

(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 7 月 23 日 (23.07.2020)



(10) 国际公布号
WO 2020/147428 A1

- (51) 国际专利分类号:
G06F 16/33 (2019.01)
- (21) 国际申请号: PCT/CN2019/120595
- (22) 国际申请日: 2019 年 11 月 25 日 (25.11.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910047156.8 2019 年 1 月 18 日 (18.01.2019) CN
- (71) 申请人: 深圳壹账通智能科技有限公司(ONE CONNECT SMART TECHNOLOGY CO., LTD. (SHENZHEN)) [CN/CN]; 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。
- (72) 发明人: 柳明辉(LIU, Minghui); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。 徐国强(XU, Guoqiang); 中国广东省深圳市前海深港

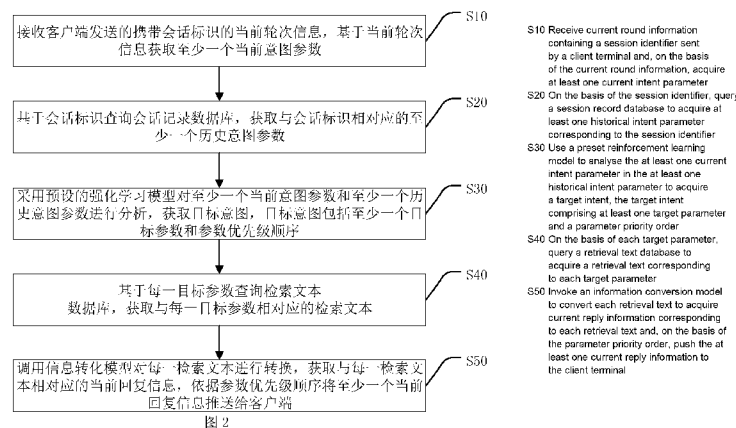
合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。 邱寒(QIU, Han); 中国广东省深圳市前海深港合作区前湾一路 1 号 A 栋 201 室, Guangdong 518000 (CN)。

(74) 代理人: 深圳众鼎专利商标代理事务所(普通合伙)(SHENZHEN ZHONGDING INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市龙岗区龙城街道中心城清林路 546 号城投商务中心 4 层/B, Guangdong 518172 (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:** INTERACTIVE CONTENT GENERATION METHOD AND APPARATUS, COMPUTER DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 交互内容生成方法、装置、计算机设备及存储介质



(57) **Abstract:** An interactive content generation method and apparatus, a computer device, and a storage medium, the method comprising: receiving current round information containing a session identifier sent by a client terminal and, on the basis of the current round information, acquiring at least one current intent parameter (S10); on the basis of the session identifier, querying a session record database to acquire at least one historical intent parameter corresponding to the session identifier (S20); using a preset reinforcement learning model to analyse the at least one current intent parameter in the at least one historical intent parameter to acquire a target intent, the target intent comprising at least one target parameter and a parameter priority order (S30); on the basis of each target parameter, querying a retrieval text database to acquire a retrieval text corresponding to each target parameter (S40); invoking an information conversion model to convert each retrieval text to acquire current reply information corresponding to each retrieval text and, on the basis of the parameter priority order, pushing the at least one current reply information to the client terminal (S50).

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)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

(57) 摘要: 一种交互内容生成方法、装置、计算机设备及存储介质, 该方法包括: 接收客户端发送的携带会话标识的当前轮次信息, 基于当前轮次信息获取至少一个当前意图参数(S10); 基于会话标识查询会话记录数据库, 获取与会话标识相对应的至少一个历史意图参数(S20); 采用预设的强化学习模型对至少一个当前意图参数和至少一个历史意图参数进行分析, 获取目标意图, 目标意图包括至少一个目标参数和参数优先级顺序(S30); 基于每一目标参数查询检索文本数据库, 获取与每一目标参数相对应的检索文本(S40); 调用信息转化模型对每一检索文件进行转换, 获取与每一检索文本相对应的当前回复信息, 依据参数优先级顺序将至少一个当前回复信息推送给客户端(S50)。

交互内容生成方法、装置、计算机设备及存储介质

本申请以 2019 年 1 月 18 日提交的申请号为 201910047156.8，名称为“交互内容生成方法、装置、计算机设备及存储介质”的中国发明申请为基础，并要求其优先权。

技术领域

本申请涉及语义分析技术领域，尤其涉及一种交互内容生成方法、装置、计算机设备及存储介质。

背景技术

随着科学技术的不断发展，信息技术、计算机技术以及人工智能技术的引入，机器人的研究已经逐步走出工业领域，逐渐扩展到了医疗、保健、家庭、娱乐以及服务行业等领域。而人们对于机器人的要求也从简单重复的机械动作提升为具有拟人问答、自主性及与其他机器人进行交互的智能机器人，人机交互也就成为决定智能机器人发展的重要因素。

目前的聊天机器人大部分侧重于和用户进行单轮交互，使得机器人并不能很好地获取用户的真实意图，因为用户聊天的重要信息可能在用户之前的对话中。这种基于单轮的聊天机制，忽略了当前用户前几轮的聊天主题和场景分析，因此其返回的回复内容可能存在偏差甚至错误。如何得到更贴近用户真实意图的聊天内容成为亟待解决的问题。

发明内容

本申请实施例提供一种交互内容生成方法、装置、计算机设备及存储介质，以解决如何得到更贴近用户真实意图的聊天内容的问题。

一种交互内容生成方法，包括：

接收客户端发送的携带会话标识的当前轮次信息，基于当前轮次信息获取至少一个当前意图参数；

基于会话标识查询会话记录数据库，获取与会话标识相对应的至少一个历史意图参数；

采用预设的强化学习模型对至少一个当前意图参数和至少一个历史意图参数进行分析，获取目标意图，目标意图包括至少一个目标参数和参数优先级顺序；

基于每一目标参数查询检索文本数据库，获取与每一目标参数相对应的检索文本；

调用信息转化模型对每一检索文本进行转换，获取与每一检索文本相对应的当前回复信息，依据参数优先级顺序将至少一个当前回复信息推送给客户端。

一种交互内容生成装置，包括：

接收当前信息模块，用于接收客户端发送的携带会话标识的当前轮次信息，基于当前轮次信息获取至少一个当前意图参数；

获取历史参数模块，用于基于会话标识查询会话记录数据库，获取与会话标识相对应的至少一个历史意图参数；

获取目标意图模块，用于采用预设的强化学习模型对至少一个当前意图参数和至少一个历史意图参数进行分析，获取目标意图，目标意图包括至少一个目标参数和参数优先级顺序；

获取检索文本模块，用于基于每一目标参数查询检索文本数据库，获取与每一目标参数相对应的检索文本；

获取回复信息模块，用于调用信息转化模型对每一检索文本进行转换，获取与每一检索文本相对应的当前回复信息，依据参数优先级顺序将至少一个当前回复信息推送给客户端。

一种计算机设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，所述处理器执行所述计算机可读指令时实现如下步骤：

接收客户端发送的携带会话标识的当前轮次信息，基于所述当前轮次信息获取至少一个

当前意图参数；

基于所述会话标识查询会话记录数据库，获取与所述会话标识相对应的至少一个历史意图参数；

采用预设的强化学习模型对至少一个所述当前意图参数和至少一个所述历史意图参数进行分析，获取目标意图，所述目标意图包括至少一个目标参数和参数优先级顺序；

基于每一所述目标参数查询检索文本数据库，获取与每一所述目标参数相对应的检索文本；

调用信息转化模型对每一所述检索文本进行转换，获取与每一所述检索文本相对应的当前回复信息，依据所述参数优先级顺序将至少一个所述当前回复信息推送给所述客户端。

一个或多个存储有计算机可读指令的可读存储介质，所述计算机可读存储介质存储有计算机可读指令，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行如下步骤：

接收客户端发送的携带会话标识的当前轮次信息，基于所述当前轮次信息获取至少一个当前意图参数；

基于所述会话标识查询会话记录数据库，获取与所述会话标识相对应的至少一个历史意图参数；

采用预设的强化学习模型对至少一个所述当前意图参数和至少一个所述历史意图参数进行分析，获取目标意图，所述目标意图包括至少一个目标参数和参数优先级顺序；

基于每一所述目标参数查询检索文本数据库，获取与每一所述目标参数相对应的检索文本；

调用信息转化模型对每一所述检索文本进行转换，获取与每一所述检索文本相对应的当前回复信息，依据所述参数优先级顺序将至少一个所述当前回复信息推送给所述客户端。

本申请的一个或多个实施例的细节在下面的附图及描述中提出。本申请的其他特征和优点将从说明书、附图以及权利要求书变得明显。

附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对本申请实施例的描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 是本申请一实施例中交互内容生成方法的应用环境示意图；

图 2 是本申请一实施例中交互内容生成方法的流程图；

图 3 是本申请一实施例中获取目标意图的实现过程示意图；

图 4 是本申请一实施例中交互内容生成方法的另一流程图；

图 5 是本申请一实施例中交互内容生成方法的另一流程图；

图 6 是本申请一实施例中交互内容生成方法的另一流程图；

图 7 是本申请一实施例中交互内容生成方法的另一流程图；

图 8 是本申请一实施例中交互内容生成装置的示意图；

图 9 是本申请一实施例中计算机设备的示意图。

具体实施方式

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清楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

