

(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
2021 年 8 月 19 日 (19.08.2021)



(10) 国际公布号
WO 2021/159656 A1

(51) 国际专利分类号:
G06F 40/30 (2020.01)

(21) 国际申请号: PCT/CN2020/098846

(22) 国际申请日: 2020 年 6 月 29 日 (29.06.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010088078.9 2020年2月12日 (12.02.2020) 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: 黄孟缘(HUANG, Mengyuan); 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: 广州三环专利商标代理有限公司(SCIHEAD IP LAW FIRM); 中国广东省广州市越秀区先烈中路 80 号汇华商贸大厦 1508 室, Guangdong 510070 (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,

(54) Title: METHOD, DEVICE, AND EQUIPMENT FOR SEMANTIC COMPLETION IN A MULTI-ROUND DIALOGUE, AND STORAGE MEDIUM

(54) 发明名称: 多轮对话中语义补齐的方法、装置、设备及存储介质

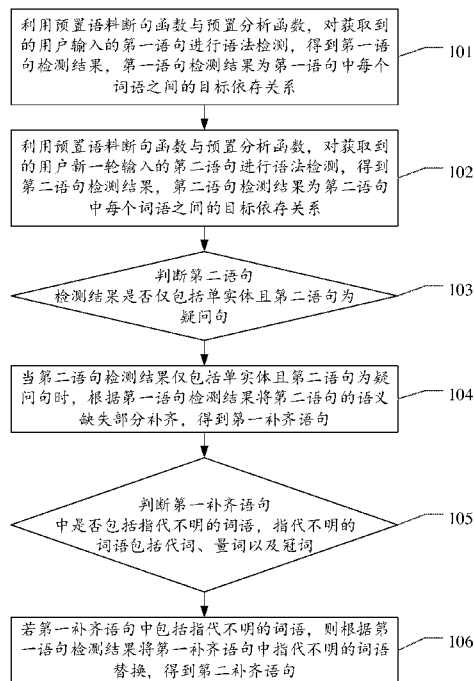


图 1

- 101 Utilize preset corpus segmentation function and preset analysis function to perform grammar check with respect to acquired first sentence inputted by user to produce first sentence check result, first sentence check result being target dependencies between every word in first sentence
- 102 Utilize preset corpus segmentation function and preset analysis function to perform grammar check with respect to acquired second sentence inputted by user in new round to produce second sentence check result, second sentence check result being target dependencies between every word in second sentence
- 103 Determine whether second sentence check result only comprises single entity and second sentence is question
- 104 When second sentence check result only comprises single entity and second sentence is question, complete semantically incomplete part of second sentence on the basis of first sentence check result to produce first completed sentence
- 105 Determine whether first completed sentence comprises unclear word, unclear word comprising pronouns, quantifiers, and articles
- 106 If first completed sentence comprises unclear word, then replace unclear word in first completed sentence on the basis of first sentence check result to produce second completed sentence

(57) Abstract: A method, device, and equipment for semantic completion in a multi-round dialogue, and a storage medium, grammar check is performed with respect to the multi-round dialogue via a preset corpus segmentation function and a preset analysis function, and semantically incomplete sentences are completed, thus increasing the accuracy of a semantic analysis result and the accuracy of searching for corresponding response information on the basis of the semantic analysis result. The method comprises: utilizing a preset corpus segmentation function and a preset analysis function to respectively perform grammar checks with respect to a first sentence

PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, 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)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

and a second sentence to produce a first sentence check result and a second sentence check result; when the second sentence check result comprises a single entity and the second sentence is a question, completing a semantically missing part of the second sentence on the basis of the first sentence check result to produce a first completed sentence (104); and if the first completed sentence comprises an unclear word, then replacing the unclear word in the first completed sentence on the basis of the first sentence check result to produce a second completed sentence (106).

(57) 摘要: 一种多轮对话中语义补齐的方法、装置、设备及存储介质, 通过预置语料断句函数与预置分析函数对多轮对话进行语法检测, 并补齐语义不全的语句, 提高了语义分析结果的准确率以及根据语义解析结果搜索对应的应答信息的准确率。上述方法包括: 利用预置语料断句函数与预置分析函数分别对第一语句和第二语句进行语法检测, 得到第一语句检测结果以及第二语句检测结果; 当第二语句检测结果包括单实体且第二语句为疑问句时, 根据第一语句检测结果将第二语句的语义缺失部分补齐, 得到第一补齐语句(104); 若第一补齐语句中包括指代不明的词语, 则根据第一语句检测结果将第一补齐语句中指代不明的词语替换, 得到第二补齐语句(106)。

多轮对话中语义补齐的方法、装置、设备及存储介质

本申请要求于2020年2月12日提交中国专利局、申请号为202010088078.9,发明名称为“多轮对话中语义补齐的方法、装置、设备及存储介质”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。

5 技术领域

本申请涉及人工智能领域,尤其涉及多轮对话中语义补齐的方法、装置、设备及存储介质。

背景技术

10 人机交互(human-computer interaction, HCI)是一门研究系统与用户之间的交互关系的学问。系统可以是各种各样的机器,也可以是计算机化的系统和软件。通常用户利用可见的人机交互界面与系统交流,并进行操作实现用户与系统之间的信息交换,以完成确定任务。人机交互是人工智能领域中重要的一部分,尤其是在客户服务或咨询获取方面,通过人机交互令用户获得需要的信息。

15 在客户服务方面,目前通过智能对话机器人与用户进行沟通,并利用智能对话机器人反馈数据来满足用户获取信息的需求。

但是发明人意识到,由于智能对话机器人获取到用户的提出要求存在语义不明的情况,如存在用户输入的要求缺少主语或宾语的情况,智能对话机器人无法准确识别出用户的意图,不能准确反馈出依据用户意图作出的应答信息,导致智能对话机器人反馈的准确率不高。

20 发明内容

本申请的提供一种多轮对话中语义补齐的方法、装置、设备及存储介质,用于解决基于多轮对话下,语句语义不全的问题,提高了语义分析结果的准确率,同时也提高了根据语义解析结果搜索对应的应答信息的准确率。

25 为本申请实施例的第一方面提供一种多轮对话中语义补齐的方法,包括:利用预置语料断句函数与预置分析函数,对获取到的用户输入的第一语句进行语法检测,得到第一语句检测结果,所述第一语句检测结果为所述第一语句中每个词语之间的目标依存关系;利用所述预置语料断句函数与所述预置分析函数,对获取到的用户新一轮输入的第二语句进行语法检测,得到第二语句检测结果,所述第二语句检测结果为所述第二语句中每个词语之间的目标依存关系;判断所述第二语句检测结果是否仅包括单实体且所述第二语句为疑问句;当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时,根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐,得到第一补齐语句;判断所述第一补齐语句中是否包括指代不明的词语,所述指代不明的词语包括代词、量词以及冠词;若所述第一补齐语句中包括指代不明的词语,则根据所述第一语句检测结果将所述第一补齐语句中指代不明的词语替换,得到第二补齐语句。

35 本申请实施例的第二方面提供一种多轮对话中语义补齐的装置,包括:第一获取单元,用于利用预置语料断句函数与预置分析函数,对获取到的用户输入的第一语句进行语法检测,得到第一语句检测结果,所述第一语句检测结果为所述第一语句中每个词语之间的目标依存关系;第二获取单元,用于利用所述预置语料断句函数与所述预置分析函数,对获取到的用户新一轮输入的第二语句进行语法检测,得到第二语句检测结果,所述第二语句检测结果为所述第二语句中每个词语之间的目标依存关系;第一判断单元,用于判断所述第二语句检测结果是否仅包括单实体且所述第二语句为疑问句;第一补齐单元,用于当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时,根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐,得到第一补齐语句;第二判断单元,用于判断所述第一补齐语句中是否包括指代不明的词语,所述指代不明的词语包括代词、量词以

40

及冠词；第二补齐单元，用于若所述第一补齐语句中包括指代不明的词语，则根据所述第一语句检测结果将所述第一补齐语句中指代不明的词语替换，得到第二补齐语句。

本申请实施例的第三方面提供了一种多轮对话中语义补齐的设备，包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现上述任一实施方式所述的基于多轮对话中语义补齐的方法，例如，实现以下步骤：利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一语句中每个词语之间的目标依存关系；利用所述预置语料断句函数与所述预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，所述第二语句检测结果为所述第二语句中每个词语之间的目标依存关系；判断所述第二语句检测结果是否仅包括单实体且所述第二语句为疑问句；当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐，得到第一补齐语句；判断所述第一补齐语句中是否包括指代不明的词语，所述指代不明的词语包括代词、量词以及冠词；若所述第一补齐语句中包括指代不明的词语，则根据所述第一语句检测结果将所述第一补齐语句中指代不明的词语替换，得到第二补齐语句。

