

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2020 年 12 月 3 日 (03.12.2020)



(10) 国际公布号
WO 2020/237872 A1

(51) 国际专利分类号:
G06F 16/35 (2019.01)

(21) 国际申请号: PCT/CN2019/103024

(22) 国际申请日: 2019 年 8 月 28 日 (28.08.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201910441486.5 2019 年 5 月 24 日 (24.05.2019) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 王丽杰(WANG, Lijie); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(74) 代理人: 北京市立方律师事务所(LIFANG & PARTNERS); 中国北京市东城区香河园街 1 号院信德京汇中心 12 层, Beijing 100028 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) **Title:** METHOD AND APPARATUS FOR TESTING ACCURACY OF SEMANTIC ANALYSIS MODEL, STORAGE MEDIUM, AND DEVICE

(54) 发明名称: 语义分析模型准确度的校验方法、装置、存储介质及设备

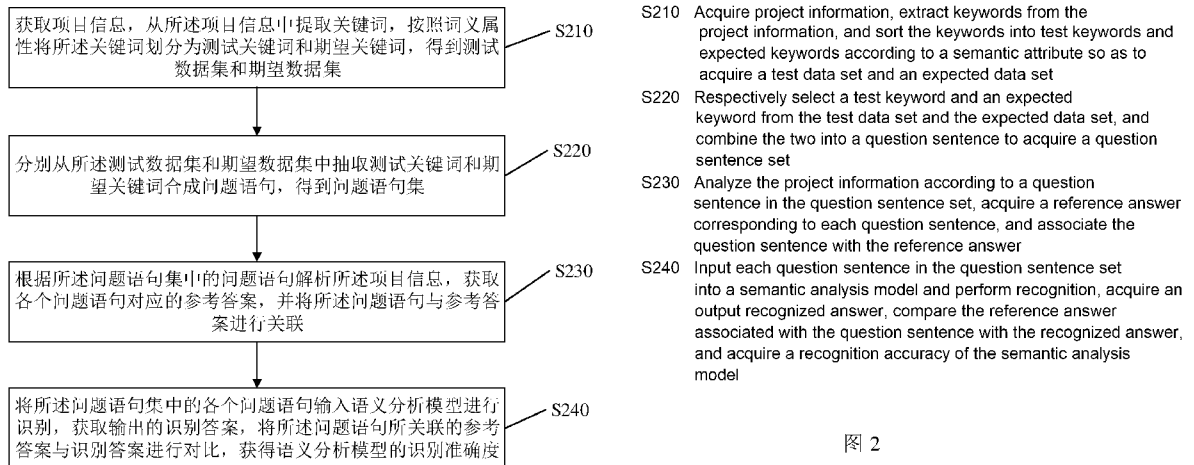


图 2

(57) **Abstract:** The present application relates to the technical field of testing, and particularly relates to a method and apparatus for testing the accuracy of a semantic analysis model, a non-volatile storage medium, and a device. The method for testing the accuracy of a semantic analysis model comprises: acquiring project information, extracting keywords from the project information, and sorting the keywords into test keywords and expected keywords according to a semantic attribute so as to acquire a test data set and an expected data set; respectively selecting a test keyword and an expected keyword from the test data set and the expected data set, and combining the two into a question sentence to acquire a question sentence set; acquiring a reference answer corresponding to each question sentence, and associating the question sentence with the reference answer; and inputting each question sentence in the question sentence set into a semantic analysis model and performing recognition, and acquiring a recognition accuracy of the semantic analysis model. The solution provided by the present application increases the number of problems in a problem set, thereby resolving the issue in which the efficiency of testing a semantic analysis model is low.



SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

根据细则4.17的声明:

- 发明人资格(细则4.17(iv))

本国际公布:

- 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本申请涉及校验技术领域, 尤其涉及一种语义分析模型准确度的校验方法、装置、非易失性存储介质及设备。其中, 语义分析模型准确度的校验方法, 包括: 获取项目信息, 从所述项目信息中提取关键词, 按照词义属性将所述关键词划分为测试关键词和期望关键词, 得到测试数据集和期望数据集; 分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句, 得到问题语句集; 获取各个问题语句对应的参考答案, 并将所述问题语句与参考答案进行关联; 将所述问题语句集中的各个问题语句输入语义分析模型进行识别, 获得语义分析模型的识别准确度。本申请提供的方案通过扩充问题集的问题, 以解决语义分析模型验证效率低下的问题。

语义分析模型准确度的校验方法、装置、存储介质及设备

本申请要求于 2019 年 5 月 24 日提交中国专利局、申请号为
5 2019104414865，发明名称为“语义分析模型准确度的校验方法、装置、
存储介质及设备”的中国专利申请的优先权，其全部内容通过引用结合在
本申请中。

技术领域

10 本申请涉及校验技术领域，具体而言，本申请涉及一种语义分析模型
准确度的校验方法、装置、存储介质及设备。

背景技术

随着人工智能的发展，计算机可以帮助人类做很多工作，计算机能够
15 帮助甚至替代人类“工作”的秘诀在于：模型的建立，训练出的模型能够
进行识别、分类等工作，但模型的建立需要大量的训练样本，发明人意识
到模型的识别精度及准确度受训练样本数量的直接影响，数据量不足会显
著影响模型效果，而且模型建立后，需要用大量的数据对生成的模型进行
准确率验证。

20 尤其是在问答系统研究领域，问答样本数据往往只有数百条到数千条
量级，且问题数据的采集及结果的对比均采用手动梳理的方式进行，问题
数据的覆盖面窄，验证效率低下。

发明内容

25 本申请提供了一种语义分析模型准确度的校验方法、装置、计算机可
读非易失性存储介质及计算机设备，以扩充问题集，解决语义分析模型验
证效率低下的问题。

为解决上述技术问题，本申请提供了一种语义分析模型准确度的校验
方法，包括：获取项目信息，从所述项目信息中提取关键词，按照词义属

性将所述关键词划分为测试关键词和期望关键词，得到测试数据集和期望数据集；分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集；根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题语句对应的参考答案，并将所述问题语句与参考答案进行关联；将所述问题语句集中的各个问题语句输入语义分析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度。

为解决上述技术问题，本申请实施例还提供了一种语义分析模型准确度的校验装置，包括：提取关键词模块，用于获取项目信息，从所述项目信息中提取关键词，按照词义属性将所述关键词划分为测试关键词和期望关键词，得到测试数据集和期望数据集；生成问题语句模块，用于分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集；获取参考答案模块，用于根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题语句对应的参考答案，并将所述问题语句与参考答案进行关联；校验模块，用于将所述问题语句集中的各个问题语句输入语义分析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度。