本申请实施例提供的交互内容生成方法，可应用在如图 1 的应用环境中，该交互内容生

成方法应用在交互内容生成系统中，该交互内容生成系统包括客户端和服务端，其中，客户端通过网络与服务端进行通信。客户端又称为用户端，是指与服务端相对应，为客户提供本地服务的程序。该客户端可安装在但不限于各种个人计算机、笔记本电脑、智能手机、平板电脑和便携式可穿戴设备等计算机设备上。服务端可以用独立的服务器或者是多个服务器组成的服务器集群来实现，用于接受用户通过客户端发送的当前轮次信息，生成聊天机器人对应的回复信息。

在一实施例中，如图2所示，提供一种交互内容生成方法，以该方法应用在图1中的服务器为例进行说明，包括如下步骤：

S10. 接收客户端发送的携带会话标识的当前轮次信息，基于当前轮次信息获取至少一个当前意图参数。

其中，当前轮次信息是用户当前轮次输入客户端的用以表达用户意图的信息。进一步地，当前轮次信息可以包括以多种形式表达的意图信息，包括但不限于文字数据、语音信息或手势动作信息等，此处不作限定。

会话标识是用以区别服务端启动的不同会话场景的标识。因服务端可与若干客户端启动若干会话场景，为了区别每一会话场景以进行会话内容分析，服务端需要给每一会话场景设置对应的会话标识。

意图是用户在当前轮次信息中表达的目的，可分为“显式意图”和“隐式意图”，它们两者间的差异十分明显：

（一）显式意图：

所谓显式意图，也就是在用户的话语中明确出现了表达意图的一类词汇，比如：“希望”、“想要”、“需要”等等。对于聊天机器人来说这种显示意图的判断难度较低，只需要识别出这些固定的意图词汇，再与句子中的其他成分作联系即可。比如用户输入“我想预定到北京的机票”，聊天机器人可以识别出意图词汇“想”，从而得到该句的意图参数为我预定机票，机票行程包括从当前位置到北京。

（二）隐式意图：

隐式意图与显式意图相反，用户的话语没有出现直接反映意图的词汇，需要聊天机器人根据文字数据来判断用户意图。对于聊天机器人来说隐式意图比较难以判断，目前用得比较多的一种方法是先把隐式意图转化成显式意图再处理。比如用户在表达“我饿了”这种隐式意图时，先转化成对应的“我想吃东西”这种显式意图形式，然后再让聊天机器人按照显式意图进行处理。

当前意图参数是将当前轮次信息转换为对应的文字数据后，从该文字数据中提取的含有实际意义的动词和名词（去停用词），以最简洁的方式表达的用户目的。其中，停用词主要包括英文字符、数字、数学字符、标点符号及使用频率高的功能词等。比如，文字数据为“今天天气真不错啊”，从该文字数据中提取的当前意图参数包括：今天、天气和不错（去除停用词“真”和“啊”）。

步骤S10中，服务端通过客户端发送的当前轮次信息可提取出至少一个当前意图参数，过滤掉当前轮次信息中无实际意义的功能词，可直接获取当前轮次信息中有效的当前意图参数，为后续结合历史意图参数获得用户的真实意图准备技术基础。

S20. 基于会话标识查询会话记录数据库，获取与会话标识相对应的至少一个历史意图参数。

其中，会话记录数据库是服务端基于每一会话场景（也即每一会话标识）保存的会话记录集合。

历史意图参数是与当前意图参数对应的，基于同一会话标识的在当前轮次会话之前保存的所有轮次的意图参数。

具体地，服务端基于同一会话标识都在会话记录数据库中对应该记录当前轮次对话产生的当前意图参数。可以理解地，每一当前意图参数被存储到会话记录数据库中即形成历史意图

参数，以备后续服务器可基于同一会话标识匹配出对应的所有历史意图参数进行真实意图分析。

步骤 S20 中，服务器可基于会话标识在会话记录数据库中直接获取对应的所有历史意图参数，无需服务器重新提取所有历史会话中的历史意图参数，加快服务器分析用户真实意图的处理速度。

S30. 采用预设的强化学习模型对至少一个当前意图参数和至少一个历史意图参数进行分析，获取目标意图，目标意图包括至少一个目标参数和参数优先级顺序。

其中，强化学习模型是预设于服务器的，用以将输入该模型的所有历史意图参数和当前意图参数进行分析，获取分析结果作为目标意图的模型。于本实施例，服务器可采用 LSTM(Long Short-Term Memory，长短期记忆网络)模型和 softmax 分类器的组合作为强化学习模型。

于本实施例针对聊天场景可设定多级目标意图，比如一级目标意图包括聊天和提问；一级目标意图又可继续划分多个二级目标意图，对于一级目标意图中的聊天意图可继续划分为生活、工作和休闲；基于二级目标意图，可继续划分为三级目标意图，根据场景需要继续细化等。

具体地，如图 3 所示，服务器通过文字数据获取目标意图的实现过程如下：

1. 将文字数据进行预处理，包括去除语料的标点符号，去除停用词（无实际含义的词）等。

2. 将预处理后的文字数据，采用 word2vec 工具生成词向量。

其中，word2vec 工具是一个 NLP (Natural Language Processing，自然语言处理) 工具，它可以将所有的自然语言中的字词向量化，转为计算机可以理解的稠密向量(Dense Vector)，用以定量地度量词与词之间的关系，挖掘词与词之间的联系。可以理解地，对于相似的词，其对应的词向量也相近。

3. 采用 LSTM 模型对词向量进行特征提取。

其中，LSTM 模型可解决自然语言处理统计方法只能考虑最近 n 个词语而忽略更久前词语的问题，旨在发现词和词之间的相关关系，在数据分析中增加时间文本内容，记住之前发生了什么，然后应用于神经网络，观察与神经网络接下来所发生的事情之间的联系，从而得出目标意图。

LSTM 的特点就是在 RNN (Recurrent Neural Network，循环神经网络) 模型以外添加了各层的阀门节点，如图 4 所示。阀门节点有 3 类：遗忘阀门(forget gate)，输入阀门(input gate)和输出阀门(output gate)。这些阀门节点可以打开或关闭，用于将判断 RNN 模型的记忆态(之前网络的状态)在上一输出层的结果是否达到阈值从而加入到当前该层的计算中。

阀门节点利用 sigmoid 函数将在 RNN 模型中该阀门节点对应的上一输出层运算结果作为输入计算；如果计算结果达到阈值则将该阀门节点输出与当前层的计算结果相乘作为下一层的输入；如果没有达到阈值则将该阀门节点对应的上一输出层运算结果遗忘掉。

LSTM 模型的记忆功能就是由这些阀门节点实现的。当阀门节点打开的时候，前面 RNN 模型的训练结果就会关联到当前的 RNN 模型进行计算，而当阀门节点关闭的时候之前的计算结果就不再影响当前的计算。因此，通过调节阀门节点的开关可以实现历史意图对最终获取目标意图的影响。当不希望历史意图对之后的意图分析产生影响，比如自然语言处理中开始分析新段落或新章节，则关掉阀门节点即可。

4. 采用 softmax 完成意图分类工作。

最后，将 LSTM 模型对词向量进行特征提取后，在神经网络中引入 softmax 分类器可定义一个新的输出层(目标意图概率分布)。基于神经网络的 LSTM 模型的输出层不一定是一个概率分布层，所以可给 LSTM 模型在最后引入 softmax 分类器，softmax 分类器能够作为一个额外的处理层把神经网络的输出(LSTM 模型的最后输出层)变成一个概率分布，也就是每一个输出都是 0 至 1 之间的小数，并且所有输出的结果之和为 1。举例说明，若 LSTM 模型的最后输出层为 y_1, y_2, y_3 和 y_4 ，服务器可采用 softmax 分类器对该 LSTM 模型的最后输出层进行处

理, 采用如下公式:

$$\text{Softmax}(y_i) = y_i' = \frac{e^{y_i}}{\sum_{j=1}^4 e^{y_j}},$$

通过 softmax 分类器可以计算出 $y_1' y_2' y_3' y_4'$ 的值, 其中, y_i' 是每一目标意图的输出概率。

可以理解地, y_i' 中数值最大的输出对应的目标意图就是当前轮次信息对应的目标意图。比如, 步骤 S10 获得文字数据“我最近看了姜文导演的电影邪不压正”, 经步骤 S20 中预设的意图识别模型进行识别后获得的目标意图是: 用户讨论电影, 其中, 电影参数包括: 电影名为邪不压正。

进一步地, 本实施例还可为目标参数设置参数优先级顺序, 也即给动词后的指定动作的对象对应的名词设置较高的参数优先级顺序, 给动作执行人设置次级参数优先级顺序。对于多数讨论场景, 人们谈论事件的重点在于动作执行的对象, 也即动作执行的对象的参数优先级顺序一般高于动作执行人。给不同的目标参数设置参数优先级顺序, 利于后续基于不同的参数优先级顺序进行概念扩展。

或者, 服务器还可设置将包括目标参数个数最多的目标意图的参数优先级的值也最高。比如, 用户通过客户端输入的文字数据为“查询龙岗租房信息, 面积 20 平米, 价格便宜的”。服务器基于该文字数据可分析该文字数据的目标意图是龙岗租房, 将该文字数据进行拆分分析:

- 1) “查询龙岗租房信息”包括的目标参数是龙岗租房;
- 2) “查询龙岗租房信息, 面积 20 平米”包括的目标参数是龙岗租房和面积;
- 3) “查询龙岗租房信息, 面积 20 平米, 价格便宜的”包括的目标参数是龙岗租房、面积和价格。