本申请实施例的第四方面提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有指令，当其在计算机上运行时，使得计算机执行上述第一方面所述的方法，例如，实现以下步骤：利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一语句中每个词语之间的目标依存关系；利用所述预置语料断句函数与所述预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，所述第二语句检测结果为所述第二语句中每个词语之间的目标依存关系；判断所述第二语句检测结果是否仅包括单实体且所述第二语句为疑问句；当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐，得到第一补齐语句；判断所述第一补齐语句中是否包括指代不明的词语，所述指代不明的词语包括代词、量词以及冠词；若所述第一补齐语句中包括指代不明的词语，则根据所述第一语句检测结果将所述第一补齐语句中指代不明的词语替换，得到第二补齐语句。

本申请实施例提高了语义分析结果的准确率，同时也提高了根据语义解析结果搜索对应的应答信息的准确率。

附图说明

图1为本申请多轮对话中语义补齐的方法的一个实施例示意图；

图2为本申请多轮对话中语义补齐的方法的另一个实施例示意图；

图3为本申请多轮对话中语义补齐的装置的一个实施例示意图；

图4为本申请多轮对话中语义补齐的装置的另一个实施例示意图；

图5为本申请多轮对话中语义补齐的设备的一个实施例示意图。

具体实施方式

本申请的提供一种多轮对话中语义补齐的方法，用于解决基于多轮对话下，语句语义不全的问题，提高了语义分析结果的准确率，同时也提高了根据语义解析结果搜索对应的应答信息的准确率。

本申请的技术方案可应用于人工智能领域或大数据技术领域，例如本申请的技术方案可通过数据平台如云计算平台实现。

为了使本技术领域的人员更好地理解本申请方案，下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行描述。

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”、“第四”

等（如果存在）是用于区别类似的对象，而不必用于描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便这里描述的实施例能够以除了在这里图示或描述的内容以外的顺序实施。此外，术语“包括”或“具有”及其任何变形，意图在于覆盖不排除他的包含，例如，包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元，而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

请参阅图 1，本申请实施例中多轮对话中语义补齐的方法一个实施例包括：

101、利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，第一语句检测结果为第一语句中每个词语之间的目标依存关系。

服务器利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，第一语句检测结果为第一语句中每个词语之间的目标依存关系。

需要说明的是，服务器利用预置语料断句函数对第一语句进行断句，以及利用预置分析函数对第一输入语句进行语法检测，均为补齐第二语句语义做准备。服务器利用预置语料断句函数对第一语句进行初步断句，断句是自然语言处理的基础，断句的准确度直接决定了服务器后面进行句子的词性标注、句法分析、词向量以及文本分析的质量。对于简单的第一语句可以直接用逗号进行断句，服务器将第一语句中的字，按照在语句中的位置进行标注，举例来说：对句子“我爱深圳因为它很美”进行标注，标注的结果为：“我 N 爱 N 深 N 圳 Y 因 N 为 N 它 N 很 N 美 N”，其中 Y 代表断句，N 代表非断句，可以视为服务器对第一语句中每个字与字之间的位置做二分类，判断第一语句是否应该断句。

需要说明的是，这里预置分析函数指的是依存分析函数，服务器利用依存分析函数对第一语句进行语法检测，依存分析函数是利用语句中单词之间的依存关系来表达语法，如：一个词语修饰另一个词语，则称该词语依赖于另一个词语，在句子中明确了词语之间的关系，服务器可以准确的判断出语句中每个词语的词性以及词语之间的关系，为对语句进行补齐做充分的准备。服务器在进行依存分析前，需要对语句进行词性标注，服务器将第一语句中的字，按照在词语中的位置进行标注，可以利用 BMES 标注法，其中，B 为一个词语开始的一个字，M 为一个词语中间的字，词语中间的字数可能为多个，E 为一个词语最后的一个字，S 为由一个字组成的词语。举例来说，将“网商银行是蚂蚁金服微贷事业部的最重要产品”进行标注，标注后结果为“BMMESBMMEBMMMESBMEBE”，对应的分词结果为“网商银行/是/蚂蚁金服/微贷事业部/的/最重要/产品”，在预置词语库中查找对应分词的词性，将分词结果进行词性标注，当句子中出现单实体（句子中仅包括主语、谓语或宾语）时，将单实体抽取出来并保存在第一语句检测结果中以备。

102、利用预置语料断句函数与预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，第二语句检测结果为第二语句中每个词语之间的目标依存关系。

服务器利用预置语料断句函数与预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，第二语句检测结果为第二语句中每个词语之间的目标依存关系。

这里的第二语句是第一语句的下文语句，第一语句作为补齐第二语句语义缺失部分的依据。这里对第二语句进行语法检测的方式与对第一语句进行语法检测的方式是相同的，将检测的结果作为第二语句检测结果。

103、判断第二语句检测结果是否仅包括单实体且第二语句为疑问句。

服务器判断第二语句检测结果是否仅包括单实体且第二语句为疑问句。

可以理解的是，用户新一轮输入的第二语句可能出现语句部分省略的情况，比如省略谓语与宾语，这时就需要服务器对新一轮输入的第二语句进行判断，若新一轮输入的第二语句中语义表达不全，则根据多轮对话中的上下文情景将新一轮输入的第二语句补充完整，因此，服务器首先判断第二语句分析结果中是否仅包含单实体且第二语句为疑问句，以此来判断新一轮输入的第二语句是否满足语义不全的条件。

104、当第二语句检测结果仅包括单实体且第二语句为疑问句时，根据第一语句检测结果将第二语句的语义缺失部分补齐，得到第一补齐语句。

当第二语句检测结果仅包括单实体且第二语句为疑问句时，服务器根据第一语句检测结果将第二语句的语义缺失部分补齐，得到第一补齐语句。

若第二语句检测结果中包含单实体且第二语句为疑问句，则服务器不能准确识别第二语句的含义，无法反馈相应的预置回应，因此服务器根据获取到的上下文情景将第二语句补齐，得到一个完整的第二语句的语义，服务器再根据第二语句的语义反馈相应的回答。服务器获取到第一语句检测结果与第二语句检测结果，两个检测结果中均包括对句子的词性分析，提取出第二语句中的单实体，作为第一目标词语，这里的单实体指的是句子中的单个字或单个词组，例如：句子“感冒呢？”，其中“感冒”一词即为该句子中的单实体。待提取出第二语句中的单实体后，在第一语句中筛选与第一目标词语语法结构相同的单实体，作为第二目标词语，这里的语法结构相同指的是第一目标词语与第二目标词语的词性相同，例如：第一目标词语为“感冒”，第二目标词语为“癌症”，第一目标词语与第二目标词语均为名词词性，可以说两者的语法结构相同。待服务器筛选出第二目标词语后，在第一语句中，服务器将第一目标词语替换为第二目标词语，得到第一补齐语句。

105、判断第一补齐语句中是否包括指代不明的词语，指代不明的词语包括代词、量词以及冠词。

服务器判断第一补齐语句中是否包括指代不明的词语，指代不明的词语包括代词、量词以及冠词。

需要说明的是，当服务器补齐新一轮输入的第二语句且得到第一补齐语句时，还要确定第一补齐语句中是否包括指代不明的词语，指代不明的词语会导致第二语句的语义不明，一般来讲，指代不明的词语包括代词、量词以及冠词。举例来说：用户的第一语句为“平安福、福满分都是医疗险么？”服务器回复第一语句为“是的。”用户新一轮输入的第二语句为“第一种能报销什么？”在用户新一轮输入的第二语句中存在指代不明的词语“第一种”，如果服务器不联系多轮对话的上下文来对用户新一轮输入的第二语句进行补充，则会导致语义理解偏差，服务器将不能够给出正确的回答，导致用户不能够继续进行下一步的操作。

106、若第一补齐语句中包括指代不明的词语，则根据第一语句检测结果将第一补齐语句中指代不明的词语替换，得到第二补齐语句。

若第一补齐语句中包括指代不明的词语，则服务器根据第一语句检测结果将第一补齐语句中指代不明的词语替换，得到第二补齐语句。

服务器若确定第一补齐语句中包括指代不明的词语，说明第一补齐语句为语义不明的语句，则服务器需要根据多轮对话中的上下文情景对第一补齐语句进行再次补齐。首先，服务器在将第二语句中提取出指代不明的词语，作为第三目标词语，其次，在第一语句中筛选与第三目标词语词性相同的词语，作为第四目标词语，最后，在第二语句中，将第四目标词语替换为第三目标词语。

本申请实施例，通过利用预置语料断句函数与预置分析函数对用户输入的多轮对话进行语法检测，以及利用多轮对话上下文的语义对语义不全的语句进行补齐，提高了语义分析结果的准确率，同时也提高了根据语义解析结果搜索对应的应答信息的准确率。