为解决上述技术问题，本申请实施例还提供了一种计算机可读非易失性存储介质，所述计算机可读非易失性存储介质用于存储计算机指令，当所述计算机指令在计算机上运行时，使得计算机可以执行一种语义分析模型准确度的校验方法，所述语义分析模型准确度的校验方法包括以下步骤：获取项目信息，从所述项目信息中提取关键词，按照词义属性将所述关键词划分为测试关键词和期望关键词，得到测试数据集和期望数据集；分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集；根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题语句对应的参考答案，并将所述问题语句与参考答案进行关联；将所述问题语句集中的各个问题语句输入语义分析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答

案进行对比，获得语义分析模型的识别准确度。

为解决上述技术问题，本申请实施例还提供了一种计算机设备，所述计算机设备包括：一个或多个处理器；存储装置，用于存储一个或多个程序，当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现一种语义分析模型准确度的校验方法，所述语义分析模型准确度的校验方法包括以下步骤：获取项目信息，从所述项目信息中提取关键词，按照词义属性将所述关键词划分为测试关键词和期望关键词，得到测试数据集和期望数据集；分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集；根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题语句对应的参考答案，并将所述问题语句与参考答案进行关联；将所述问题语句集中的各个问题语句输入语义分析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度。

本申请实施例提供了一种语义分析模型准确度的校验方法、装置、非易失性存储介质及设备，通过从项目信息中提取的大量关键词，利用关键词的组合生成大批量的问题语句，这些问题语句能够覆盖整个项目信息，覆盖面广且数量丰富，提高语义识别模型校验准确性的同时也提高了校验效率。

附图说明

图 1 为本申请一个实施例提供的语义分析模型准确度的校验方法的实施环境图；

图 2 为本申请一个实施例提供的语义分析模型准确度的校验方法的流程示意图；

图 3 为本申请一个实施例提供的分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集的流程示意图；

图 4 为本申请一个实施例提供的分别从所述测试数据集和期望数据

集中抽取测试关键词和期望关键词合成问题语句的流程示意图；

图 5 为本申请另一个实施例提供的语义分析模型准确度的校验方法的流程示意图；

图 6 为本申请一种实施例提供的语义分析模型准确度的校验装置的结构示意图；

图 7 为本申请一种实施例提供的计算机设备的结构示意图。

具体实施方式

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

本领域技术人员可以理解，除非特意声明，这里使用的单数形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是，本申请的说明书中使用的措辞“包括”是指存在所述特征、整数、步骤、操作、元件和/或组件，但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和/或它们的组。

可以理解，本申请所使用的术语“第一”、“第二”等可在本文中用于描述各种元件，但这些元件不受这些术语限制。这些术语仅用于将第一个元件与另一个元件区分。举例来说，在不脱离本申请的范围的情况下，可以将第一直播视频图像称为第二直播视频图像，且类似地，可将第二直播视频图像称为第一直播视频图像。

图 1 为一个实施例中提供的语义分析模型准确度的校验方法的实施环境图，在该实施环境中，包括用户终端、服务器端。

本实施例提供的语义分析模型准确度的校验方案，服务器端获取项目信息，从项目信息中提取关键词，按照词义属性将提取出的关键词划分为测试关键词和期望关键词，得到测试数据集和期望数据集，分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集；根据所述问题语句集中的问题语句解析所述项目信息，获

取各个问题语句对应的参考答案，并将所述问题语句与参考答案进行关联；将所述问题语句集中的各个问题语句输入语义分析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的准确度。

5 需要说明的是，用户终端可为智能手机、平板电脑、笔记本电脑、台式计算机等，服务器端可以由具有处理功能的计算机设备来实现，但并不局限于此。服务器端与用户终端可以通过蓝牙、USB(Universal Serial Bus，通用串行总线)或者其他通讯连接方式进行网络连接，本申请在此不做限制。

10 在一个实施例中，图2为本申请实施例提供的语义分析模型准确度的校验方法的流程示意图，该语义分析模型准确度的校验方法可以应用于上述的服务器端，包括如下步骤：

 步骤 S210，获取项目信息，从所述项目信息中提取关键词，按照词义属性将所述关键词划分为测试关键词和期望关键词，得到测试数据集和
15 期望数据集；

 步骤 S220，分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集；

 步骤 S230，根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题语句对应的参考答案，并将所述问题语句与参考答案进行关
20 联；

 步骤 S240，将所述问题语句集中的各个问题语句输入语义分析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度。

 自然语言处理 (Natural Language Processing, NLP)，是计算机科学、
25 人工智能、语言学关注计算机和人类自然语言之间的相互作用的领域，研究人与计算机之间用自然语言进行有效通信的各种理论和方法。语义分析模型应用于自然语言处理领域，能够实现计算机识别人类自然语言。

 本申请方案的实施基础是已经获得项目信息，基于该项目信息进行后续处理，该项目信息可以是某个科研项目或研发项目，该项目信息中可以

包含文字、图表等形式表示的信息，从这些项目信息中提取关键词，按照关键词的词义属性，将提取出的关键词进行分类，首先将关键词划分为测试关键词和期望关键词，测试关键词，如：A 公司、B 行业、男女、售价、入职等等，期望关键词，如：数量、比例、日期等。

5 分别从测试数据集和期望数据集抽取测试关键词和期望关键词，组成问题语句，如：男女比例，抽取的方式多样，可以随机抽取，也可以按照预设规则进行抽取，从项目信息中抽取大量关键词，对关键词进行组合生成多个问题语句，组成问题语句集合，根据提取出来的问题语句解析项目信息，获得各个问题语句的参考答案，并将问题语句对应的参考答案进行
10 关联，依次提取问题语句集中的问题语句作为语义分析模型的输入，获得模型输出的识别答案，对比同一问题的参考答案与识别答案，若识别答案与参考答案一致，则为识别正确，若识别答案与参考答案不一致，则识别错误，根据正确识别的问题占全部问题的比例，获得语义分析模型的识别准确度。

15 本申请实施例提供的语义分析模型准确度的校验方案，提取项目信息中的关键词，对抽取出的关键词进行组合生成大量问题语句，将大量的问题语句作为语义分析模型的测试问题，进行语义分析模型的识别准确度校验。本方案基于项目信息提取大量关键词，不同关键词的组合能够生成大批量的问题语句，基于整个项目信息的关键词获得的问题语句的覆盖面广，
20 问题语句集的问题数量丰富，有利于提高模型校验准确性的同时也提高了校验效率。

为了更清楚本申请提供的语义分析模型准确度的校验方案及其技术效果，接下来以多个实施例对其具体方案进行详细阐述。

步骤 S220 的分别从所述测试数据集和期望数据集中抽取测试关键词
25 和期望关键词合成问题语句，得到问题语句集，可以通过如下方式获得，其流程示意图如图 3 所示，包括如下子步骤：