由上述分析可以看出, 第三组包括的目标参数最多, 也即第三组的参数优先级顺序为 3, 依次类推, 第二组的参数优先级顺序为 2, 第一组的参数优先级顺序为 1。服务器基于参数优先级顺序进行回复时, 可按参数优先级顺序的从大到小的顺序进行查找, 也即将第三组查询到的租房结果首先返回给客户端。

举例说明在本实施例中, 服务器通过目标意图获取目标参数的实现过程:

比如, 文字数据“我最近看了姜文导演的电影邪不压正”, 经步骤 S30 的识别后可确认该文字数据对应的目标意图为动作人谈休闲活动。在动作人谈休闲活动的目标意图模板中可获取该模板的目标参数为“动作人”和“休闲活动”。在文字数据“我最近看了姜文导演的电影邪不压正”中提取与上述目标参数分别对应的词语: 动作人对应“我”(也即用户), 休闲活动“电影邪不压正”。其中, 动作人对应的参数优先级顺序为 1 级, 休闲活动对应的参数优先级顺序为 2 级, 则在后续将目标参数作为回复信息回复给客户端时, 选取参数优先级顺序的值最大的目标参数先回复给客户端。

步骤 S30 中, 服务器能够通过预设的强化学习模型识别用户的目标意图, 精准地判断出用户的需求, 并基于目标参数对应的参数优先级顺序最大的优先回复, 利于后续基于该目标参数得到相应的回复信息, 准确地围绕用户的关注点推进聊天, 以保持聊天机器人和用户之间的聊天黏性和实用性。

S40. 基于每一目标参数查询检索文本数据库, 获取与每一目标参数相对应的检索文本。

其中, 联网数据库即为在线检索库, 比如, 百度、搜狗、谷歌或 360 搜索等。检索文本就是与目标参数对应的回复文本, 比如, 目标参数为“今天天气”, 服务器通过联网数据库查询在线天气可获“今天天气”对应的天气查询结果, 该天气查询结果就是检索文本。

步骤 S40 中, 服务器可通过联网数据库获取目标参数对应的检索文本, 给用户提供与当前轮次信息相关的文本信息, 提高回复内容的准确性。

S50. 调用信息转化模型对每一检索文本进行转换, 获取与每一检索文本相对应的当前回

复信息，依据参数优先级顺序将至少一个当前回复信息推送给客户端。

具体地，本步骤中的信息转化模型是将文本信息转化为传递给用户的回复信息的形式模型。比如，用户以语音形式输入当前轮次信息，则本步骤采用语音转化模型将检索文本转化为语音信息输出给用户，或者也可以按客户端的喜好设置直接以文本形式直接输出给客户端。

信息转化模型是将文本数据表达的含义与特定表达方式相互转化的模型，比如，将语音信息转化为对应的文字数据，将文字数据转化为对应的手语动作等。于本实施例，信息转化模型可基于多种表达式进行转换，因此其包括多个转化模型。比如，将语音信息转化为文字信息的语音转化模型，比如 RNN-HMM (Recurrent Neural Network-Hidden Markov Model 即循环神经网络-隐马尔科夫) 模型或 LSTM-HMM (Long Short-Term Memory, 长短期记忆网络-隐马尔科夫) 模型等机器声学模型，或者将手势信息转化为文字信息的手势转化模型，比如 FLDCRFs (Fuzzy based Latent dynamic Condition Random Fields, 模糊隐动态条件随机场) 等机器手势识别模型。应当理解地，本实施例中 RNN-HMM 机器声学模型、LSTM-HMM 机器声学模型和 FLDCRFs 机器手势识别模型等机器学习模型为公知技术，此处不再赘述。

举例说明步骤 S60 中将检索文本转化为当前回复信息的实现过程，在步骤 S10 中，若用户通过客户端的麦克风输入当前轮次信息（语音信息）：“我最近看了姜文导演的新电影邪不压正”。服务器通过客户端获取麦克风采集的语音格式的当前轮次信息，因此可确定当前轮次信息为语音信息，即可采用 RNN-HMM 机器声学模型将该语音信息转化为“我最近看了姜文导演的新电影邪不压正”这一文字数据。步骤 S60 中，为了保持与用户交流方式的一致性，提高聊天趣味性，聊天计算机可将步骤 S50 得到的检索文本同样采用 RNN-HMM 机器声学模型，将检索文本转化为语音信息（当前回复信息）返回给客户端。

步骤 S50 中，服务器可按照参数优先级顺序将检索文本对应的当前回复信息返回给客户端，利于基于用户的关注点直接回复，及时给客户端返回准确地可供参考的当前回复信息，以提高聊天机器人和用户之间的聊天内容的关联性；服务器可根据场景需要设置为输出的回复信息的形式与用户输入的当前轮次信息的方式一致，保持聊天习惯一致性，提高用户与聊天机器人的聊天趣味性。

本实施例提供的交互内容生成方法中，服务器通过接收客户端发送的当前轮次信息，分析得到至少一个当前意图参数，将当前意图参数结合会话记录数据库获取至少一个历史意图参数后得到目标意图和对应的回复信息，可保障聊天机器人及时生成准确的回复信息回复给客户端，避免仅依靠当前意图参数获得不准确甚至无关的回复信息，提升聊天机器人和客户端之间的互动关联性和可靠性。

在一具体实施方式中，如图 4 所示，步骤 S10 中，即接收客户端发送的当前轮次信息，包括：

S11. 接收客户端发送的当前轮次信息，调用信息转化模型对当前轮次信息进行识别，获取文字数据。

其中，文字数据就是以特定方式发送的当前轮次信息对应含义的文字信息，比如，用户通过客户端的摄像头发送点头动作作为当前轮次信息，服务器可通过调用信息转化模型（于本实施例为动作识别模型）将点头动作对应转化为文字数据“同意”。

步骤 S11 中，服务器通过将接受到的客户端发送的当前轮次信息转换为文字数据，利于服务器后续基于对文字数据进行进一步处理，以获取与当前轮次信息对应的回复信息返回给客户端。

S12. 采用预设语言处理模型对文字数据进行解析处理，获取至少一个当前意图参数。

其中，本实施例的预设语言处理模型可采用 word2vec，一个 NLP (Natural Language Processing, 自然语言处理) 工具，它可以将所有的自然语言中的字词向量化，转为计算机可以理解的稠密向量 (Dense Vector)，用以定量地度量词与词之间的关系，挖掘词与词之间的联系。可以理解地，对于相似的词，其对应的词向量也相近。

本步骤中采用预设语言处理模型对文字数据进行解析处理，获取至少一个当前意图参数的实现过程在步骤 S10 中已进行详细描述，为了避免重复，此处不再赘述。

步骤 S12 中，服务器可采用预设语言处理模型解析文字数据获得至少一个当前意图参数，可提取出用户在当前轮次表达的最简洁的当前意图，利于后续服务器基于该当前意图结合所有历史意图参数获取用户的真实意图。

步骤 S11 至 S12 中，服务器通过将接受到的客户端发送的当前轮次信息转换为文字数据，利于服务器后续基于对文字数据进行进一步处理，以获取与当前轮次信息对应的回复信息返回给客户端。服务器可采用预设语言处理模型解析文字数据获得至少一个当前意图参数，可提取出用户在当前轮次表达的最简洁的当前意图，利于后续服务器基于该当前意图结合所有历史意图参数获取用户的真实意图。

在一具体实施方式中，如图 5 所示，步骤 S30 中，即采用预设的强化学习模型对至少一个当前意图参数和至少一个历史意图参数进行分析，获取目标意图，包括：

S31. 基于至少一个当前意图参数和至少一个历史意图参数，分别获取当前词向量矩阵和历史词向量矩阵。

其中，词向量其实是将词映射到一个语义空间，得到的语义映射矩阵。举例说明，中心词 A 映射出周边词 BCDAEFG，得到一个参数矩阵 W1，然后将中心词 L 映射到周边词 BCDLEFG，得到参数矩阵 W2。若 W2 和 W1 距离相近，则说明 A 和 L 能映射到同样的周边词，可能是近义词。

词向量训练的预处理步骤：

1. 对输入的文本（当前意图参数或历史意图参数）生成一个词汇表，统计词汇表中每个词的词频，按照词频从高到低排序，取最频繁的 V 个词，构成一个词汇表。每个词存在一个维度是 V 的 one-hot 向量。其中，one-hot 向量用以表征向量中每一个元素都关联着词库中的一个单词，指定词的向量表示为：其在向量中对应的元素设置为 1，其他的元素设置为 0（如果该词在词汇表中出现过，则向量中词汇表中对应的位置为 1，其他位置全为 0）。如果词汇表中不出现，则向量为全 0。

2. 将输入文本的每个词都生成一个 one-hot 向量，保留每个词的原始位置，因为是上下文相关的。

3. 确定词向量的维数 N。

具体地，本实施例可采用 CBOW 模型（连续词袋模型，将词语的上下文作为输入，来预测这个词本身）获取词向量，实现步骤如下：

1. 确定窗口大小 window，对每个词生成 $2 \times \text{window}$ 个训练样本， $(i - \text{window}, i)$ ， $(i - \text{window} + 1, i)$ ， $\dots$ ， $(i + \text{window} - 1, i)$ ， $(i + \text{window}, i)$ 。

2. 确定 batch_size，batch_size 的大小须是 $2 \times \text{window}$ 的整数倍，以确保每个 batch 包含了一个词汇对应的所有样本。