请参阅图2，本申请实施例中多轮对话中语义补齐的方法另一个实施例包括：

201、获取用户输入的第一语句。

服务器获取用户输入的第一语句。在人机交互中，服务器获取用户输入的语句，并将用户输入的语句作为第一语句，为下文第二语句的补齐做准备。

202、利用预置语料断句函数将第一语句进行语句分割，得到第一输入语句。

5 服务器利用预置语料断句函数将第一语句进行语句分割，得到第一输入语句。具体的，服务器对第一语句进行断句，得到分段语句；服务器将分段语句中的分段语料与预置语料进行匹配，预置语料是依据业务数据建立在预置意图规则库中的语料；若分段语料与预置语料相匹配，则服务器在分段语料的前后位置对分段语句进行分割，得到分割语句，并将分割语句作为第一输入语句；若分段语料与预置语料不相配，则服务器直接将分段语料作为第一输入语句。

10 需要说明的是，服务器利用预置语料断句函数对第一语句进行断句，以及利用预置分析函数对第一输入语句进行语法检测，均为补齐第二语句语义做准备。服务器利用预置语料断句函数对第一语句进行初步断句，断句是自然语言处理的基础，断句的准确度直接决定服务器后面进行词性标注、句法分析、词向量以及文本分析的质量。对于简单的第一语句可直接用逗号进行断句，服务器将第一语句中的字，按照在语句中的位置进行标注，举例来说：对句子“我爱深圳因为它很美”进行标注，标注的结果为：“我N爱N深N圳Y因N为N它N很N美N”，其中Y代表断句，N代表非断句，可视为服务器对第一语句中每个字与字之间的位置做二分类，判断是否应该断句。

15 服务器再将第一语句进行断句后，得到分开的分段语句，再将分段语句中的分段语料与预置语料进行匹配，这里的分段语句是不能够在进行语句断句的语句，因此进一步对分段语句中的分段语料进行匹配，明确相应的语料是否需要进一步进行分割。这里的预置语料是依据业务数据建立在预置意图规则库中的语料，如：预置意图规则库为保险库，预置语料可为保单、被保人身份证号等，是服务器依据相关业务数据构成语料。若分段语料与预置语料相匹配，则服务器在分段语料的前后位置对分段语句进行分割，得到分割语句，并将分割语句作为第一输入语句；若分段语料与预置语料不相配，则服务器直接将分段语料作为第一输入语句。举例来说：对句子“感冒可以投保e生保么？”进行分段语料匹配后分割成“感冒/可以/投保/e生保/么？”，其中“感冒”、“投保”、“e生保”均为预置意图规则库中的预置语料，服务器不用再对匹配到预置语料的分段语料进行分割，直接在匹配到的分段语料的前后位置上进行分割，节省了服务器对语句分割的时间。

20 203、利用预置分析函数对第一输入语句进行语法检测，得到第一语句检测结果，第一语句检测结果为第一输入语句中每个词语之间的目标依存关系。

25 服务器利用预置分析函数对第一输入语句进行语法检测，得到第一语句检测结果，第一语句检测结果为第一输入语句中每个词语之间的目标依存关系。具体的，服务器对第一输入语句中的词语进行词性标注以及实体抽取，得到第一语句标注结果；服务器计算每个词语之间的依存概率，词语为第一语句标注结果中的词语，依存概率为预置依存关系出现的频次；服务器确定每个词语之间的目标依存关系，依存概率权重最大的概率所对应的预置依存关系为词语之间的目标依存关系；服务器获取第一语句检测结果，第一语句检测结果为第一输入语句中每个词语之间的目标依存关系。

30 需要说明的是，这里预置分析函数指的是依存分析函数，服务器利用依存分析函数对第一语句进行语法检测，依存分析函数是利用语句中单词之间的依存关系来表达语法，如：一个词语修饰另一个词语，则称该词语依赖于另一个词语，在句子中明确了词语之间的关系，服务器可以准确的判断出语句中每个词语的词性以及词语之间的关系，为对语句进行补齐做充分的准备。服务器进行依存分析前，需要对语句进行词性标注，服务器将第一语句中的字，按照在词语中的位置进行标注，可以利用 BMES 标注法，其中，B 为一个词语

开始的一个字，M 为一个词语中间的字，词语中间的字数可能为多个，E 为一个词语最后的一个字，S 为由一个字组成的词语。举例来说，将“网商银行是蚂蚁金服微贷事业部的最重要产品”进行标注，标注后结果为“BMMESBMMEBMMMMESBMEBE”，对应的分词结果为“网商银行/是/蚂蚁金服/微贷事业部/的/最重要/产品”，在预置词语库中查找对应分词的词性，将分词结果进行词性标注，当句子中出现单实体（句子中仅包括主语、谓语或宾语）时，将单实体抽取出来并保存在第一语句检测结果中以备。

204、利用预置语料断句函数与预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，第二语句检测结果为第二语句中每个词语之间的目标依存关系。

服务器利用预置语料断句函数与预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，第二语句检测结果为第二语句中每个词语之间的目标依存关系。

这里的第二语句是第一语句的下文语句，第一语句作为补齐第二语句语义缺失部分的依据。这里对第二语句进行语法检测的方式与对第一语句进行语法检测的方式是相同的，将检测的结果作为第二语句检测结果。

205、判断第二语句检测结果是否仅包括单实体且第二语句为疑问句。

服务器判断第二语句检测结果是否仅包括单实体且第二语句为疑问句。

可以理解的是，用户新一轮输入的第二语句可能出现语句部分省略的情况，比如省略谓语与宾语，这时就需要服务器对新一轮输入的第二语句进行判断，若新一轮输入的第二语句中语义表达不全，则根据多轮对话中的上下文情景将新一轮输入的第二语句补充完整，因此，服务器首先判断第二语句分析结果中是否仅包含单实体且第二语句为疑问句，来判断新一轮输入的第二语句是否满足语义不全的条件。

206、当第二语句检测结果仅包括单实体且第二语句为疑问句时，根据第一语句检测结果将第二语句的语义缺失部分补齐，得到第一补齐语句。

当第二语句检测结果仅包括单实体且第二语句为疑问句时，服务器根据第一语句检测结果将第二语句的语义补全，得到第一补齐语句。具体的，当第二语句检测结果仅包括单实体且第二语句为疑问句时，服务器提取出第二语句检测结果中的单实体，并将单实体作为第一目标词语；服务器在第一语句检测结果中筛选与第一目标词语语法结构相同的单实体，得到第二目标词语；在第一语句中，服务器利用第一目标词语替换第二目标词语，得到第一补齐语句。

若第二语句检测结果中仅包含单实体且第二语句为疑问句，则服务器不能准确识别第二语句的含义，无法反馈相应的预置回应，因此服务器根据获取到的上下文情景将第二语句的语义缺失部分补齐，得到完整的第二语句的语义，服务器再根据第二语句的语义反馈相应的回答。服务器获取到第一语句检测结果与第二语句检测结果，两个检测结果中均包括对句子的词性分析，提取出第二语句中的单实体，作为第一目标词语，并在第一语句检测结果中筛选与第一目标词语语法结构相同的单实体，作为第二目标词语，在第一语句中，服务器将第一目标词语替换为第二目标词语，得到第一补齐语句。

举例来说：用户输入的第一语句为“感冒可以投保 e 生保么？”，服务器对用户输入的第一语句进行词性标注与依存分析，得到“感冒/nhd 可以/c 投保/vn e 生保/nbx 么/y？”这里的 nhd 以及 nbx 是预置语料的相关词性，例如：预置语料所在的预置意图规则库为保险库，nhd 代表保险库中的疾病一类词，nbx 则代表保险库中保险名称一类词；c 代表助动词；vn 代表动名词；y 代表语气助词，提取出来的单实体为：感冒/nhd 投保/vn e 生保/nbx。用户输入新一轮的第二语句为“癌症哪？”服务器对这句话进行解析，得到“癌症/nhd 哪/y”很明显第二语句检测结果中仅包括单实体且第二语句以反问语气结尾，满足第二语句语义

不全的条件，则服务器对第二语句进行语义补充，根据第一语句检测结果以及第二语句检测结果，得到感冒/nhd 与癌症/nhd 是同一类词语，将两者在第一语句中互换，得到“癌症可以投保 e 生保么？”将第二语句补齐。

207、判断第一补齐语句中是否包括指代不明的词语，指代不明的词语包括代词、量词以及冠词。

服务器判断第一补齐语句中是否包括指代不明的词语，指代不明的词语包括代词、量词以及冠词。具体的，服务器获取第一补齐语句的检测结果，第一补齐语句的检测结果为第二语句检测结果与第一语句检测结果的结合；服务器判断第一补齐语句的检测结果中是否包括代词、量词以及冠词。

需要说明的是，当服务器补齐新一轮输入的第二语句且得到第一补齐语句时，还要判断第一补齐语句中是否包括指代不明的词语，指代不明的词语会导致第二语句的语义不明，一般来讲，指代不明的词语包括代词、量词以及冠词。举例来说：用户的第一语句为“平安福、福满分都是医疗险么？”服务器回复第一语句为“是的。”用户新一轮输入的第二语句为“第一种能报销什么？”在用户新一轮输入的第二语句中存在指代不明的词语“第一种”，如果服务器不联系多轮对话的上下文来对用户新一轮输入的第二语句进行补充，则会导致语义理解偏差，服务器将不能够给出正确的回答，导致用户不能够继续进行下一步的操作。