S310，从测试数据集中抽取至少一个测试关键词，从期望数据集中抽取一个期望关键词；

S320，将抽取出来的所述测试关键词与期望关键词进行组合，生成问

题语句；

S330，将生成的问题语句集合起来形成问题语句集。

其中，将抽取出来的测试关键词与期望关键词进行组合生成问题语句，可以通过至少两种方式进行，其一，随机从所述测试数据集中抽取至少一个测试关键词，随机从所述期望数据集中抽取一个期望关键词，将抽取出来的测试关键词和期望关键词进行随机组合，生成问题语句。

其中，测试数据集中可抽取至少一个测试关键词，可以为两个或两个以上的测试关键词，如抽取两个测试关键词：A 公司、男女，或抽取三个测试关键词：A 公司、男女、近三个月，即使抽取出的期望关键词为同一个，也至少对应两个不同的问题语句。

其二，通过预先设定抽取规则，如按照词义属性，将测试数据集划分为多个测试子集，依次从一个测试子集中抽取一个关键词，组成问题语句，具体如下：

步骤 S220 的分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，可以通过如下方式获得，其流程示意图如图 4 所示，包括如下子步骤：

S410，将测试数据集划分为第一测试数据集及第二测试数据集；

S420，依次抽取第一测试数据集中的第一测试关键词、第二测试数据集中的第二测试关键词及期望关键词生成问题语句。

将测试数据集划分为多个数据种类，可以划分为：来源信息，如文件名、公司名或行业名，时间信息，事件信息等等，如项目信息有如下信息：A 公司于 2018 年 3 月至 7 月的入职人数为 6 个，提取其中的关键词：A 公司、2018 年 3 月至 7 月、入职、人数，按照数据种类将测试数据集划分为来源信息集、时间信息集、事件信息集，其中，来源信息关键词为：A 公司，时间信息关键词为：2018 年 3 月至 7 月，事件信息关键词为：入职，期望关键词为：人数。依次提取多个测试数据分类集中的关键词及期望关键词，生成问题语句：A 公司 2018 年 3 月至 7 月入职人数？

值得说明的是，可以将测试数据集划分为三个测试数据集甚至更多，可以包含第三测试数据集、第四测试数据集等，划分的测试数据集越多，

表明测试数据划分的类别越详细，对应的参考答案越精准，解析项目信息获得参考答案的效率越高。

进一步地，将测试数据集划分为多个测试子集，将测试子集进行随机组合，并将随机组合的测试关键词与期望关键词进行随机组合。从项目相关信息中可以提取到大量关键词，测试数据集包括多个测试子集，不同测试子集之间进行随机组合，采用该种方式能够增大测试数据的组合数量，测试数据集与期望数据集之间进行随机组合，进一步增大了生成的问题语句的数量，按照这种方式生成的问题语句数量很容易达到百万数量级，满足模型训练或模型验证的需求。举例说明本方案：如测试子集有 3 个，每个测试子集中均设有 1 个测试关键词，其随机组合方式有 $(C_3^1 + C_3^2 + C_3^3)$ ，期望数据集中设有 3 个期望关键词，则问题数据组合有 $3 \times (C_3^1 + C_3^2 + C_3^3)$ ，按照该种组合方式，有限的关键词就能生成大量问题语句，且该种组合方式及问题语句的生成方式能够根据设定程序自动运行，降低问题生成过程中的人工成本及问题生成难度。

综上，本实施例提供的方案基于提取出的大量关键词采用数据集随机组合的方式使得构建的问题语句集合中问题语句数量呈爆炸式增长，获得问题语句的效率，且容易获得大量的测试样本，满足语义分析模型的识别准确率的验证需求。

进一步地，步骤 S220 的分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句的步骤之后，还包括：

S221，调用语法规则对所述问题语句进行语句润化处理，以使所述问题语句符合语法规则。

结合上一示例，提取测试关键词和期望关键词组成的问题语句是：A 公司 2018 年 3 月至 7 月入职人数？该语句的各关键词之间无连接词，不符合人的语法习惯，因此，需要将该问题语句进行语句润化，调用现有的语法规则，将连接词补入欠缺的位置，润化后的问题语句如：A 公司于 2018 年 3 月至 7 月的入职人数是多少？

润化之后的问题语句更加符合人的语法规则，避免问题语句产生歧义，进而导致无法利用该问题语句进行语义分析模型的识别准确性。

进一步地，步骤 S220 的得到问题语句集的步骤之后，还包括：S222，对所述问题语句集中的问题语句进行语义分析，剔除其中无意义的问题语句。

若问题语句的生成方式是通过对测试数据集与期望数据集的随机组合，可能会出现无意义的问题数据组合，如测试数据集中的测试关键词包括：A 公司、2018 年 3 月至 7 月、入职、男女，期望关键词：人数、比例，任意抽取上述测试数据集与期望数据集组成的问题语句：2018 年 3 月至 7 月比例，该问题语句由于缺乏必要的定语变得无意义，为了提高问题语句的质量，将该部分无意义的问题语句剔除。

在一种实施例中，在步骤 S220 的得到问题语句集的步骤之后，首先对问题语句进行润化处理，使问题语句符合语法规则，再对润化后的问题语句进行剔除处理，避免剔除由于不符合语法规则的问题语句，提高问题语句的筛选通过率。

步骤 S230 的根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题语句对应的参考答案，并将所述问题语句与参考答案进行关联，项目信息可以采用文字、图表等形式表示，解析文字、图表等形式表征的信息，提取其中的关键词及包含至少一个关键词标签的预设参考答案，如：项目信息中有如下文字记录的信息：甲公司 2017 年的入职人数是 6 个，提取其中的关键词：甲公司、2017 年、入职、人数，预设参考答案为：6 个，该预设参考答案的关键词标签为：甲公司、2017 年、入职、人数。根据提取出来的问题语句解析项目信息，如：问题语句中的关键词为：甲公司、2017 年、入职、人数，根据问题语句的关键词与预设参考答案的关键词标签进行对比，若问题语句的关键词与预设参考答案的关键词标签完全匹配，则该预设参考答案为该问题语句的参考答案，并将问题语句与对应的参考答案进行关联存储，以便后续调取该问题语句的参考答案，将其与问题语句的识别答案进行对比。

若解析项目信息并未获得问题语句对应的参考答案，也可以剔除该部分的问题语句，该种方案能够保证问题语句集中的问题语句都设有参考答案，提高了问题语句的质量，有利于获得准确的语义分析模型的识别准确

性。

一种实施例中，若遍历项目信息并未获得问题语句的参考答案，所述获取各个问题语句对应的参考答案的步骤，可以通过如下方式进行，S231，将该问题语句划归到未解问题语句集合中，为所述未解问题语句集中的问题语句设置统一参考答案。如设置的统一参考答案形如“这个问题超纲了”或“这是个好问题”等，为未解问题语句集中的问题语句设置统一答案，能够保证每个问题语句均对应应有参考答案，避免影响语义分析模型准确率的评估结果。