3. 训练算法有两种：层次 Softmax 和 Negative Sampling。

4. 神经网络迭代训练一定次数，得到输入层到维数为 N 的隐藏层的参数矩阵，矩阵中每一行的转置即是对应词的维数为 V 的词向量矩阵。

步骤 S31 中，服务器可分别获取当前意图参数和历史意图参数对应的当前词向量矩阵和历史词向量矩阵，为后续服务器基于词向量矩阵获取目标意图准备技术基础。

S32. 采用预设的强化学习模型对当前词向量矩阵和历史词向量矩阵进行分析，获取目标意图。

具体地，本步骤中可采用步骤 S30 中的 LSTM（Long Short-Term Memory，长短期记忆网络）模型和 softmax 分类器的组合作为强化学习模型，对当前词向量矩阵和历史词向量矩阵进行分析进行解析处理，获取目标意图。

将当前词向量矩阵和历史词向量矩阵输入到预设的强化学习模型的输入阀门，具体的实现过程已在前述步骤 S30 中进行详细描述，为了避免重复，此处不再赘述。

步骤 S32 中, 服务器能够通过预设的强化学习模型识别用户的目标意图, 精准地判断出用户的需求, 获取与用户需求更贴合的聊天内容, 减少聊天机器人和用户之间答非所问的次数。

步骤 S31 至 S32 中, 服务器可分别获取当前意图参数和历史意图参数对应的当前词向量矩阵和历史词向量矩阵, 为后续服务器基于词向量矩阵获取目标意图准备技术基础。服务器能够通过预设的强化学习模型识别用户的目标意图, 精准地判断出用户的需求, 获取与用户需求更贴合的聊天内容, 减少聊天机器人和用户之间答非所问的次数。

在一具体实施方式中, 如图 6 所示, 步骤 40 中, 基于每一目标参数查询联网数据库, 获取与每一目标参数相对应的检索文本, 具体包括如下步骤:

S41. 若目标参数的属性为产品属性, 则基于每一目标参数查询与产品属性相对应的产品数据库, 获取与每一目标参数相对应的产品推荐信息作为检索文本。

其中, 产品属性是内容提供商通过服务器给每一目标参数标记的事物属性。其中, 内容提供商是提供聊天机器人所在服务器的提供商, 该提供商可同时售卖其它实物产品或提供各种服务。

可以理解地, 当本聊天机器人应用于内容提供商的营销场景, 聊天机器人可在用户沟通的恰当时机对内容提供商提供的营销产品或服务进行推介, 利于采用该聊天机器人系统的内容提供商提高产品或服务销售转化率。比如, 目标参数为“保险”, 若保险属于内容提供商的自由产品, 内容提供商可给“保险”标注为产品属性。

具体地, 服务器可将每一目标意图模板中的实物名称或服务名称与内容提供商自身提供的实物或服务类型进行关联, 用以在服务器获取到的目标参数包括其中的实物名称或服务名称时, 内容提供商可根据关联关系给用户推出对应的产品或服务介绍。

产品数据库是保存内容提供商可提供的所有产品或服务的数据库。产品推荐信息是产品数据库中给每一产品或服务记录的产品内容介绍。

优选的, 产品推荐信息可包括产品概要介绍和产品详细介绍, 用以服务器首先推送给客户端产品概要介绍。当服务器接收到客户端发送的产品详细介绍请求时, 表明用户有意愿在继续了解产品的详细情况时, 可再继续给客户端发送产品对应的产品详细介绍。

步骤 S41 中, 服务器可预先给每一目标参数标注是否属于产品属性的标签, 利于服务器获取到目标参数即刻可进行判断, 无需再对每一目标参数进行分析是否为产品属性, 利于缩短服务器的判定时间。服务器可基于产品数据库获取目标参数对应的产品推荐信息, 用以后续推送给用户, 提升采用本聊天机器人的实际适用性和聊天内容可扩展性。

优选地, 在步骤 S41 之后, 即在获取与每一目标参数相对应的产品推荐信息作为检索文本之后, 交互内容生成方法还包括: 将产品推荐信息按预设产品推送格式推送给客户端。

其中, 预设产品推送格式是服务器预先设置的将产品在聊天界面推送给客户端的格式, 比如, 服务器可设定推送产品名列表, 或者每一产品对应产品简单介绍等。

进一步地, 当服务器接收到客户端选中的产品名, 可将该产品名与步骤 S51 中建立的产品详细介绍关联, 以将产品详细介绍再推送给客户端, 避免初次即推送给客户端的用户不感兴趣的产品的详细介绍。

本步骤中, 服务器可按预设产品推送格式将产品推荐信息发送给客户端, 保持聊天内容涉及到产品或服务时的格式一致性。

S42. 若目标参数的属性不为产品属性, 基于每一目标参数查询联网数据库, 获取与每一目标参数相对应的检索文本。

其中, 联网数据库即为在线检索库, 比如, 百度、搜狗、谷歌或 360 搜索等。检索文本就是与目标参数对应的回复文本, 比如, 目标参数为“今天天气”, 服务器通过联网数据库查询在线天气可获“今天天气”对应的天气查询结果, 该天气查询结果就是检索文本。

步骤 S42 中，服务器可通过联网数据库获取目标参数对应的检索文本，给用户提供与当前轮次信息相关的检索文本，提高回复内容的准确性；同时，服务器无需在本地保存与目标参数对应的检索文本，可节省服务器本地的存储空间。

步骤 S41 至 S42 中，服务器可预先给每一目标参数标注是否属于产品属性的标签，利于服务器获取到目标参数即可进行判断，无需再对每一目标参数进行分析是否为产品属性，利于缩短服务器的判定时间。服务器可基于产品数据库获取目标参数对应的产品推荐信息，用以后续推送给用户，提升采用本聊天机器人的实际适用性和聊天内容可扩展性。服务器可通过联网数据库获取目标参数对应的检索文本，给用户提供与当前轮次信息相关的检索文本，提高回复内容的准确性；同时，服务器无需在本地保存与目标参数对应的检索文本，可节省服务器本地的存储空间。

在一具体实施方式中，如图 7 所示，在步骤 S50 之前，即在所述调用信息转化模型对每一所述检索文本进行转换之前，所述交互内容生成还包括：

S501. 获取用户个人信息，基于用户个人信息生成用户画像。

其中，用户个人信息是服务器收集的静态数据和动态数据，其中，静态数据是用户在服务器注册时主动输入的长时间内不会变化的个人信息，比如，用户的地域、年龄、性别、文化、职业和收入等；动态数据是服务器与用户互动过程中对用户行为进行分析得到的数据，比如，生活习惯或消费习惯等。总体来讲，为了获取用户画像需要用到如下用户个人信息：

- (1) 人口属性：包括性别、年龄等基本信息；
- (2) 兴趣特征：浏览内容、收藏内容、阅读咨询、购买物品偏好等；
- (3) 消费特征：与消费相关的特征；
- (4) 位置特征：用户所处城市、所处居住区域、用户移动轨迹等；
- (5) 设备属性：使用的终端特征等；
- (6) 行为数据：访问时间、浏览路径等用户在网站的行为日志数据；
- (7) 社交数据：用户社交相关数据。

用户画像是服务器根据用户个人信息而抽象出来的高度精炼的标签，也即用户特征标识。通过给用户“贴”标签可以利用高度概括、容易理解的特征来描述用户，利于服务器根据标签（标准化信息）进一步处理。

具体地，基于用户个人信息生成用户画像的实现过程如下：

一、对用户个人信息中的某项特征进行抽象分类和概括形成标签，该标签的标签值具备可分类性。

例：将用户个人信息中“男”、“女”这类特征进行抽象概括，统称为“性别”，“性别”即一个标签；

二、穷举标签值（Tag Value），以使该标签包括对应的所有可能的情况的值。

例：对于标签“性别”，其标签值可分为“男”、“女”和“未知”；

对于标签“年龄”，其标签值可分为“0-18”、“18-35”、“35-60”、“60-100”等。

三、构建用户画像（User Profile）。根据步骤一和步骤二创造的标签，提取用户个人信息中与每一标签对应的标签值。

例：用户画像包括的标签有性别、年龄、手机品牌、居住地和爱好等。小明是用户画像的一个实例，小明的用户画像的输出结果为：“男”、“18-35”、“iPhone”、“北京”和“足球”。

步骤 S501 中，服务器可基于用户个人信息构建用户画像，给后续服务器基于该用户画像分析用户的用户性格准备技术基础。

S502. 分析用户画像，获取与用户画像相对应的用户性格。

其中，用户性格是一个人对现实的稳定的态度，以及与这种态度相应的，习惯化了的行为方式中表现出来的人格特征。性格一经形成便比较稳定，但是并非一成不变，而是可塑性的。性格不同于气质，更多体现了人格的社会属性，个体之间的人格差异的核心是性格的差

异。

具体地，通过对用户画像进行分析和挖掘，还可以揭示用户的内心需求和用户性格等潜在属性。比如，用户画像中有关购物标签包括平均浏览时间和平均对比同类次数，三个用户的网上购物标签记录情况如下：