208、若第一补齐语句中包括指代不明的词语，则根据第一语句检测结果将第一补齐语句中指代不明的词语替换，得到第二补齐语句。

若第一补齐语句中包括指代不明的词语，则服务器根据第一语句检测结果将第一补齐语句中指代不明的词语替换，得到第二补齐语句。具体的，若第一补齐语句的检测结果包括指代不明的词语，则服务器提取出指代不明的词语，并将指代不明的词语作为第三目标词语；服务器在第一语句检测结果中筛选与第三目标词语语法结构相同的词语，得到第四目标词语；在第二语句中，服务器利用第四目标词语替换第三目标词语，得到第二补齐语句。

服务器若确定第一补齐语句中包括指代不明的词语，说明第一补齐语句为语义不明的语句，则服务器需要根据多轮对话中的上下文情景对第一补齐语句再次补齐。首先，服务器在将第二语句中提取出指代不明的词语，作为第三目标词语，其次，在第一语句中筛选与第三目标词语词性相同的词语，作为第四目标词语，最后，在第二语句中将第四目标词语替换为第三目标词语。

举例来说：用户的第一语句为“平安福、福满分，都是医疗险么？”，服务器对用户输入的第一语句进行词性标注与依存分析，得到“平安福/nbx、福满分/nbx，是/v 医疗险/nbx 么/y？”这里的 nbx 是预置语料的相关词性，例如：预置语料所在的预置意图规则库为保险库，nbx 则代表保险库中保险名称一类词；v 代表动词；y 代表语气助词。用户新一轮输入的第二语句为“前者能报销什么？”，服务器对用户新一轮输入的第二语句进行词性标注与依存分析，得到“前者/r 能/v 报销/v 什么/y？”这里的 r 代表代词；v 代表动词；y 代表语气助词，在用户新一轮输入的第二语句中存在指代不明的词语“前者”，服务器联系多轮对话的上下文来对用户新一轮输入的第二语句进行补充，结合用户第一语句的检测结果，这里的前者指代的是“平安福”，将“平安福”直接带入第二语句中，得到“平安福能报销什么？”得到完整的第二补齐语句，使得服务器能够根据第二补齐语句进行反馈与回答。

本申请实施例，通过利用预置语料断句函数与预置分析函数对用户输入的多轮对话进行语法检测，以及利用多轮对话上下文的语义对语义不全的语句进行补齐，提高了语义分析结果的准确率，同时也提高了根据语义解析结果搜索对应的应答信息的准确率。

上面对本申请实施例中多轮对话中语义补齐的方法进行了描述，下面对本申请实施例中多轮对话中语义补齐的装置进行描述，请参阅图 3，本申请实施例中多轮对话中语义补

齐的装置一个实施例包括：

第一获取单元 301，用于利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，第一语句检测结果为第一语句中每个词语之间的目标依存关系；

5 第二获取单元 302，用于利用预置语料断句函数与预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，第二语句检测结果为第二语句中每个词语之间的目标依存关系；

第一判断单元 303，用于判断第二语句检测结果是否仅包括单实体且第二语句为疑问句；

10 第一补齐单元 304，用于当第二语句检测结果仅包括单实体且第二语句为疑问句时，根据第一语句检测结果将第二语句的语义缺失部分补齐，得到第一补齐语句；

第二判断单元 305，用于判断第一补齐语句中是否包括指代不明的词语，指代不明的词语包括代词、量词以及冠词；

15 第二补齐单元 306，用于若第一补齐语句中包括指代不明的词语，则根据第一语句检测结果将第一补齐语句中指代不明的词语替换，得到第二补齐语句。

本申请实施例中，第一获取单元 301 利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，第一语句检测结果为第一语句中每个词语之间的目标依存关系；第二获取单元 302 利用预置语料断句函数与预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，第二语句检测结果为第二语句中每个词语之间的目标依存关系；第一判断单元 303 判断第二语句检测结果是否仅包括单实体且第二语句为疑问句；第一补齐单元 304 当第二语句检测结果仅包括单实体且第二语句为疑问句时，根据第一语句检测结果将第二语句的语义缺失部分补齐，得到第一补齐语句；第二判断单元 305 判断第一补齐语句中是否包括指代不明的词语，指代不明的词语包括代词、量词以及冠词；第二补齐单元 306 若第一补齐语句中包括指代不明的词语，则根据第一语句检测结果将第一补齐语句中指代不明的词语替换，得到第二补齐语句。

本申请实施例，通过利用预置语料断句函数与预置分析函数对用户输入的多轮对话进行语法检测，以及利用多轮对话上下文的语义对语义不全的语句进行补齐，提高了语义分析结果的准确率，同时也提高了根据语义解析结果搜索对应的应答信息的准确率。

30 请参阅图 4，本申请实施例中多轮对话中语义补齐的装置另一个实施例包括：

第一获取单元 301，用于利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，第一语句检测结果为第一语句中每个词语之间的目标依存关系；

35 第二获取单元 302，用于利用预置语料断句函数与预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，第二语句检测结果为第二语句中每个词语之间的目标依存关系；

第一判断单元 303，用于判断第二语句检测结果是否仅包括单实体且第二语句为疑问句；

40 第一补齐单元 304，用于当第二语句检测结果仅包括单实体且第二语句为疑问句时，根据第一语句检测结果将第二语句的语义缺失部分补齐，得到第一补齐语句；

第二判断单元 305，用于判断第一补齐语句中是否包括指代不明的词语，指代不明的词语包括代词、量词以及冠词；

第二补齐单元 306，用于若第一补齐语句中包括指代不明的词语，则根据第一语句检测结果将第一补齐语句中指代不明的词语替换，得到第二补齐语句。

可选的, 第一获取单元 301 包括:

获取模块 3011, 用于获取用户输入的第一语句;

分割模块 3012, 用于利用预置语料断句函数将第一语句进行语句分割, 得到第一输入语句;

5 检测模块 3013, 用于利用预置分析函数对第一输入语句进行语法检测, 得到第一语句检测结果, 第一语句检测结果为第一输入语句中每个词语之间的目标依存关系。

可选的, 分割模块 3012 具体用于:

对第一语句进行断句, 得到分段语句;

10 将分段语句中的分段语料与预置语料进行匹配, 预置语料是依据业务数据建立在预置意图规则库中的语料;

若分段语料与预置语料相匹配, 则在分段语料的前后位置对分段语句进行分割, 得到分割语句, 并将分割语句作为第一输入语句;

若分段语料与预置语料不相配, 则直接将分段语料作为第一输入语句。

可选的, 检测模块 3013 具体用于:

15 对第一输入语句中的词语进行词性标注以及实体抽取, 得到第一语句标注结果;

计算每个词语之间的依存概率, 词语为第一语句标注结果中的词语, 依存概率为预置依存关系出现的频次;

确定每个词语之间的目标依存关系, 依存概率权重最大的概率所对应的预置依存关系为词语之间的目标依存关系;

20 获取第一语句检测结果, 第一语句检测结果为第一输入语句中每个词语之间的目标依存关系。

可选的, 第一补齐单元 304 具体用于:

当第二语句检测结果仅包括单实体且第二语句为疑问句时, 提取出第二语句检测结果中的单实体, 并将单实体作为第一目标词语;

25 在第一语句检测结果中筛选与第一目标词语语法结构相同的单实体, 得到第二目标词语;

在第一语句中, 利用第一目标词语替换第二目标词语, 得到第一补齐语句。

可选的, 第二判断单元 305 具体用于:

30 获取第一补齐语句的检测结果, 第一补齐语句的检测结果为第二语句检测结果与第一语句检测结果的结合;

判断第一补齐语句的检测结果中是否包括代词、量词以及冠词。

可选的, 第二补齐单元 306 具体用于:

若第一补齐语句的检测结果包括指代不明的词语, 则提取出指代不明的词语, 并将指代不明的词语作为第三目标词语;

35 在第一语句检测结果中筛选与第三目标词语语法结构相同的词语, 得到第四目标词语;

在第二语句中, 利用第四目标词语替换第三目标词语, 得到第二补齐语句。

40 本申请实施例中, 第一获取单元 301 包括: 获取模块 3011, 用于获取用户输入的第一语句; 分割模块 3012, 用于利用预置语料断句函数将第一语句进行语句分割, 得到第一输入语句; 检测模块 3013, 用于利用预置分析函数对第一输入语句进行语法检测, 得到第一语句检测结果, 第一语句检测结果为第一输入语句中每个词语之间的目标依存关系; 第二获取单元 302 利用预置语料断句函数与预置分析函数, 对获取到的用户新一轮输入的第二语句进行语法检测, 得到第二语句检测结果, 第二语句检测结果为第二语句中每个词语之间的目标依存关系; 第一判断单元 303 判断第二语句检测结果是否仅包括单实体且第二语句为疑问句; 第一补齐单元 304 当第二语句检测结果仅包括单实体且第二语句为疑问句时,