优选地，步骤 S230 的为所述未解问题语句集中的问题语句设置统一参考答案的步骤之前，还包括：对所述问题语句进行润化处理，对润化后的问题语句进行语义分析，保留润化处理后具有实际意义的问题语句。

经过该步骤之后，问题语句集中的问题语句均为符合语法规则且有意义的问题语句，对该部分问题语句进行参考答案的获取，对其中未解的问题语句设置统一参考答案，以便根据该问题语句进行语义分析模型准确度的测试。

值得说明的是，项目信息中对应应有参考答案的问题语句可以作为模型识别准确性的测试数据，项目信息中未有对应参考答案的问题语句，也可以作为语义分析模型的测试问题，若语义分析模型对未解问题语句集中的问题语句的识别答案是表征无正确答案的意思，则认为该语义分析模型的识别结果是正确的。

基于上述润化处理及过滤处理，本实施例提供了一种语义分析模型准确度的校验方案，其流程示意图如图 5 所示，步骤 S220 得到问题语句集之后，进行步骤 S221，调用语法规则对所述问题语句进行语句润化处理，以使所述问题语句符合语法规则，然后，再对润化后的问题语句进行 S222，对所述问题语句集中的问题语句进行语义分析，剔除其中无意义的问题语句，获得符合语法规则且有意义的问题语句，接下来进行步骤 S230 的根据所述问题语句集中的问题语句解析所述项目信息，判断项目信息中是否存储有问题语句对应的参考答案，若遍历项目信息并未获得问题语句对应的参考答案，则进行步骤 S231，将该问题语句划归到未解问题语句集合

中，为所述未解问题语句集中的问题语句设置统一参考答案。若项目信息中已存储有问题语句对应的参考答案或经过步骤 S231 设置有统一参考答案，将问题语句与参考答案进行关联，关联之后进行步骤 S240，将所述问题语句集中的各个问题语句输入语义分析模型件识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度。

本实施例提供的方案中，问题语句经过润化和过滤，使得问题语句均为符合语法规则且有意义的问题语句，为各问题语句设置参考答案时，为项目信息中并未存储的问题语句设置统一参考答案，以保证每个问题语句均能作为语义分析模型的测试问题，提高问题语句的质量，与剔除该部分问题语句相比，保证问题语句的数量规模，进而有利于获得准确的模型识别准确度，以及获得模型准确度校验的效率。

一种实施例中，将问题语句与参考答案进行关联之后，依次提取问题语句集中的各个问题语句，将所述问题语句输入语义分析模型中，获取输出的识别答案；调取与所述问题语句相关联的参考答案，对比所述参考答案与识别答案。步骤 S240 的将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度的步骤，包括：

若所述参考答案与识别答案一致，则标记为识别正确；若所述参考答案与识别答案不一致，则标记为识别错误；

统计识别正确的问题语句数量与问题语句集中问题语句的数量，计算所述识别正确的问题语句在问题语句集中的数量占比，获得语义分析模型的识别准确度。

本申请实施例提供的获得语义分析模型的识别准确度的方案，通过对参考答案与识别答案进行一致性判断，其中一致性判断为语义一致，如：识别答案为“6人”，若参考答案为“六人”，则表示参考答案与识别答案一致，若一致，则标记该问题语句为识别正确，统计识别正确的问题语句在问题语句集中的占比，该处问题语句集中的数量是均输入语义识别模型的问题语句的数量，本方案能够简单明了地获取语义分析模型的识别准确度。

以上为本申请提供的语义分析模型准确度的校验方法实施例，针对于该方法，下面阐述与其对应的语义分析模型准确度的校验装置的实施例。

本申请实施例还提供了一种语义分析模型准确度的校验装置，其结构示意图如图 6 所示，包括：提取关键词模块 610、生成问题语句模块 620、
5 获取参考答案模块 630、校验模块 640，具体如下：

提取关键词模块 610，用于获取项目信息，从所述项目信息中提取关键词，按照词义属性将所述关键词划分为测试关键词和期望关键词，得到测试数据集和期望数据集；

生成问题语句模块 620，用于分别从所述测试数据集和期望数据集中
10 抽取测试关键词和期望关键词合成问题语句，得到问题语句集；

获取参考答案模块 630，用于根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题语句对应的参考答案，并将所述问题语句与参考答案进行关联；

校验模块 640，用于将所述问题语句集中的各个问题语句输入语义分
15 析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度。

关于上述实施例中的语义分析模型准确度的校验装置，其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述，此处将不做详细阐述说明。

20 进一步地，本申请实施例还提供一种计算机可读非易失性存储介质，其上存储有计算机指令，该计算机指令被处理器执行时实现上述任意一项所述的语义分析模型准确度的校验方法的步骤。其中，所述非易失性存储介质包括但不限于任何类型的盘（包括软盘、硬盘、光盘、CD-ROM、和磁光盘）、ROM（Read-Only Memory，只读存储器）、RAM（Random Access
25 Memory，随机存储器）、EPROM（Erasable Programmable Read-Only Memory，可擦写可编程只读存储器）、EEPROM（Electrically Erasable Programmable Read-Only Memory，电可擦可编程只读存储器）、闪存、磁性卡片或光线卡片。也就是，非易失性存储介质包括由设备（例如，计算机）以能够读的形式存储或传输信息的任何介质。可以是只读存储器，磁

盘或光盘等。

更进一步地，本申请实施例还提供一种计算机设备，所述计算机设备包括：

一个或多个处理器；

5 存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现上述任意一项所述的语义分析模型准确度的校验方法的步骤。

图 7 是根据一示例性实施例示出的一种用于计算机设备 700 的框图。

10 例如，计算机设备 700 可以被提供为一服务器。参照图 7，计算机设备 700 包括处理组件 722，其进一步包括一个或多个处理器，以及由存储器 732 所代表的存储器资源，用于存储可由处理组件 722 的执行的指令，例如应用程序。存储器 732 中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外，处理组件 722 被配置为执行指令，以执行
15 上述语义分析模型准确度的校验方法的步骤。

计算机设备 700 还可以包括一个电源组件 726 被配置为执行计算机设备 700 的电源管理，一个有线或无线网络接口 750 被配置为将计算机设备 700 连接到网络，和一个输入输出（I/O）接口 758。计算机设备 700 可以操作基于存储在存储器 732 的操作系统，例如 Windows Server™，Mac OS
20 X™，Unix™，Linux™，FreeBSD™ 或类似。应该理解的是，虽然附图的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，其可以以其他的顺序执行。而且，附图的流程图中的至少一部分步骤可以包括多个子步骤或者多个阶段，这些子步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，
25 其执行顺序也不必然是依次进行，而是可以与其他步骤或者其他步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