	平均浏览时间（分钟）	平均对比次数
用户 A	7	3
用户 B	15	12
用户 C	25	20

购物性格参考模板（基于实际情况统计生成）：

冲动型：平均浏览时间少于 10 分钟，平均对比次数少于 5 次。

理性型：平均浏览时间介于 10 分钟至 20 分钟，平均对比次数介于 5 次至 20 次。

犹豫型：平均浏览时间大于 20 分钟，平均对比次数大于 20 次。

服务器基于三个用户的购物标签和购物性格参考模板进行对比可得如下结果：

用户 A 总在短时间（平均浏览时间少于 10 分钟）内比较少量的商品（平均对比次数少于 5 次）下单，则用户 A 的购物性格为冲动型。

用户 B 总在适量时间内（平均浏览时间介于 10 分钟至 20 分钟）比较少量同类商品（平均对比次数介于 5 次至 20 次）后下单，则用户 B 的购物性格为理性型。

用户 C 总在长时间（平均浏览时间大于 20 分钟）大量地浏览很多商品（平均对比次数大于 20 次）后下单，则用户 C 的购物性格为犹豫型。

同样地，服务器还可对用户画像进行分析获得用户的聊天性格，聊天性格包括：兴奋、平静或低沉等。需要补充的是，服务器还可通过分析用户输入的当前轮次信息中的声音特质结合聊天性格参考模板来判断用户性格。其中，声音特征的聊天性格参考模板（基于实际情况统计生成）如下：

1) 掺杂呼吸且脆弱的声音：如果用户是男性，属于年轻艺术家的类型；如果用户是女性，虽性格偏女性化，且集美丽、小巧、开朗于一身，这类女性比较容易兴奋。

2) 无力的声音：如果用户是男性，无特定性格；如果用户是女性，则具有社交能力强、感性、幽默等特性。

3) 平淡的声音：用户无论男女，性格特征均为偏男性化、精神状态欠佳、冷淡、畏缩不前，当前状态不太积极。

4) 声调低且粗犷的声音：如果用户是男性，应该拥有敏锐的观察力，且现实、圆滑、成熟、干练、适应性强；如果用户是女性，性格特质主要表现为喜欢偷懒或病怏怏等。

5) 爽朗的声音：如果用户是男性，性格干练且自尊心强；如果用户是女性，则性格活泼、善于社交、自尊心强，缺乏幽默感。

6) 语速快：用户无论男女，该类人群的性格均较活泼，且社交能力强。

7) 抑扬顿挫且洪亮的声音：用户无论男女，该类人群精力旺盛。

步骤 S502 中，服务器可对步骤 S601 中获取的用户画像根据实际经验进行分析，获得用户对应的用户性格，为后续服务器基于不同的用户性格采用不同的回复模式准备技术基础。

S503. 基于用户性格查询预置回复模式库，获取与用户性格对应的目标回复模式。

其中，回复模式库是预置于服务器的，针对不同的用户性格的回复信息的模式组成的数据库。

目标回复模式是针对用户性格进行回复的聊天模式，本实施例主要应用于语音聊天，因

此回复模式即为聊天机器人与用户聊天时的语气回复模式。

具体地，根据心理学研究聊天时语气同步的双方，更加容易被对方接受。也就是说如果用户说话语气慢条斯理，若聊天机器人跟着用户的节奏聊天，可增加用户与聊天机器人继续聊天的意愿。基于此，服务器可根据用户性格搭配对应的聊天语气，以进一步增强用户和聊天机器人之间的聊天趣味性和情感交互性。

进一步地，服务器可将每一用户性格配置一种语气回复模式，比如，对于热情、兴奋或积极等用户性格的语气回复模式为快速且热情的语气回复模式；对于冷静、平淡或成熟的用户性格的语气回复模式为中速且平静的语气回复模式等。

步骤 S503 中，服务器可根据用户性格在预置回复模式库中匹配对应的目标回复模式，以进一步增强用户和聊天机器人之间的聊天趣味性和情感交互性。

S504. 将目标回复模式添加到信息转化模型，以更新信息转化模型。

具体地，服务器将步骤 S603 获得的目标回复模式（可以为语气回复模式），比如快速且热情的语气回复模式等发送给信息转化模型，以使信息转化模型将概念文本转换为与目标回复模式相匹配的当前回复信息。进一步地，信息转化模型可通过调整语速和说话的音调实现获取与目标回复模式相匹配的当前回复信息。

步骤 S504 中，服务器可结合目标回复模式和信息转化模型获得更新的信息转化模型，用以调整与目标回复模式相匹配的语气，将对应的回复信息发送给客户端，增强聊天机器人和用户之间的情感交互性，提高聊天内容的有效性。

步骤 S501 至 S504 中，服务器可根据实际经验对用户画像进行分析，获得用户对应的用户性格，并根据用户性格在预置回复模式库中匹配对应的目标回复模式，以进一步增强用户和聊天机器人之间的聊天趣味性和情感交互性。服务器可将调整后与目标回复模式相匹配的当前回复信息发送给客户端，增强聊天机器人和用户之间的情感交互性，提高聊天内容的有效性。

本实施例提供的交互内容生成方法中，服务器通过接收客户端发送的当前轮次信息，分析得到至少一个当前意图参数，将当前意图参数结合会话记录数据库获取至少一个历史意图参数后得到目标意图和对应的回复信息，可保障聊天机器人及时生成准确的回复信息回复给客户端，避免仅依靠当前意图参数获得不准确甚至无关的回复信息，提升聊天机器人和客户端之间的互动关联性和可靠性。

进一步地，服务器通过将接受到的客户端发送的当前轮次信息转换为文字数据，利于服务器后续基于对文字数据进行进一步处理，以获取与当前轮次信息对应的回复信息返回给客户端。服务器可采用预设语言处理模型解析文字数据获得至少一个当前意图参数，可提取出用户在当前轮次表达的最简洁的当前意图，利于后续服务器基于该当前意图结合所有历史意图参数获取用户的真实意图。

进一步地，服务器可分别获取当前意图参数和历史意图参数对应的当前词向量矩阵和历史词向量矩阵，为后续服务器基于词向量矩阵获取目标意图准备技术基础。服务器能够通过预设的强化学习模型识别用户的目标意图，精准地判断出用户的需求，获取与用户需求更贴合的聊天内容，减少聊天机器人和用户之间答非所问的次数。

进一步地，服务器可预先给每一目标参数标注是否属于产品属性的标签，利于服务器获取到目标参数即刻可进行判断，无需再对每一目标参数进行分析是否为产品属性，利于缩短服务器的判定时间。服务器可基于产品数据库获取目标参数对应的产品推荐信息，用以后续推送给用户，提升采用本聊天机器人的实际适用性和聊天内容可扩展性。服务器可通过联网数据库获取目标参数对应的检索文本，给用户与当前轮次信息相关的检索文本，提高回复内容的准确性；同时，服务器无需在本地保存与目标参数对应的检索文本，可节省服务器本地的存储空间。

进一步地，服务器可根据实际经验对用户画像进行分析，获得用户对应的用户性格，并根据用户性格在预置回复模式库中匹配对应的目标回复模式，以进一步增强用户和聊天机器

人之间的聊天趣味性和情感交互性。服务器可将调整后与目标回复模式相匹配的当前回复信息发送给客户端，增强聊天机器人和用户之间的情感交互性，提高聊天内容的有效性。

应理解，上述实施例中各步骤的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不对本申请实施例的实施过程构成任何限定。

在一实施例中，提供一种交互内容生成装置，该交互内容生成装置与上述实施例中交互内容生成方法一一对应。如图8所示，该交互内容生成装置包括接收当前信息模块10、获取历史参数模块20、获取目标意图模块30、获取目标意图模块30、获取检索文本模块40和获取回复信息模块50。各功能模块详细说明如下：

接收当前信息模块10，用于接收客户端发送的携带会话标识的当前轮次信息，基于当前轮次信息获取至少一个当前意图参数。

获取历史参数模块20，用于基于会话标识查询会话记录数据库，获取与会话标识相对应的至少一个历史意图参数。

获取目标意图模块30，用于采用预设的强化学习模型对至少一个当前意图参数和至少一个历史意图参数进行分析，获取目标意图，目标意图包括至少一个目标参数和参数优先级顺序。

获取检索文本模块40，用于基于每一目标参数查询检索文本数据库，获取与每一目标参数相对应的检索文本。

获取回复信息模块50，用于调用信息转化模型对每一检索文本进行转换，获取与每一检索文本相对应的当前回复信息，依据参数优先级顺序将至少一个当前回复信息推送给客户端。

优选地，接收当前信息模块10包括接收轮次信息单元11和获取意图参数单元12。

接收轮次信息单元11，用于接收客户端发送的当前轮次信息，调用信息转化模型对当前轮次信息进行识别，获取文字数据。

获取意图参数单元12，用于采用预设语言处理模型对文字数据进行解析处理，获取至少一个当前意图参数。

优选地，获取目标意图模块包括分别获取当前词向量矩阵和历史词向量矩阵单元、获取词性向量矩阵和获取目标意图单元。

分别获取当前词向量矩阵和历史词向量矩阵单元，用于基于至少一个当前意图参数和至少一个历史意图参数，分别获取当前词向量矩阵和历史词向量矩阵。

获取词性向量矩阵，用于基于当前意图参数和历史意图参数的词性标注，获取与当前词向量矩阵和历史词向量矩阵分别对应的当前词性向量矩阵和历史词性向量矩阵。

获取目标意图单元，用于采用预设的强化学习模型对当前词向量矩阵和历史词向量矩阵进行分析，获取目标意图。

优选地，获取检索文本模块包括获取产品信息单元和获取检索文本单元。

获取产品信息单元，用于若目标参数的属性为产品属性，则基于每一目标参数查询与产品属性相对应的产品数据库，获取与每一目标参数相对应的产品推荐信息作为检索文本。

获取检索文本单元，用于若目标参数的属性不为产品属性，基于每一目标参数查询联网数据库，获取与每一目标参数相对应的检索文本。

优选地，该交互内容生成装置还包括推送产品信息模块。

推送产品信息模块，用于将产品推荐信息按预设产品推送格式推送给客户端。

优选地，该交互内容生成装置还包括推送产品信息模块和获取问题文本模块。

推送产品信息模块，用于判断检索文本是否包含至少两个目标选项。

获取问题文本模块，用于若包含至少两个目标选项，则获取与至少两个目标选项相对应的问题文本，调用信息转化模型对问题文本进行转换，获取当前回复信息，将当前回复信息推送给客户端。