根据第一语句检测结果将第二语句的语义缺失部分补齐，得到第一补齐语句；第二判断单元 305 判断第一补齐语句中是否包括指代不明的词语，指代不明的词语包括代词、量词以及冠词；第二补齐单元 306 若第一补齐语句中包括指代不明的词语，则根据第一语句检测结果将第一补齐语句中指代不明的词语替换，得到第二补齐语句。

5 本申请实施例，通过利用预置语料断句函数与预置分析函数对用户输入的多轮对话进行语法检测，以及利用多轮对话上下文的语义对语义不全的语句进行补齐，提高了语义分析结果的准确率，同时也提高了根据语义解析结果搜索对应的应答信息的准确率。

10 上面图 3 至图 4 从模块化功能实体的角度对本申请实施例中的多轮对话中语义补齐的装置进行详细描述，下面从硬件处理的角度对本申请实施例中多轮对话中语义补齐的设备进行详细描述。

下面结合图 5 对多轮对话中语义补齐的设备的各个构成部件进行具体的介绍：

15 图 5 是本申请实施例提供的一种多轮对话中语义补齐的设备的结构示意图，该多轮对话中语义补齐的设备 500 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上处理器（central processing units, CPU）501（例如，一个或一个以上处理器）和存储器 509，一个或一个以上存储应用程序 507 或数据 506 的存储介质 508（例如一个或一个以上海量存储设备）。其中，存储器 509 和存储介质 508 可以是短暂存储或持久存储。存储在存储介质 508 的程序可以包括一个或一个以上模块（图示没标出），每个模块可以包括对多轮对话中语义补齐的设备中的一系列指令操作。更进一步地，处理器 501 可以设置为与存储介质 508 通信，在多轮对话中语义补齐的设备 500 上执行存储介质 508 中的一系列指令操作。

20 多轮对话中语义补齐的设备 500 还可以包括一个或一个以上电源 502，一个或一个以上有线或无线网络接口 503，一个或一个以上输入输出接口 504，和/或，一个或一个以上操作系统 505，例如 Windows Serve, Mac OS X, Unix, Linux, FreeBSD 等等。本领域技术人员可以理解，图 5 中示出的多轮对话中语义补齐的设备结构并不构成对多轮对话中语义补齐的设备的限定，可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。

下面结合图 5 对多轮对话中语义补齐的设备的各个构成部件进行具体的介绍：

30 处理器 501 是多轮对话中语义补齐的设备的控制中心，可以按照多轮对话中语义补齐的方法进行处理。处理器 501 利用各种接口和线路连接整个多轮对话中语义补齐的设备的各个部分，通过运行或执行存储在存储器 509 内的软件程序和/或模块，以及调用存储在存储器 509 内的数据，利用预置语料断句函数与预置分析函数对用户输入的多轮对话进行语法检测，以及利用多轮对话上下文的语义对语义不全的语句进行补齐，提高了语义分析结果的准确率。存储介质 508 和存储器 509 都是存储数据的载体，本申请实施例中，存储介质 508 可以是指储存容量较小，但速度快的内存储器，而存储器 509 可以是储存容量大，但储存速度慢的外存储器。

40 存储器 509 可用于存储软件程序以及模块，处理器 501 通过运行存储在存储器 509 的软件程序以及模块，从而执行多轮对话中语义补齐的设备 500 的各种功能应用以及数据处理。存储器 509 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；存储数据区可存储根据多轮对话中语义补齐的设备的使用所创建的数据等。此外，存储器 509 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，例如至少一个磁盘存储器件、闪存器件、或其他非易失性固态存储器件。在本申请实施例中提供的多轮对话中语义补齐的程序和接收到的数据流存储在存储器中，当需要使用时，处理器 501 从存储器 509 中调用。

在计算机上加载和执行计算机程序指令时，全部或部分地产生按照本申请实施例的流

程或功能。计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。计算机指令可以存储在计算机可读存储介质中，或者从一个计算机可读存储介质向另一计算机可读存储介质传输，例如，计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、双绞线）或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。计算机可读存储介质可以是计算机能够存储的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。可用介质可以是磁性介质，（例如，软盘、硬盘、磁带）、光介质（例如，光盘）、或者半导体介质（例如固态硬盘（solid state disk, SSD））等。

可选的，该计算机可读存储介质（或存储介质）可以是非易失性的，也可以是易失性的。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统，装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

以上，以上实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

权利要求

1、一种多轮对话中语义补齐的方法，其中，包括：

利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一语句中每个词语之间的目标依存关系；

利用所述预置语料断句函数与所述预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，所述第二语句检测结果为所述第二语句中每个词语之间的目标依存关系；

判断所述第二语句检测结果是否仅包括单实体且所述第二语句为疑问句；

当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐，得到第一补齐语句；

判断所述第一补齐语句中是否包括指代不明的词语，所述指代不明的词语包括代词、量词以及冠词；

若所述第一补齐语句中包括指代不明的词语，则根据所述第一语句检测结果将所述第一补齐语句中指代不明的词语替换，得到第二补齐语句。

2、根据权利要求1所述的方法，其中，所述利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一语句中每个词语之间的目标依存关系包括：

获取用户输入的第一语句；

利用预置语料断句函数将所述第一语句进行语句分割，得到第一输入语句；

利用预置分析函数对所述第一输入语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一输入语句中每个词语之间的目标依存关系。

3、根据权利要求2所述的方法，其中，所述利用预置语料断句函数将所述第一语句进行语句分割，得到第一输入语句包括：

对所述第一语句进行断句，得到分段语句；

将所述分段语句中的分段语料与预置语料进行匹配，预置语料是依据业务数据建立在预置意图规则库中的语料；

若所述分段语料与所述预置语料相匹配，则在所述分段语料的前后位置对所述分段语句进行分割，得到分割语句，并将所述分割语句作为第一输入语句；

若所述分段语料与所述预置语料不相配，则直接将所述分段语料作为第一输入语句。

4、根据权利要求2所述的方法，其中，所述利用预置分析函数对所述第一输入语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一输入语句中每个词语之间的目标依存关系包括：

对所述第一输入语句中的词语进行词性标注以及实体抽取，得到第一语句标注结果；

计算每个所述词语之间的依存概率，所述词语为所述第一语句标注结果中的词语，所述依存概率为预置依存关系出现的频次；

确定每个所述词语之间的目标依存关系，所述依存概率权重最大的概率所对应的所述预置依存关系为所述词语之间的所述目标依存关系；

获取第一语句检测结果，所述第一语句检测结果为所述第一输入语句中每个所述词语之间的所述目标依存关系。

5、根据权利要求1所述的方法，其中，所述当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐，得到第一补齐语句包括：

当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，提取出所述第二

语句检测结果中的单实体，并将所述单实体作为第一目标词语；

在所述第一语句检测结果中筛选与所述第一目标词语语法结构相同的所述单实体，得到第二目标词语；

在所述第一语句中，利用所述第一目标词语替换所述第二目标词语，得到第一补齐语句。

6、根据权利要求 1-5 任一项所述的方法，其中，所述判断所述第一补齐语句中是否包括指代不明的词语，所述指代不明的词语包括代词、量词以及冠词包括：

获取第一补齐语句的检测结果，所述第一补齐语句的检测结果为所述第二语句检测结果与所述第一语句检测结果的结合；

判断所述第一补齐语句的检测结果中是否包括代词、量词以及冠词。

7、根据权利要求 6 所述的方法，其中，所述若所述第一补齐语句中包括指代不明的词语，则根据所述第一语句检测结果将所述第一补齐语句中指代不明的词语替换，得到第二补齐语句包括：

若所述第一补齐语句的检测结果包括指代不明的词语，则提取出所述指代不明的词语，并将所述指代不明的词语作为第三目标词语；

在所述第一语句检测结果中筛选与所述第三目标词语语法结构相同的词语，得到第四目标词语；

在所述第二语句中，利用所述第四目标词语替换所述第三目标词语，得到第二补齐语句。

8、一种多轮对话中语义补齐的装置，其中，包括：

第一获取单元，用于利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一语句中每个词语之间的目标依存关系；

第二获取单元，用于利用所述预置语料断句函数与所述预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，所述第二语句检测结果为所述第二语句中每个词语之间的目标依存关系；

第一判断单元，用于判断所述第二语句检测结果是否仅包括单实体且所述第二语句为疑问句；

第一补齐单元，用于当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐，得到第一补齐语句；

第二判断单元，用于判断所述第一补齐语句中是否包括指代不明的词语，所述指代不明的词语包括代词、量词以及冠词；

第二补齐单元，用于若所述第一补齐语句中包括指代不明的词语，则根据所述第一语句检测结果将所述第一补齐语句中指代不明的词语替换，得到第二补齐语句。

9、一种多轮对话中语义补齐的设备，其中，包括：

存储器和至少一个处理器，所述存储器中存储有指令，所述存储器和所述至少一个处理器通过线路互联；

所述至少一个处理器调用所述存储器中的所述指令，以使得所述多轮对话中语义补齐的设备执行以下步骤：

利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一语句中每个词语之间的目标依存关系；

利用所述预置语料断句函数与所述预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，所述第二语句检测结果为所述第二语句中每