应该理解的是，在本申请各实施例中的各功能单元可集成在一个处理模块中，也可以各个单元单独物理存在，也可以两个或两个以上单元集成

于一个模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。

以上所述仅是本申请的部分实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本申请原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也应视为本申请的保护范围。

5

权利要求书

1.一种语义分析模型准确度的校验方法，包括：

获取项目信息，从所述项目信息中提取关键词，按照词义属性将所述
5 关键词划分为测试关键词和期望关键词，得到测试数据集和期望数据集；

分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集；

根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题语句对应的参考答案，并将所述问题语句与参考答案进行关联；

10 将所述问题语句集中的各个问题语句输入语义分析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度。

2.根据权利要求1所述的语义分析模型准确度的校验方法，所述分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句的步骤，包括：
15

将测试数据集划分为第一测试数据集及第二测试数据集；

依次抽取第一测试数据集中的第一测试关键词、第二测试数据集中的第二测试关键词及期望关键词生成问题语句。

3.根据权利要求1所述的语义分析模型准确度的校验方法，所述分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句的步骤之后，还包括：
20

调用语法规则对所述问题语句进行语句润化处理，以使所述问题语句符合语法规则。

4.根据权利要求1或3所述的语义分析模型准确度的校验方法，所述得到问题语句集的步骤之后，还包括：
25

对所述问题语句集中的问题语句进行语义分析，剔除其中无意义的问题语句。

5.根据权利要求1所述的语义分析模型准确度的校验方法，所述分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问

题语句，得到问题语句集的步骤，包括：

从测试数据集中抽取至少一个测试关键词，从期望数据集中抽取一个期望关键词；

将抽取出来的所述测试关键词与期望关键词进行组合，生成问题语句；

5 将生成的问题语句集合起来形成问题语句集。

6.根据权利要求1所述的语义分析模型准确度的校验方法，若遍历项目信息并未获得问题语句的参考答案，所述获取各个问题语句对应的参考答案的步骤，包括：

10 将问题语句划归为未解问题语句集合中，为所述未解问题语句集中的问题语句设置统一参考答案。

7.根据权利要求1所述的语义分析模型准确度的校验方法，所述将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度的步骤，包括：

若所述参考答案与识别答案一致，则标记为识别正确；

15 分别统计识别正确的问题语句数量与问题语句集中问题语句的数量，计算识别正确的问题语句在问题语句集中的数量占比，获得语义分析模型的识别准确度。

8.一种语义分析模型准确度的校验装置，包括：

20 提取关键词模块，用于获取项目信息，从所述项目信息中提取关键词，按照词义属性将所述关键词划分为测试关键词和期望关键词，得到测试数据集和期望数据集；

生成问题语句模块，用于分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集；

25 获取参考答案模块，用于根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题语句对应的参考答案，并将所述问题语句与参考答案进行关联；

校验模块，用于将所述问题语句集中的各个问题语句输入语义分析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度。

9.一种计算机可读非易失性存储介质，所述计算机可读非易失性存储介质用于存储计算机指令，当其在计算机上运行时，使得计算机可以执行一种语义分析模型准确度的校验方法，所述语义分析模型准确度的校验方法包括以下步骤：

5 获取项目信息，从所述项目信息中提取关键词，按照词义属性将所述关键词划分为测试关键词和期望关键词，得到测试数据集和期望数据集；

 分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集；

 根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题
10 语句对应的参考答案，并将所述问题语句与参考答案进行关联；

 将所述问题语句集中的各个问题语句输入语义分析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度。

10.根据权利要求 9 所述的非易失性存储介质，所述分别从所述测试
15 数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句的步骤，包括：

 将测试数据集划分为第一测试数据集及第二测试数据集；

 依次抽取第一测试数据集中的第一测试关键词、第二测试数据集中的第二测试关键词及期望关键词生成问题语句。

20 11.根据权利要求 9 所述的非易失性存储介质，所述分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句的步骤之后，还包括：

 调用语法规则对所述问题语句进行语句润化处理，以使所述问题语句符合语法规则。

25 12.根据权利要求 9 或 11 所述的非易失性存储介质，所述得到问题语句集的步骤之后，还包括：

 对所述问题语句集中的问题语句进行语义分析，剔除其中无意义的问题语句。

13.根据权利要求 9 所述的非易失性存储介质，所述分别从所述测试

数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集的步骤，包括：

从测试数据集中抽取至少一个测试关键词，从期望数据集中抽取一个期望关键词；

- 5 将抽取出来的所述测试关键词与期望关键词进行组合，生成问题语句；
将生成的问题语句集合起来形成问题语句集。

14.根据权利要求 9 所述的非易失性存储介质，若遍历项目信息并未获得问题语句的参考答案，所述获取各个问题语句对应的参考答案的步骤，包括：

- 10 将问题语句划归为未解问题语句集合中，为所述未解问题语句集中的问题语句设置统一参考答案。

15.根据权利要求 9 所述的非易失性存储介质，所述将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度的步骤，包括：

- 15 若所述参考答案与识别答案一致，则标记为识别正确；

分别统计识别正确的问题语句数量与问题语句集中问题语句的数量，计算识别正确的问题语句在问题语句集中的数量占比，获得语义分析模型的识别准确度。

16.一种计算机设备，所述计算机设备包括：

- 20 一个或多个处理器；

存储装置，用于存储一个或多个程序，

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现一种语义分析模型准确度的校验方法，所述一种语义分析模型准确度的校验方法包括以下步骤：

- 25 获取项目信息，从所述项目信息中提取关键词，按照词义属性将所述关键词划分为测试关键词和期望关键词，得到测试数据集和期望数据集；

分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集；

根据所述问题语句集中的问题语句解析所述项目信息，获取各个问题

语句对应的参考答案，并将所述问题语句与参考答案进行关联；

将所述问题语句集中的各个问题语句输入语义分析模型进行识别，获取输出的识别答案，将所述问题语句所关联的参考答案与识别答案进行对比，获得语义分析模型的识别准确度。

- 5 17.根据权利要求 16 所述的计算机设备，所述分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句的步骤，包括：

将测试数据集划分为第一测试数据集及第二测试数据集；

依次抽取第一测试数据集中的第一测试关键词、第二测试数据集中的第二测试关键词及期望关键词生成问题语句。

- 10 18.根据权利要求 16 所述的计算机设备，所述分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句的步骤之后，还包括：