关于交互内容生成装置的具体限定可以参见上文中对于交互内容生成方法的限定，在此不再赘述。上述交互内容生成装置中的各个模块可全部或部分通过软件、硬件及其组合来实

现。上述各模块可以硬件形式内嵌于或独立于计算机设备中的处理器中，也可以以软件形式存储于计算机设备中的存储器中，以便于处理器调用执行以上各个模块对应的操作。

在一实施例中，提供了一种计算机设备，该计算机设备可以是服务器，其内部结构图可以如图 9 所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中，该计算机设备的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于交互内容生成方法相关的数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令被处理器执行时以实现一种交互内容生成方法。

在一实施例中，提供一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机可读指令，处理器执行计算机可读指令时实现上述实施例交互内容生成方法，例如图 2 所示 S10 至步骤 S50。或者，处理器执行计算机可读指令时实现上述实施例中交互内容生成装置的各模块/单元的功能，例如图 8 所示模块 10 至模块 50 的功能。为避免重复，此处不再赘述。

在一实施例中，一个或多个存储有计算机可读指令的可读存储介质，所述计算机可读存储介质存储有计算机可读指令，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器执行时实现上述实施例交互内容生成方法，例如图 2 所示 S10 至步骤 S50。或者，该计算机可读指令被处理器执行时实现上述装置实施例中交互内容生成装置中各模块/单元的功能，例如图 8 所示模块 10 至模块 50 的功能。为避免重复，此处不再赘述。本实施例中的可读存储介质包括非易失性可读存储介质和易失性可读存储介质。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机可读指令来指令相关的硬件来完成，该计算机可读指令可存储于一非易失性可读存储介质也可以存储在易失性可读存储介质中，该计算机可读指令在执行时，可包括如上述各方法的实施例的流程。其中，本申请各实施例中所使用的对存储器、存储、数据库或其它介质的任何引用，均可包括非易失性和/或易失性存储器。非易失性存储器可包括只读存储器（ROM）、可编程 ROM（PROM）、电可编程 ROM（EPROM）、电可擦除可编程 ROM（EEPROM）或闪存。易失性存储器可包括随机存取存储器（RAM）或者外部高速缓冲存储器。作为说明而非局限，RAM 以多种形式可得，诸如静态 RAM（SRAM）、动态 RAM（DRAM）、同步 DRAM（SDRAM）、双数据率 SDRAM（DDRSDRAM）、增强型 SDRAM（ESDRAM）、同步链路（Synchlink）DRAM（SLDRAM）、存储器总线（Rambus）直接 RAM（RDRAM）、直接存储器总线动态 RAM（DRDRAM）、以及存储器总线动态 RAM（RDRAM）等。

所属领域的技术人员可以清楚地了解到，为了描述的方便和简洁，仅以上述各功能单元、模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能单元、模块完成，即将所述装置的内部结构划分成不同的功能单元或模块，以完成以上描述的全部或者部分功能。

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，均应包含在本申请的保护范围之内。

权 利 要 求 书

1. 一种交互内容生成方法，其特征在于，包括：

接收客户端发送的携带会话标识的当前轮次信息，基于所述当前轮次信息获取至少一个当前意图参数；

基于所述会话标识查询会话记录数据库，获取与所述会话标识相对应的至少一个历史意图参数；

采用预设的强化学习模型对至少一个所述当前意图参数和至少一个所述历史意图参数进行分析，获取目标意图，所述目标意图包括至少一个目标参数和参数优先级顺序；

基于每一所述目标参数查询检索文本数据库，获取与每一所述目标参数相对应的检索文本；

调用信息转化模型对每一所述检索文本进行转换，获取与每一所述检索文本相对应的当前回复信息，依据所述参数优先级顺序将至少一个所述当前回复信息推送给所述客户端。

2. 如权利要求 1 所述的交互内容生成方法，其特征在于，所述接收客户端发送的携带会话标识的当前轮次信息，基于所述当前轮次信息获取至少一个当前意图参数，包括：

接收客户端发送的当前轮次信息，调用信息转化模型对所述当前轮次信息进行识别，获取文字数据；

采用预设语言处理模型对所述文字数据进行解析处理，获取至少一个当前意图参数。

3. 如权利要求 1 所述的交互内容生成方法，其特征在于，所述采用预设的强化学习模型对至少一个所述当前意图参数和至少一个所述历史意图参数进行分析，获取目标意图，包括：

基于至少一个所述当前意图参数和至少一个所述历史意图参数，分别获取当前词向量矩阵和历史词向量矩阵；

采用预设的强化学习模型对所述当前词向量矩阵和所述历史词向量矩阵进行分析，获取目标意图。

4. 如权利要求 1 所述的交互内容生成方法，其特征在于，所述基于每一所述目标参数查询检索文本数据库，获取与每一所述目标参数相对应的检索文本，包括：

若所述目标参数的属性为产品属性，则基于每一所述目标参数查询与所述产品属性相对应的产品数据库，获取与每一所述目标参数相对应的产品推荐信息作为检索文本；

若所述目标参数的属性不为产品属性，基于每一所述目标参数查询联网数据库，获取与每一所述目标参数相对应的检索文本。

5. 如权利要求 4 所述的交互内容生成方法，其特征在于，若所述目标参数的属性为产品属性，在所述获取与每一所述目标参数相对应的检索文本之后，所述交互内容生成方法还包括：

将所述产品推荐信息按预设产品推送格式推送给所述客户端。

6. 如权利要求 1 所述的交互内容生成方法，其特征在于，在所述调用信息转化模型对每一所述检索文本进行转换之前，所述交互内容生成方法还包括：

获取用户个人信息，基于所述用户个人信息生成用户画像；

分析所述用户画像，获取与所述用户画像相对应的用户性格；

基于所述用户性格查询预置回复模式库，获取与所述用户性格对应的目标回复模式；

将所述目标回复模式添加到所述信息转化模型，以更新所述信息转化模型。

7. 一种交互内容生成装置，其特征在于，所述交互内容生成装置包括：

接收当前信息模块，用于接收客户端发送的携带会话标识的当前轮次信息，基于所述当前轮次信息获取至少一个当前意图参数；

获取历史参数模块，用于基于所述会话标识查询会话记录数据库，获取与所述会话标识相对应的至少一个历史意图参数；

获取目标意图模块，用于采用预设的强化学习模型对至少一个所述当前意图参数和至少一个所述历史意图参数进行分析，获取目标意图，所述目标意图包括至少一个目标参数和参数优先级顺序；

获取检索文本模块，用于基于每一所述目标参数查询检索文本数据库，获取与每一所述目标参数相对应的检索文本；

获取回复信息模块，用于调用信息转化模型对每一所述检索文本进行转换，获取与每一所述检索文本相对应的当前回复信息，依据所述参数优先级顺序将至少一个所述当前回复信息推送给所述客户端。

8. 如权利要求 7 所述的交互内容生成装置，其特征在于，接收当前信息模块包括：

接收轮次信息单元，用于接收客户端发送的当前轮次信息，调用信息转化模型对所述当前轮次信息进行识别，获取文字数据；

获取意图参数单元，用于采用预设语言处理模型对所述文字数据进行解析处理，获取至少一个当前意图参数。

9. 一种计算机设备，包括存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机可读指令，其特征在于，所述处理器执行所述计算机可读指令时实现如下步骤：

接收客户端发送的携带会话标识的当前轮次信息，基于所述当前轮次信息获取至少一个当前意图参数；

基于所述会话标识查询会话记录数据库，获取与所述会话标识相对应的至少一个历史意图参数；

采用预设的强化学习模型对至少一个所述当前意图参数和至少一个所述历史意图参数进行分析，获取目标意图，所述目标意图包括至少一个目标参数和参数优先级顺序；

基于每一所述目标参数查询检索文本数据库，获取与每一所述目标参数相对应的检索文本；

调用信息转化模型对每一所述检索文本进行转换，获取与每一所述检索文本相对应的当前回复信息，依据所述参数优先级顺序将至少一个所述当前回复信息推送给所述客户端。

10. 如权利要求 9 所述的计算机设备，其特征在于，所述接收客户端发送的携带会话标识的当前轮次信息，基于所述当前轮次信息获取至少一个当前意图参数，包括：

接收客户端发送的当前轮次信息，调用信息转化模型对所述当前轮次信息进行识别，获取文字数据；

采用预设语言处理模型对所述文字数据进行解析处理，获取至少一个当前意图参数。

11. 如权利要求 9 所述的计算机设备，其特征在于，所述采用预设的强化学习模型对至少一个所述当前意图参数和至少一个所述历史意图参数进行分析，获取目标意图，包括：