个词语之间的目标依存关系；

判断所述第二语句检测结果是否仅包括单实体且所述第二语句为疑问句；

当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐，得到第一补齐语句；

5 判断所述第一补齐语句中是否包括指代不明的词语，所述指代不明的词语包括代词、量词以及冠词；

若所述第一补齐语句中包括指代不明的词语，则根据所述第一语句检测结果将所述第一补齐语句中指代不明的词语替换，得到第二补齐语句。

10 10、根据权利要求9所述的设备，其中，所述利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果时，具体执行以下步骤：

获取用户输入的第一语句；

利用预置语料断句函数将所述第一语句进行语句分割，得到第一输入语句；

15 利用预置分析函数对所述第一输入语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一输入语句中每个词语之间的目标依存关系。

11、根据权利要求10所述的设备，其中，所述利用预置语料断句函数将所述第一语句进行语句分割，得到第一输入语句时，具体执行以下步骤：

对所述第一语句进行断句，得到分段语句；

20 将所述分段语句中的分段语料与预置语料进行匹配，预置语料是依据业务数据建立在预置意图规则库中的语料；

若所述分段语料与所述预置语料相匹配，则在所述分段语料的前后位置对所述分段语句进行分割，得到分割语句，并将所述分割语句作为第一输入语句；

若所述分段语料与所述预置语料不相配，则直接将所述分段语料作为第一输入语句。

25 12、根据权利要求10所述的设备，其中，所述利用预置分析函数对所述第一输入语句进行语法检测，得到第一语句检测结果时，具体执行以下步骤：

对所述第一输入语句中的词语进行词性标注以及实体抽取，得到第一语句标注结果；

计算每个所述词语之间的依存概率，所述词语为所述第一语句标注结果中的词语，所述依存概率为预置依存关系出现的频次；

30 确定每个所述词语之间的目标依存关系，所述依存概率权重最大的概率所对应的所述预置依存关系为所述词语之间的所述目标依存关系；

获取第一语句检测结果，所述第一语句检测结果为所述第一输入语句中每个所述词语之间的所述目标依存关系。

35 13、根据权利要求9所述的设备，其中，所述当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐，得到第一补齐语句时，具体执行以下步骤：

当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，提取出所述第二语句检测结果中的单实体，并将所述单实体作为第一目标词语；

在所述第一语句检测结果中筛选与所述第一目标词语语法结构相同的所述单实体，得到第二目标词语；

40 在所述第一语句中，利用所述第一目标词语替换所述第二目标词语，得到第一补齐语句。

14、根据权利要求9-13任一项所述的设备，其中，所述判断所述第一补齐语句中是否包括指代不明的词语时，具体执行以下步骤：

获取第一补齐语句的检测结果，所述第一补齐语句的检测结果为所述第二语句检测结

果与所述第一语句检测结果的结合；

判断所述第一补齐语句的检测结果中是否包括代词、量词以及冠词。

15、根据权利要求 14 所述的设备，其中，所述若所述第一补齐语句中包括指代不明的词语，则根据所述第一语句检测结果将所述第一补齐语句中指代不明的词语替换，得到第二补齐语句时，具体执行以下步骤：

若所述第一补齐语句的检测结果包括指代不明的词语，则提取出所述指代不明的词语，并将所述指代不明的词语作为第三目标词语；

在所述第一语句检测结果中筛选与所述第三目标词语语法结构相同的词语，得到第四目标词语；

在所述第二语句中，利用所述第四目标词语替换所述第三目标词语，得到第二补齐语句。

16、一种计算机可读存储介质，其中，包括指令，当所述指令在计算机上运行时，使得计算机执行以下步骤：

利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一语句中每个词语之间的目标依存关系；

利用所述预置语料断句函数与所述预置分析函数，对获取到的用户新一轮输入的第二语句进行语法检测，得到第二语句检测结果，所述第二语句检测结果为所述第二语句中每个词语之间的目标依存关系；

判断所述第二语句检测结果是否仅包括单实体且所述第二语句为疑问句；

当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐，得到第一补齐语句；

判断所述第一补齐语句中是否包括指代不明的词语，所述指代不明的词语包括代词、量词以及冠词；

若所述第一补齐语句中包括指代不明的词语，则根据所述第一语句检测结果将所述第一补齐语句中指代不明的词语替换，得到第二补齐语句。

17、根据权利要求 16 所述的计算机可读存储介质，其中，所述利用预置语料断句函数与预置分析函数，对获取到的用户输入的第一语句进行语法检测，得到第一语句检测结果时，具体执行以下步骤：

获取用户输入的第一语句；

利用预置语料断句函数将所述第一语句进行语句分割，得到第一输入语句；

利用预置分析函数对所述第一输入语句进行语法检测，得到第一语句检测结果，所述第一语句检测结果为所述第一输入语句中每个词语之间的目标依存关系。

18、根据权利要求 17 所述的计算机可读存储介质，其中，所述利用预置语料断句函数将所述第一语句进行语句分割，得到第一输入语句时，具体执行以下步骤：

对所述第一语句进行断句，得到分段语句；

将所述分段语句中的分段语料与预置语料进行匹配，预置语料是依据业务数据建立在预置意图规则库中的语料；

若所述分段语料与所述预置语料相匹配，则在所述分段语料的前后位置对所述分段语句进行分割，得到分割语句，并将所述分割语句作为第一输入语句；

若所述分段语料与所述预置语料不相配，则直接将所述分段语料作为第一输入语句。

19、根据权利要求 17 所述的计算机可读存储介质，其中，所述利用预置分析函数对所述第一输入语句进行语法检测，得到第一语句检测结果时，具体执行以下步骤：

对所述第一输入语句中的词语进行词性标注以及实体抽取，得到第一语句标注结果；

计算每个所述词语之间的依存概率，所述词语为所述第一语句标注结果中的词语，所述依存概率为预置依存关系出现的频次；

确定每个所述词语之间的目标依存关系，所述依存概率权重最大的概率所对应的所述预置依存关系为所述词语之间的所述目标依存关系；

5 获取第一语句检测结果，所述第一语句检测结果为所述第一输入语句中每个所述词语之间的所述目标依存关系。

20、根据权利要求 16 所述的计算机可读存储介质，其中，所述当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，根据所述第一语句检测结果将所述第二语句的语义缺失部分补齐，得到第一补齐语句时，具体执行以下步骤：