调用语法规则对所述问题语句进行语句润化处理，以使所述问题语句符合语法规则。

- 15 19.根据权利要求 16 或 18 所述的计算机设备，所述得到问题语句集的步骤之后，还包括：

对所述问题语句集中的问题语句进行语义分析，剔除其中无意义的问题语句。

- 20 20.根据权利要求 16 所述的计算机设备，所述分别从所述测试数据集和期望数据集中抽取测试关键词和期望关键词合成问题语句，得到问题语句集的步骤，包括：

从测试数据集中抽取至少一个测试关键词，从期望数据集中抽取一个期望关键词；

将抽取出来的所述测试关键词与期望关键词进行组合，生成问题语句；

- 25 将生成的问题语句集合起来形成问题语句集。

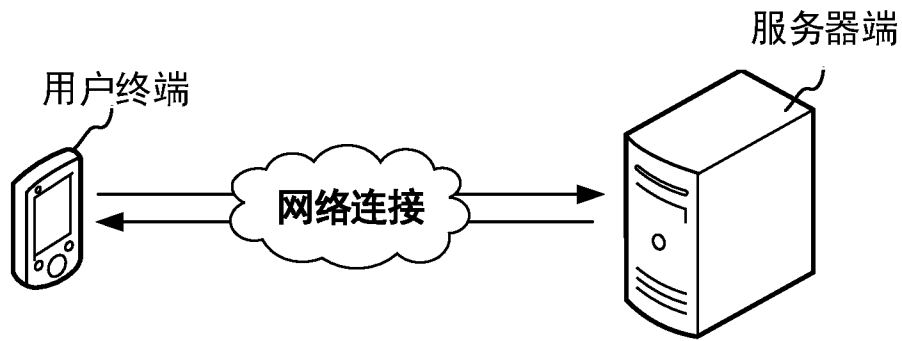


图 1

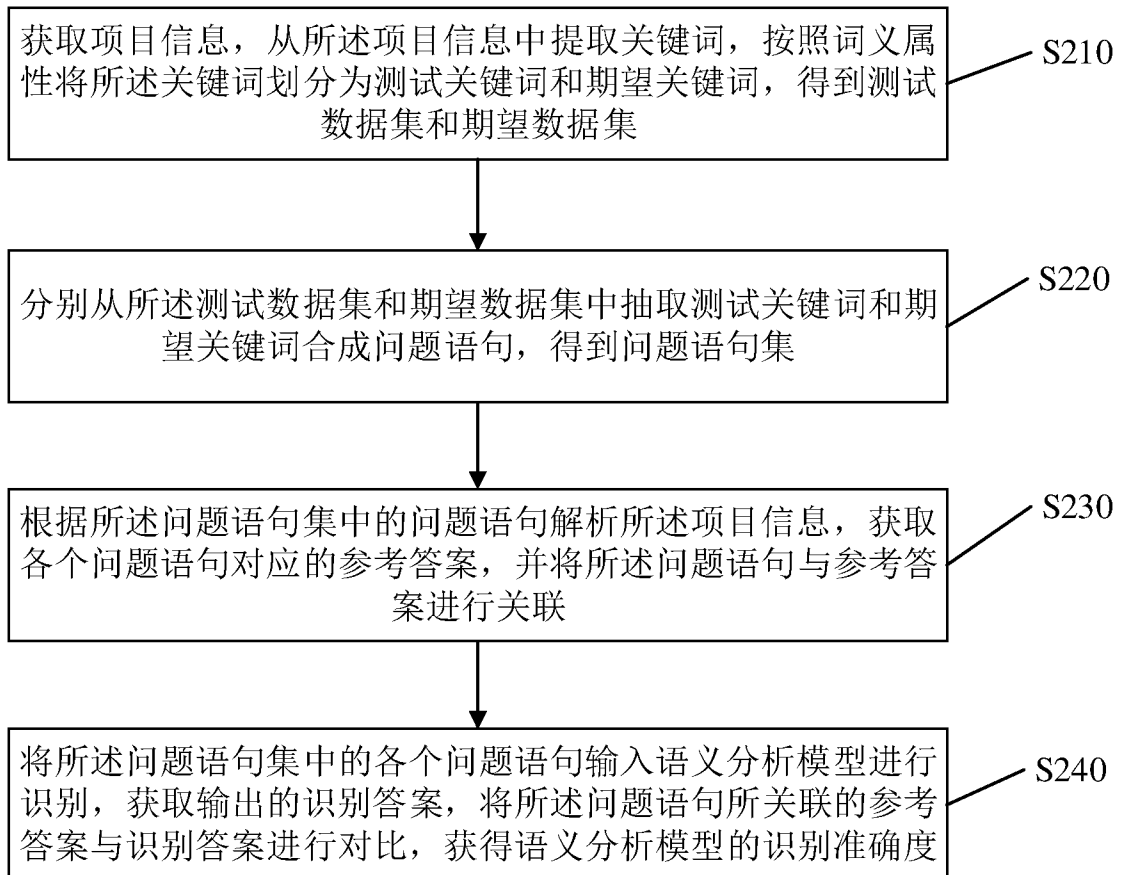


图 2

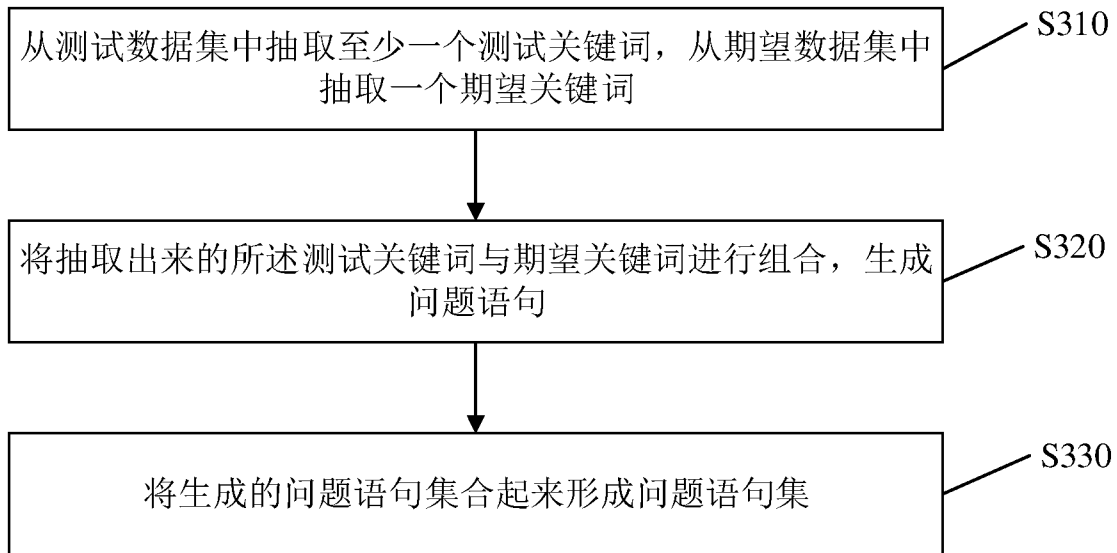


图 3

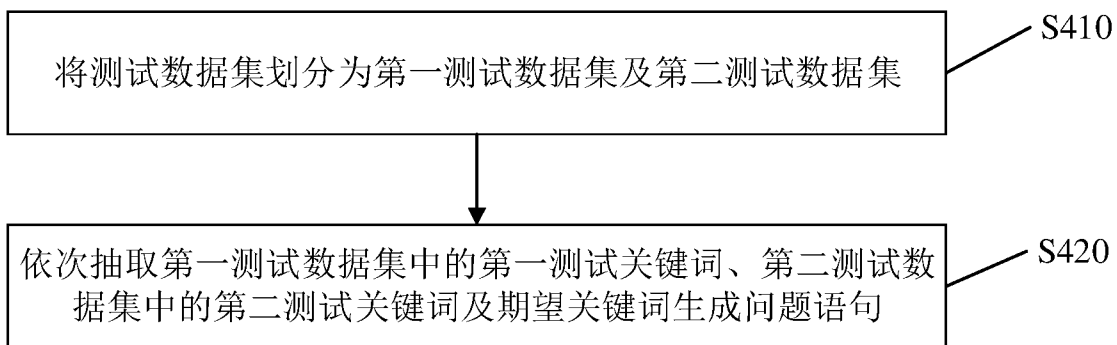


图 4

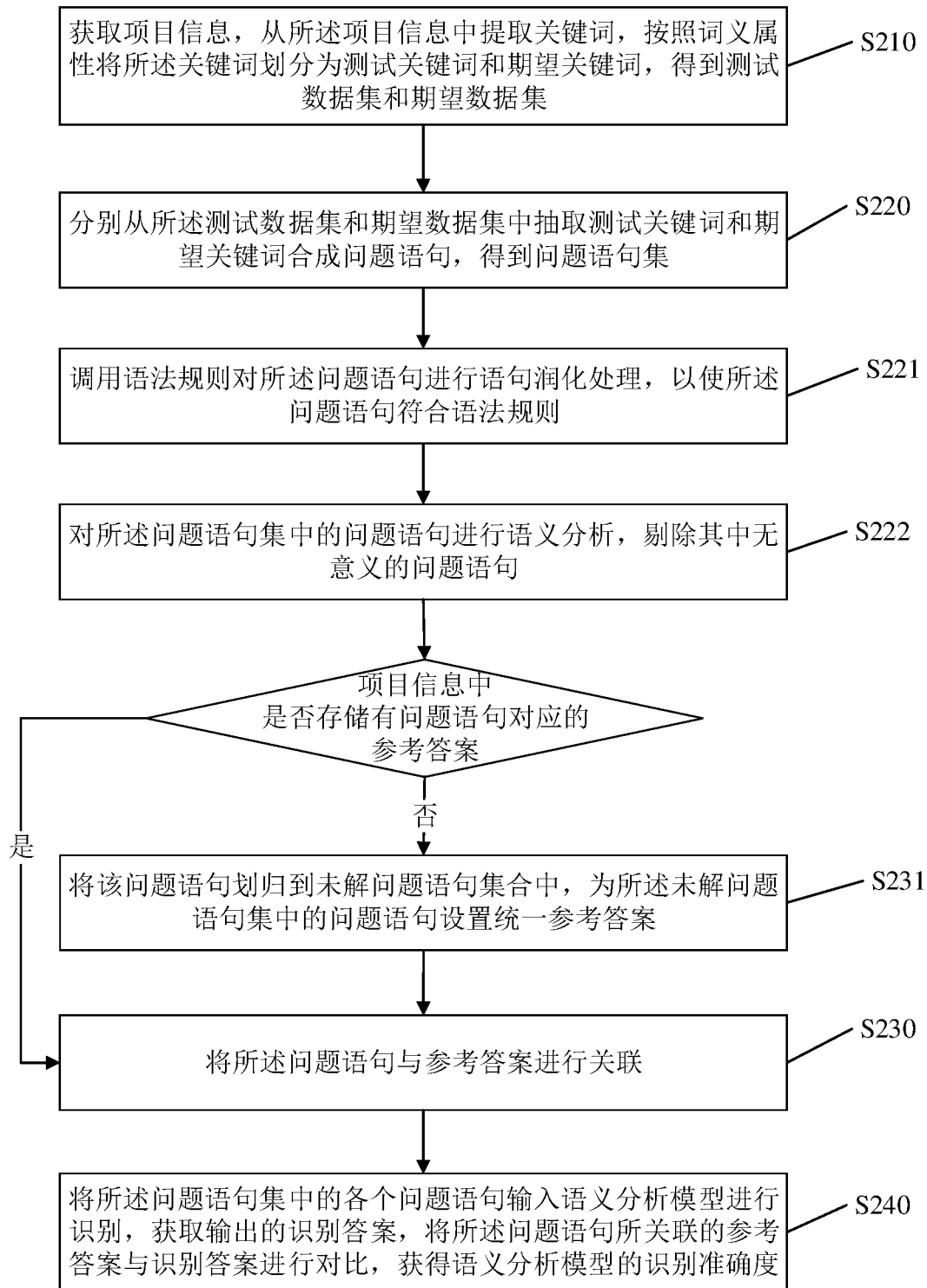


图 5

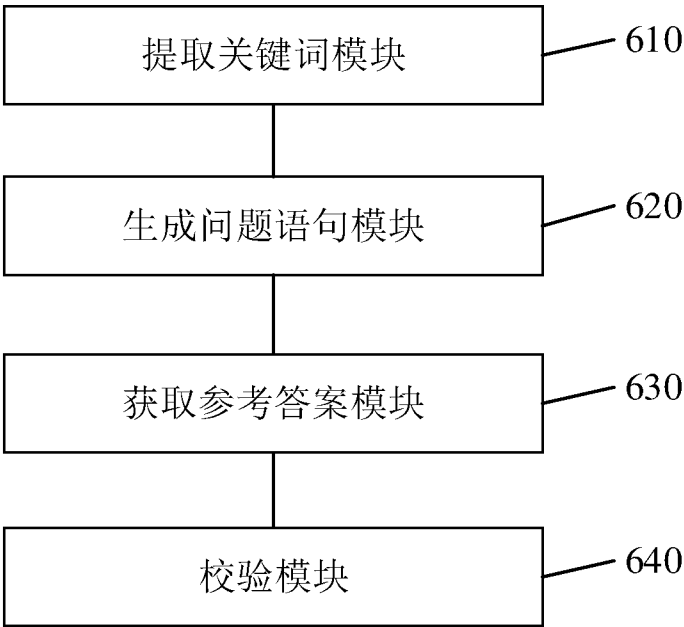


图 6

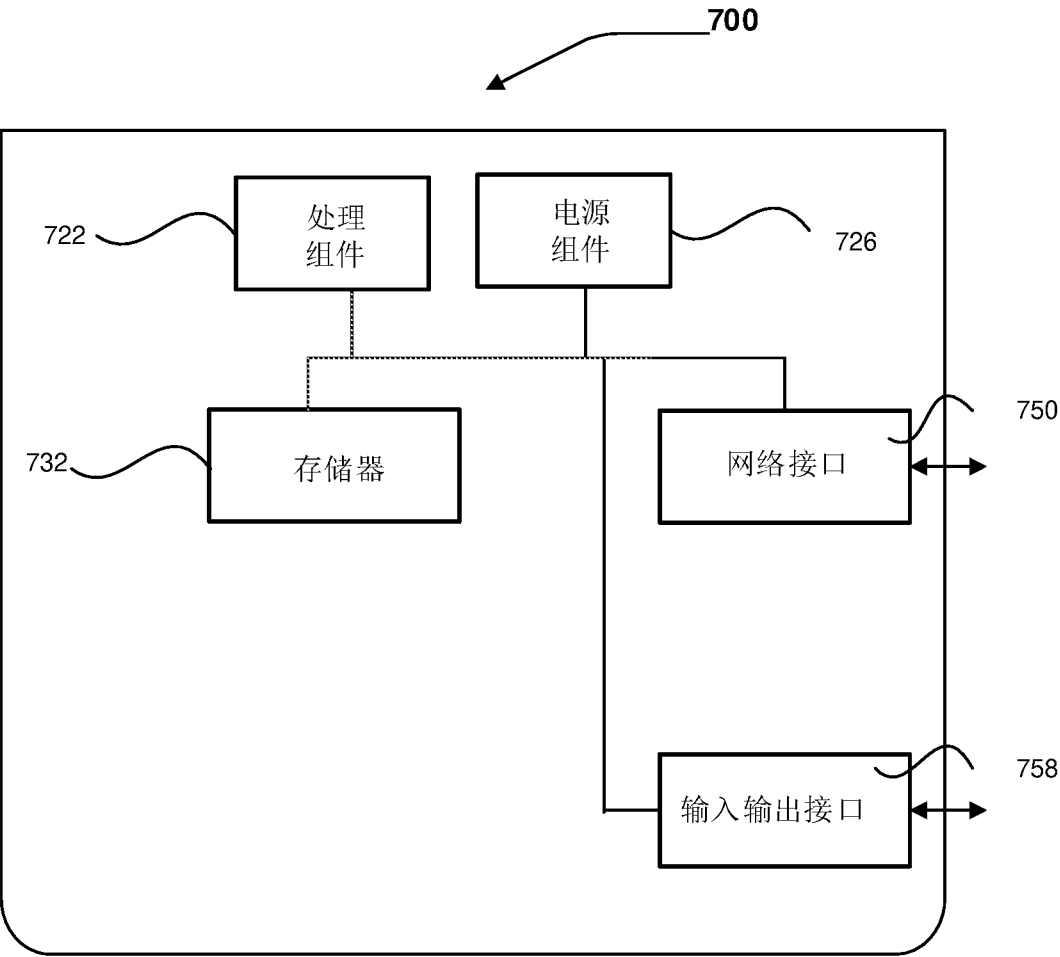


图 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/103024

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/35(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F16

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI, VEN, USTXT, EPTXT, WOTXT: 语义, 语句, 关键词, 答案, 答语, 识别, 准确度, 可靠度, semantic, sentence, keyword, answer, recognition, accuracy, reliability

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2017140290 A1 (IBM) 18 May 2017 (2017-05-18) claims 1-20 and description, paragraphs 21-112	1-20
X	US 2017140289 A1 (IBM) 18 May 2017 (2017-05-18) description, paragraphs 21-98 and claims 1-21	1-20
A	CN 108960574 A (BAIDU ONLINE NETWORK TECHNOLOGY BEIJING CO., LTD.) 07 December 2018 (2018-12-07) entire document	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