基于至少一个所述当前意图参数和至少一个所述历史意图参数，分别获取当前词向量矩阵和历史词向量矩阵；

采用预设的强化学习模型对所述当前词向量矩阵和所述历史词向量矩阵进行分析，获取目标意图。

12. 如权利要求 9 所述的计算机设备，其特征在于，所述基于每一所述目标参数查询检索文本数据库，获取与每一所述目标参数相对应的检索文本，包括：

若所述目标参数的属性为产品属性，则基于每一所述目标参数查询与所述产品属性相对应的产品数据库，获取与每一所述目标参数相对应的产品推荐信息作为检索文本；

若所述目标参数的属性不为产品属性，基于每一所述目标参数查询联网数据库，获取

与每一所述目标参数相对应的检索文本。

13. 如权利要求 12 所述的计算机设备, 其特征在于, 若所述目标参数的属性为产品属性, 在所述获取与每一所述目标参数相对应的检索文本之后, 所述处理器执行所述计算机程序时还实现如下步骤:

将所述产品推荐信息按预设产品推送格式推送给所述客户端。

14. 如权利要求 9 所述的计算机设备, 其特征在于, 在所述调用信息转化模型对每一所述检索文本进行转换之前, 所述处理器执行所述计算机程序时还实现如下步骤:

获取用户个人信息, 基于所述用户个人信息生成用户画像;

分析所述用户画像, 获取与所述用户画像相对应的用户性格;

基于所述用户性格查询预置回复模式库, 获取与所述用户性格对应的目标回复模式;

将所述目标回复模式添加到所述信息转化模型, 以更新所述信息转化模型。

15. 一个或多个存储有计算机可读指令的可读存储介质, 所述计算机可读存储介质存储有计算机可读指令, 其特征在于, 所述计算机可读指令被一个或多个处理器执行时, 使得所述一个或多个处理器执行如下步骤:

接收客户端发送的携带会话标识的当前轮次信息, 基于所述当前轮次信息获取至少一个当前意图参数;

基于所述会话标识查询会话记录数据库, 获取与所述会话标识相对应的至少一个历史意图参数;

采用预设的强化学习模型对至少一个所述当前意图参数和至少一个所述历史意图参数进行分析, 获取目标意图, 所述目标意图包括至少一个目标参数和参数优先级顺序;

基于每一所述目标参数查询检索文本数据库, 获取与每一所述目标参数相对应的检索文本;

调用信息转化模型对每一所述检索文本进行转换, 获取与每一所述检索文本相对应的当前回复信息, 依据所述参数优先级顺序将至少一个所述当前回复信息推送给所述客户端。

16. 如权利要求 15 所述的可读存储介质, 其特征在于, 所述接收客户端发送的携带会话标识的当前轮次信息, 基于所述当前轮次信息获取至少一个当前意图参数, 包括:

接收客户端发送的当前轮次信息, 调用信息转化模型对所述当前轮次信息进行识别, 获取文字数据;

采用预设语言处理模型对所述文字数据进行解析处理, 获取至少一个当前意图参数。

17. 如权利要求 15 所述的可读存储介质, 其特征在于, 所述采用预设的强化学习模型对至少一个所述当前意图参数和至少一个所述历史意图参数进行分析, 获取目标意图, 包括:

基于至少一个所述当前意图参数和至少一个所述历史意图参数, 分别获取当前词向量矩阵和历史词向量矩阵;

采用预设的强化学习模型对所述当前词向量矩阵和所述历史词向量矩阵进行分析, 获取目标意图。

18. 如权利要求 15 所述的可读存储介质, 其特征在于, 所述基于每一所述目标参数查询检索文本数据库, 获取与每一所述目标参数相对应的检索文本, 包括:

若所述目标参数的属性为产品属性, 则基于每一所述目标参数查询与所述产品属性相对应的产品数据库, 获取与每一所述目标参数相对应的产品推荐信息作为检索文本;

若所述目标参数的属性不为产品属性, 基于每一所述目标参数查询联网数据库, 获取与每一所述目标参数相对应的检索文本。

19. 如权利要求 18 所述的可读存储介质, 其特征在于, 若所述目标参数的属性为产品属性, 在所述获取与每一所述目标参数相对应的检索文本之后, 所述计算机可读指令被一

个或多个处理器执行时，使得所述一个或多个处理器还执行如下步骤：

将所述产品推荐信息按预设产品推送格式推送给所述客户端。

20. 如权利要求 15 所述的可读存储介质，其特征在于，在所述调用信息转化模型对每一所述检索文本进行转换之前，所述计算机可读指令被一个或多个处理器执行时，使得所述一个或多个处理器还执行如下步骤：

获取用户个人信息，基于所述用户个人信息生成用户画像；

分析所述用户画像，获取与所述用户画像相对应的用户性格；

基于所述用户性格查询预置回复模式库，获取与所述用户性格对应的目标回复模式；

将所述目标回复模式添加到所述信息转化模型，以更新所述信息转化模型。

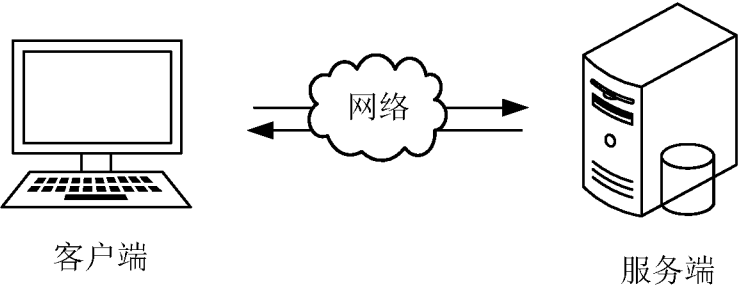


图 1

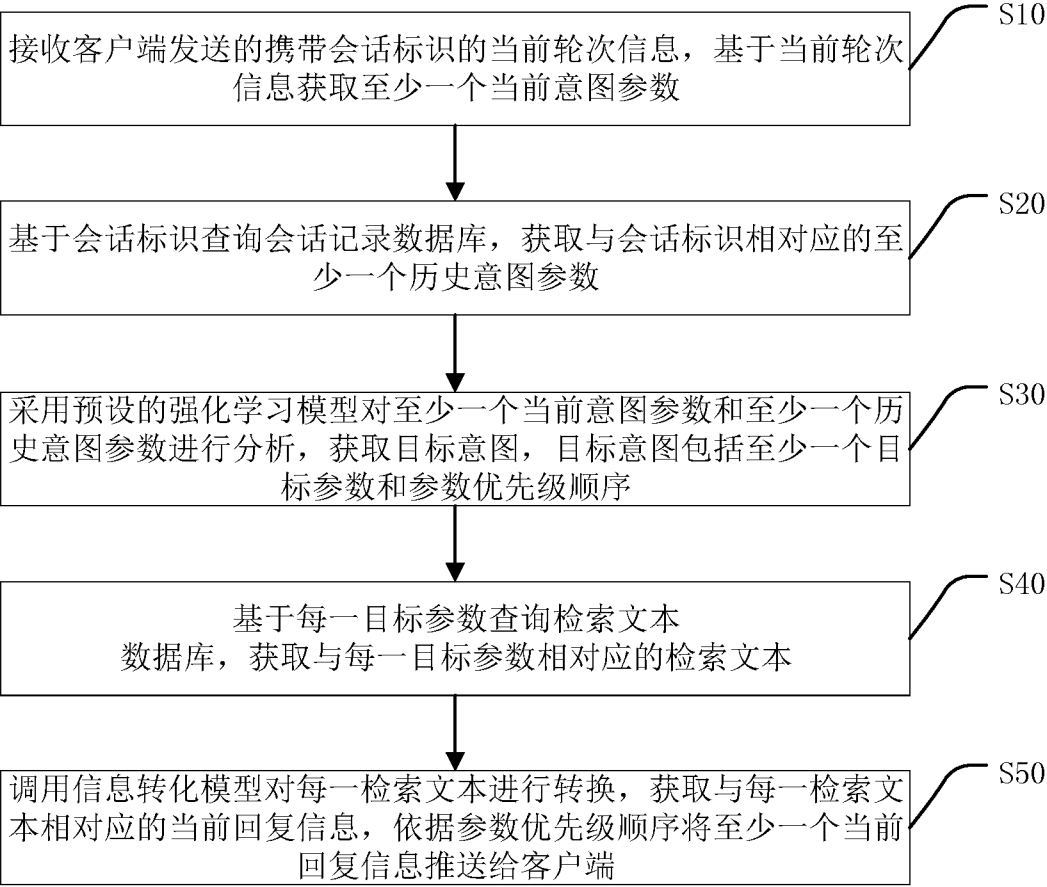
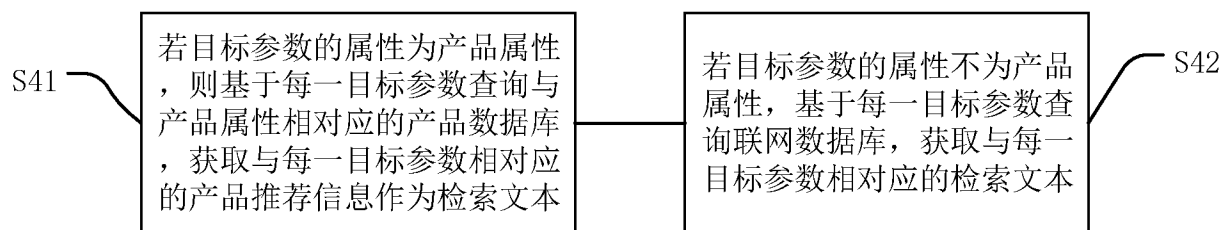
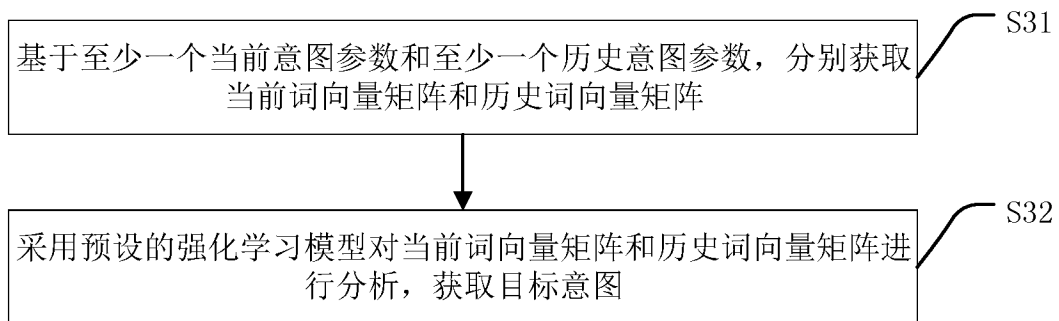
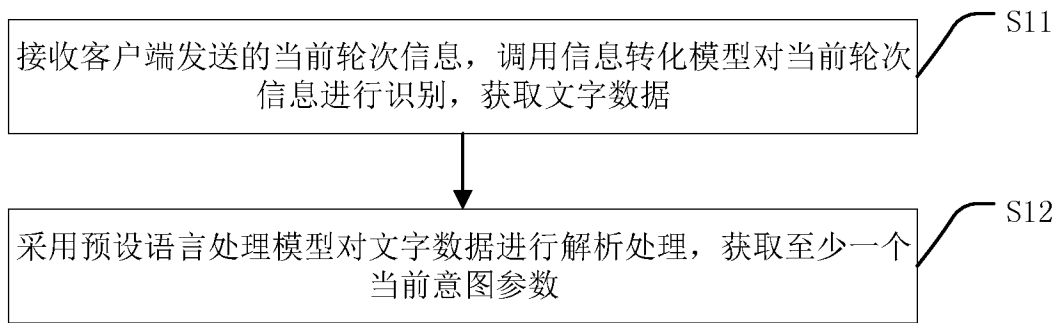