10 当所述第二语句检测结果仅包括单实体且所述第二语句为疑问句时，提取出所述第二语句检测结果中的单实体，并将所述单实体作为第一目标词语；

在所述第一语句检测结果中筛选与所述第一目标词语语法结构相同的所述单实体，得到第二目标词语；

15 在所述第一语句中，利用所述第一目标词语替换所述第二目标词语，得到第一补齐语句。

— 1/4 —

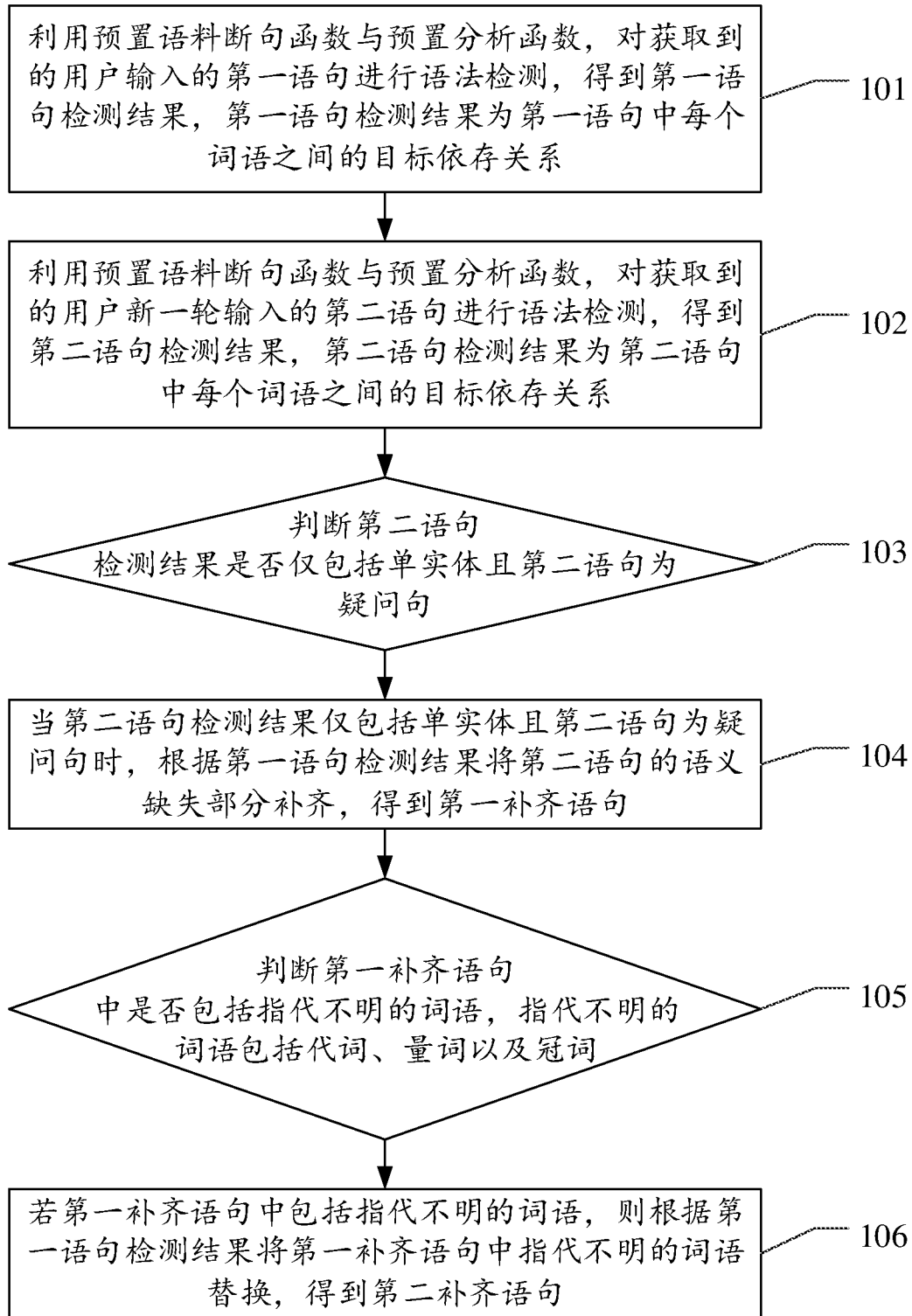


图 1

-2/4-

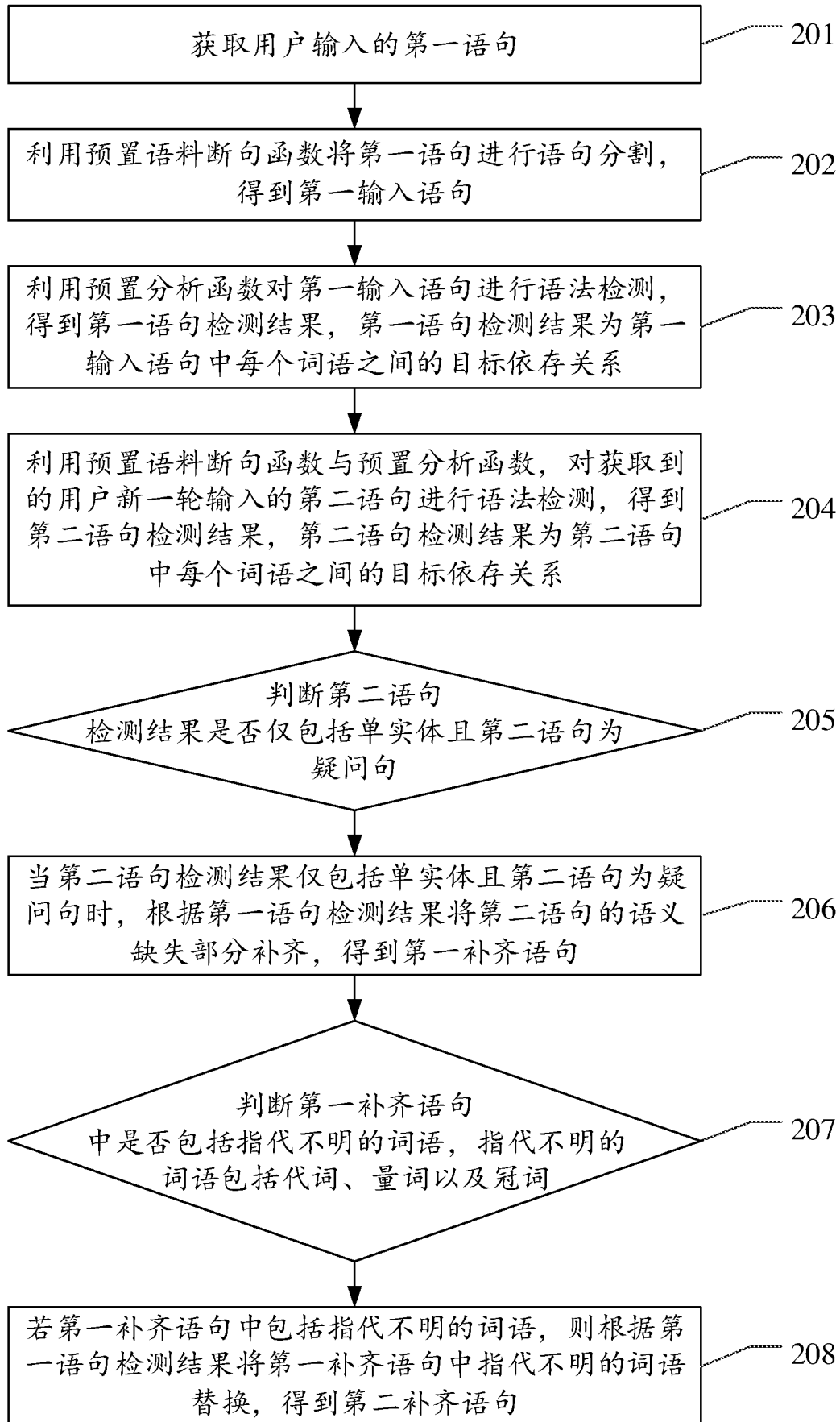


图 2

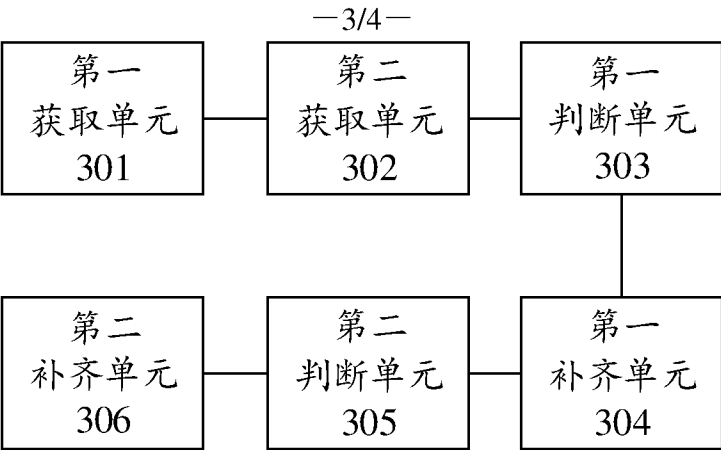


图 3

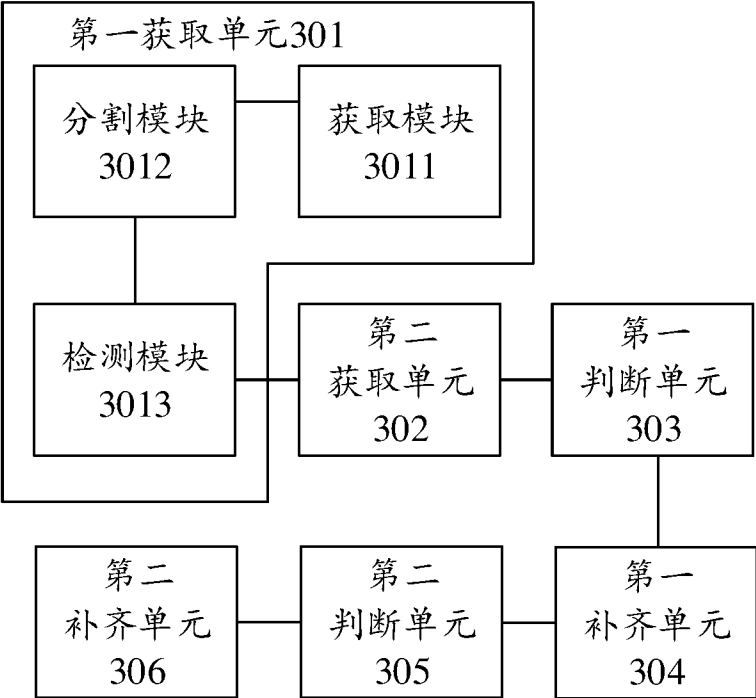


图 4

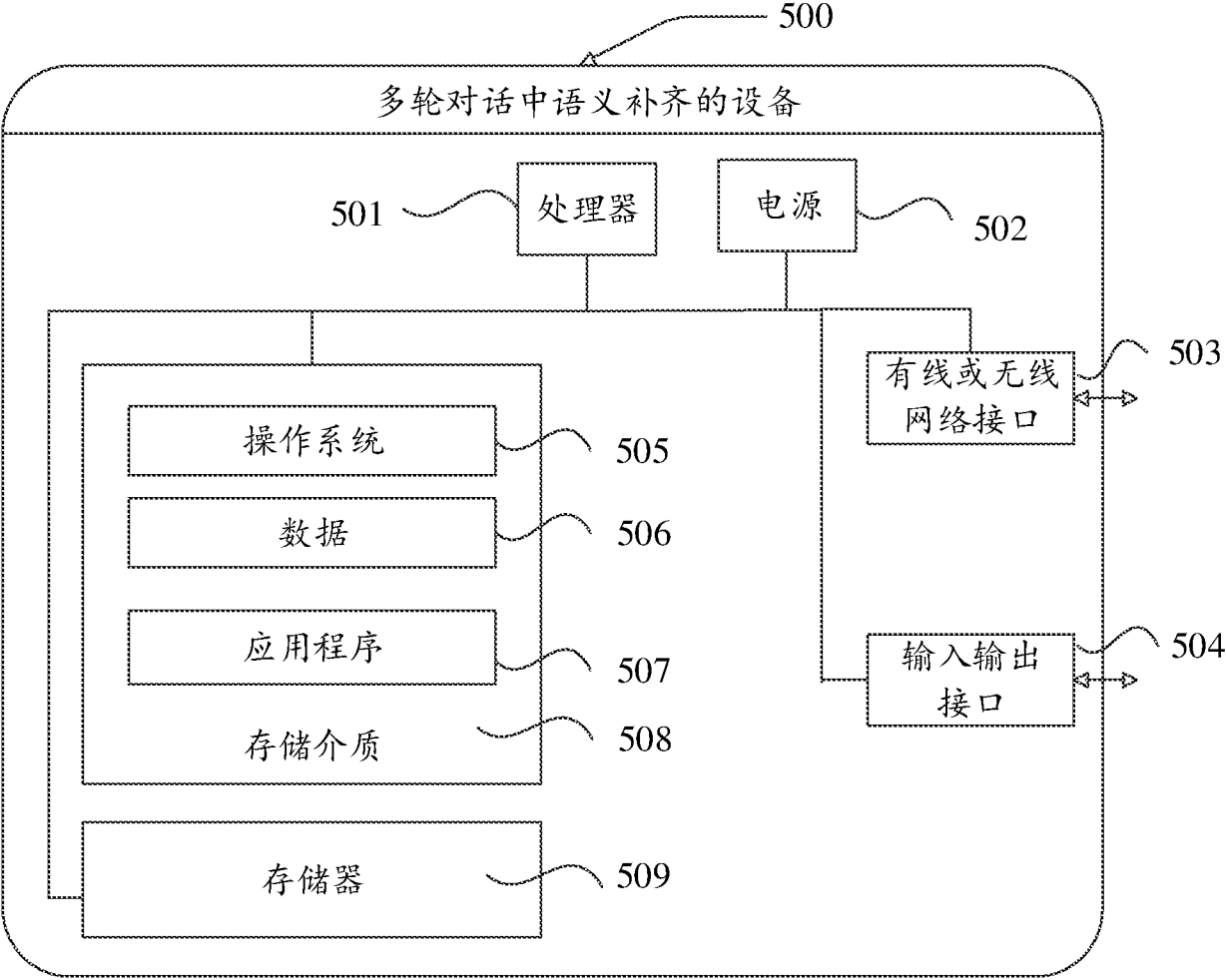


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/098846

A. CLASSIFICATION OF SUBJECT MATTER

G06F 40/30(2020.01)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)

G06F

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)

CNPAT; CNKI; WPI; EPODOC: 对话, 多轮, 语义, 语句, 补齐, 意图, 上下文, 历史, 记录, 缺失, 替代, 替换, 指代, dialogue, multi-turn, semantics, sentence, fill, intent, context, history, record, miss, substitute, replace, refer

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111325034 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 23 June 2020 (2020-06-23) claims 1-10, description, paragraphs [0005]-[0023]	1-20
A	CN 107402913 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 28 November 2017 (2017-11-28) description, paragraphs [0034]-[0169]	1-20
A	US 2019243900 A1 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) 08 August 2019 (2019-08-08) entire document	1-20
A	CN 109697282 A (ALIBABA GROUP HOLDING LIMITED) 30 April 2019 (2019-04-30) entire document	1-20
A	CN 109918494 A (SHENZHEN GOWILD ROBOTICS CO., LTD. et al.) 21 June 2019 (2019-06-21) 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

30 October 2020

Date of mailing of the international search report

18 November 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/CN2020/098846

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111325034	A	23 June 2020	None			
CN	107402913	A	28 November 2017	JP	2019504395	A	14 February 2019
				EP	3460678	A1	27 March 2019
				WO	2017197947	A1	23 November 2017
				US	2018307671	A1	25 October 2018
				KR	20180078318	A	09 July 2018
US	2019243900	A1	08 August 2019	CN	108536708	A	14 September 2018
				WO	2018157805	A1	07 September 2018
CN	109697282	A	30 April 2019	TW	201935273	A	01 September 2019
				US	2020250378	A1	06 August 2020
				WO	2019076286	A1	25 April 2019
CN	109918494	A	21 June 2019	None			

国际检索报告

国际申请号

PCT/CN2020/098846

A. 主题的分类 G06F 40/30 (2020.01) 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) G06F 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)) CNPAT; CNKI; WPI; EPDOC: 对话, 多轮, 语义, 语句, 补齐, 意图, 上下文, 历史, 记录, 缺失, 替代, 替换, 指代, dialogue, multi-turn, semantics, sentence, fill, intent, context, history, record, miss, substitute, replace, refer																				
