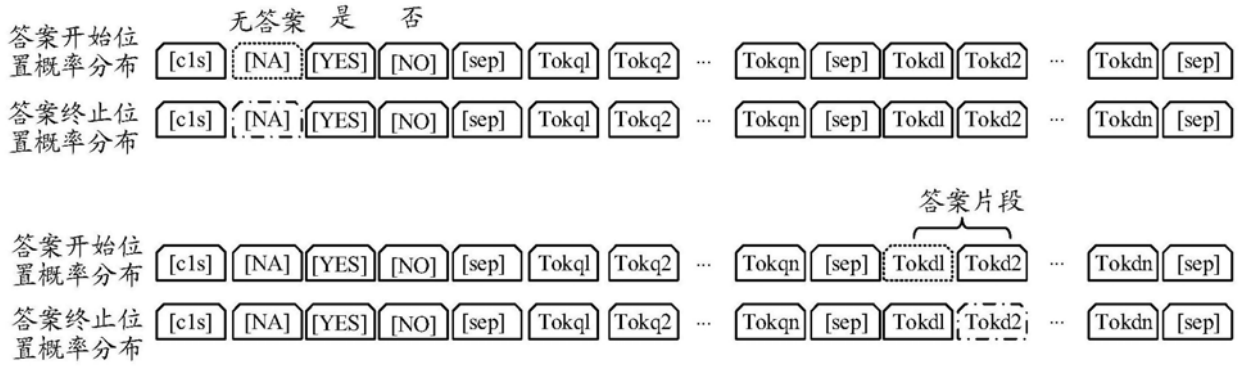


图14

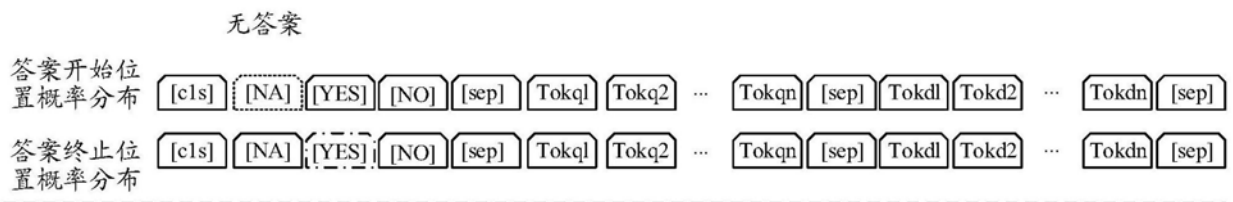


图15



图16