10 February 2020

Date of mailing of the international search report

02 March 2020

Name and mailing address of the ISA/CN

**China National Intellectual Property Administration (ISA/
CN)**
**No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
 100088**
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/103024

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)
US	2017140290	A1	18 May 2017	US	10282678 B2	07 May 2019
US	2017140289	A1	18 May 2017	None		
CN	108960574	A	07 December 2018	None		

国际检索报告

国际申请号

PCT/CN2019/103024

A. 主题的分类 G06F 16/35 (2019.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类														
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06F16 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) CNABS, CNTXT, CNKI, VEN, USTXT, EPTXT, WOTXT: 语义, 语句, 关键词, 答案, 答语, 识别, 准确度, 可靠度, semantic, sentence, keyword, answer, recognition, accuracy, reliability														
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>US 2017140290 A1 (IBM) 2017年 5月 18日 (2017 - 05 - 18) 权利要求1-20及说明书第21-112段</td> <td>1-20</td> </tr> <tr> <td>X</td> <td>US 2017140289 A1 (IBM) 2017年 5月 18日 (2017 - 05 - 18) 说明书第21-98段及权利要求1-21</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 108960574 A (百度在线网络技术北京有限公司) 2018年 12月 7日 (2018 - 12 - 07) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	US 2017140290 A1 (IBM) 2017年 5月 18日 (2017 - 05 - 18) 权利要求1-20及说明书第21-112段	1-20	X	US 2017140289 A1 (IBM) 2017年 5月 18日 (2017 - 05 - 18) 说明书第21-98段及权利要求1-21	1-20	A	CN 108960574 A (百度在线网络技术北京有限公司) 2018年 12月 7日 (2018 - 12 - 07) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求												
X	US 2017140290 A1 (IBM) 2017年 5月 18日 (2017 - 05 - 18) 权利要求1-20及说明书第21-112段	1-20												
X	US 2017140289 A1 (IBM) 2017年 5月 18日 (2017 - 05 - 18) 说明书第21-98段及权利要求1-21	1-20												
A	CN 108960574 A (百度在线网络技术北京有限公司) 2018年 12月 7日 (2018 - 12 - 07) 全文	1-20												
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。														
<table border="0"> <tr> <td> * 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 </td> <td> “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件 </td> </tr> </table>			* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件										
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件													
国际检索实际完成的日期 2020年 2月 10日		国际检索报告邮寄日期 2020年 3月 2日												
ISA/CN的名称和邮寄地址 中国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 王国海 电话号码 86-(20)-28958137												

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/103024

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
US	2017140290	A1	2017年 5月 18日	US 10282678 B2	2019年 5月 7日
US	2017140289	A1	2017年 5月 18日	无	
CN	108960574	A	2018年 12月 7日	无	