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>PX</td> <td>CN 111325034 A (平安科技深圳有限公司) 2020年 6月 23日 (2020 - 06 - 23) 权利要求1-10, 说明书第[0005]-[0023]段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 107402913 A (腾讯科技深圳有限公司) 2017年 11月 28日 (2017 - 11 - 28) 说明书第[0034]-[0169]段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 2019243900 A1 (TENCENT TECHNOLOGY SHENZHEN COMPANY LIMITED) 2019年 8月 8日 (2019 - 08 - 08) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 109697282 A (阿里巴巴集团控股有限公司) 2019年 4月 30日 (2019 - 04 - 30) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 109918494 A (深圳狗尾草智能科技有限公司 等) 2019年 6月 21日 (2019 - 06 - 21) 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	PX	CN 111325034 A (平安科技深圳有限公司) 2020年 6月 23日 (2020 - 06 - 23) 权利要求1-10, 说明书第[0005]-[0023]段	1-20	A	CN 107402913 A (腾讯科技深圳有限公司) 2017年 11月 28日 (2017 - 11 - 28) 说明书第[0034]-[0169]段	1-20	A	US 2019243900 A1 (TENCENT TECHNOLOGY SHENZHEN COMPANY LIMITED) 2019年 8月 8日 (2019 - 08 - 08) 全文	1-20	A	CN 109697282 A (阿里巴巴集团控股有限公司) 2019年 4月 30日 (2019 - 04 - 30) 全文	1-20	A	CN 109918494 A (深圳狗尾草智能科技有限公司 等) 2019年 6月 21日 (2019 - 06 - 21) 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																		
PX	CN 111325034 A (平安科技深圳有限公司) 2020年 6月 23日 (2020 - 06 - 23) 权利要求1-10, 说明书第[0005]-[0023]段	1-20																		
A	CN 107402913 A (腾讯科技深圳有限公司) 2017年 11月 28日 (2017 - 11 - 28) 说明书第[0034]-[0169]段	1-20																		
A	US 2019243900 A1 (TENCENT TECHNOLOGY SHENZHEN COMPANY LIMITED) 2019年 8月 8日 (2019 - 08 - 08) 全文	1-20																		
A	CN 109697282 A (阿里巴巴集团控股有限公司) 2019年 4月 30日 (2019 - 04 - 30) 全文	1-20																		
A	CN 109918494 A (深圳狗尾草智能科技有限公司 等) 2019年 6月 21日 (2019 - 06 - 21) 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 10月 30日		国际检索报告邮寄日期 2020年 11月 18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 王璐 电话号码 86-(10)-53961303																		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/098846

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)
CN	111325034	A	2020年 6月 23日	无			
CN	107402913	A	2017年 11月 28日	JP	2019504395	A	2019年 2月 14日
				EP	3460678	A1	2019年 3月 27日
				WO	2017197947	A1	2017年 11月 23日
				US	2018307671	A1	2018年 10月 25日
				KR	20180078318	A	2018年 7月 9日
US	2019243900	A1	2019年 8月 8日	CN	108536708	A	2018年 9月 14日
				WO	2018157805	A1	2018年 9月 7日
CN	109697282	A	2019年 4月 30日	TW	201935273	A	2019年 9月 1日
				US	2020250378	A1	2020年 8月 6日
				WO	2019076286	A1	2019年 4月 25日
CN	109918494	A	2019年 6月 21日	无			