图 2



图 3



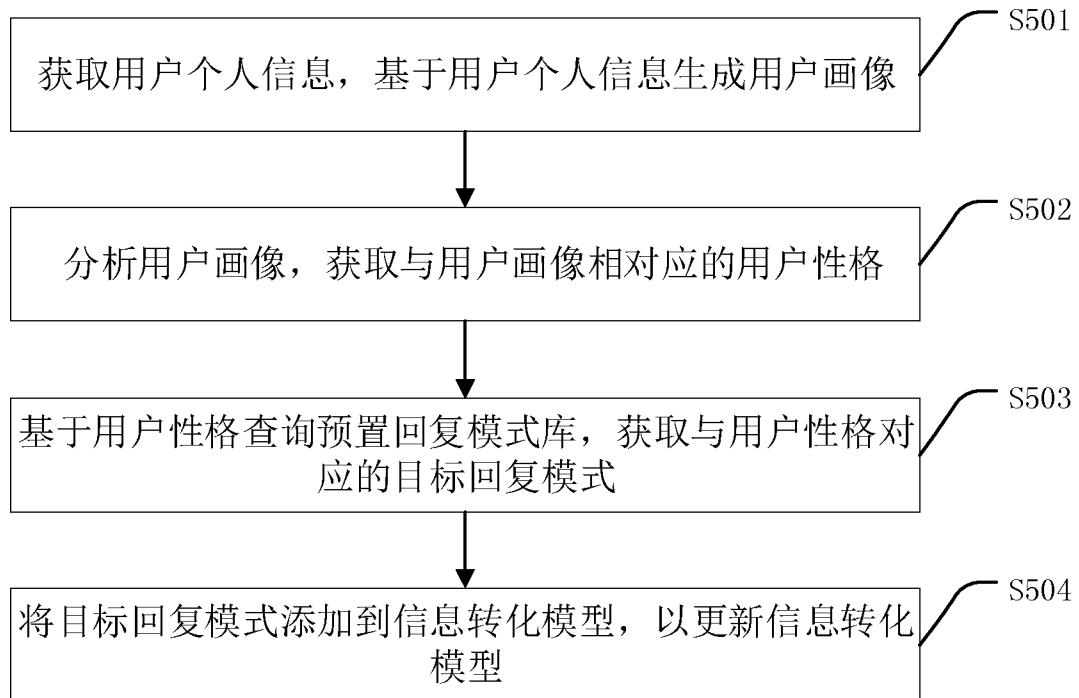


图 7

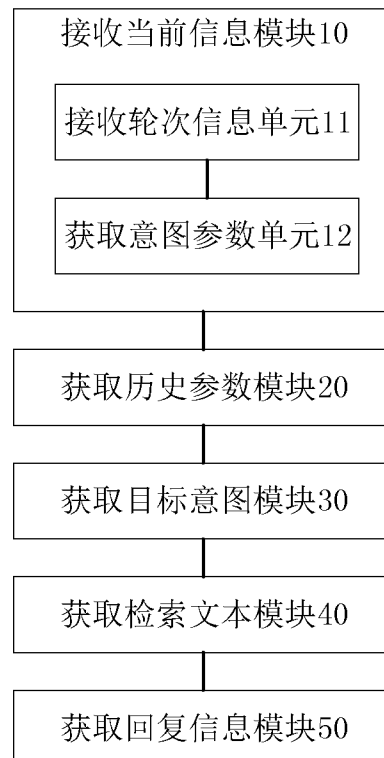


图 8

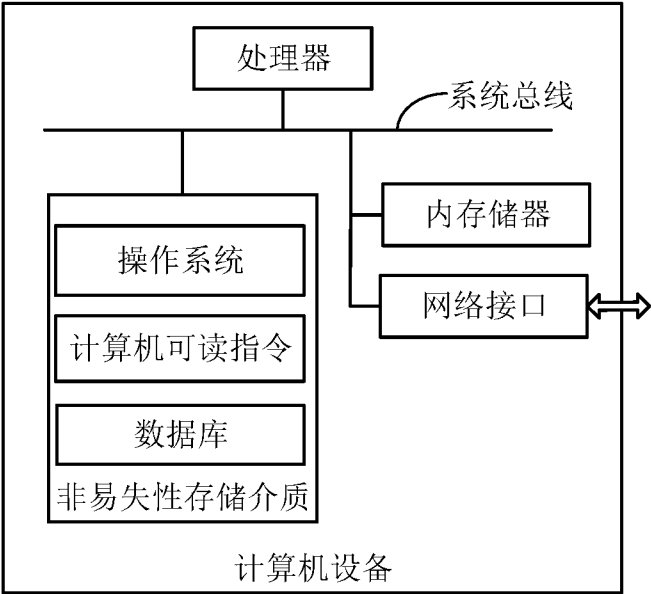


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/120595

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/33(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)

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, WPI, EPODOC, IEEE, CNKI: 会话, 问答, 意图, 识别, 上下文, 历史, 搜索, 查询, 优先级, 顺序, 转化, 转换, dialog, conversation, intention, distinguish, recognise, context, history, search, priority, convert

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109857848 A (SHENZHEN ONECONNECT TECHNOLOGY CO., LTD.) 07 June 2019 (2019-06-07) claims 1-10, and description, paragraphs [0205]-[0207]	1-20
Y	CN 106997342 A (SHANGHAI BENYING NETWORK TECHNOLOGY CO., LTD.) 01 August 2017 (2017-08-01) description, paragraphs [0020]-[0042] and [0094]-[0097]	1-20
Y	CN 107918634 A (SHANGHAI ONECONNECT FINANCIAL TECHNOLOGY CO., LTD.) 17 April 2018 (2018-04-17) description, paragraphs [0092]-[0100]	1-20
A	CN 105068661 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 18 November 2015 (2015-11-18) entire document	1-20
A	US 2018129646 A1 (MICROSOFT TECHNOLOGY LICENSING, LLC) 10 May 2018 (2018-05-10) 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

19 February 2020

Date of mailing of the international search report

28 February 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/CN2019/120595

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109857848	A	07 June 2019	None			
CN	106997342	A	01 August 2017	None			
CN	107918634	A	17 April 2018	HK	1250541	A0	21 December 2018
CN	105068661	A	18 November 2015	WO	2017041372	A1	16 March 2017
US	2018129646	A1	10 May 2018	US	2012253788	A1	04 October 2012

国际检索报告

国际申请号

PCT/CN2019/120595

A. 主题的分类

G06F 16/33 (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)

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, WPI, EPDOC, IEEE, CNKI: 会话, 问答, 意图, 识别, 上下文, 历史, 搜索, 查询, 优先级, 顺序, 转化, 转换, dialog, conversation, intention, distinguish, recognise, context, history, search, priority, convert

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 109857848 A (深圳壹账通智能科技有限公司) 2019年 6月 7日 (2019 - 06 - 07) 权利要求1-10, 说明书第[0205]-[0207]段	1-20
Y	CN 106997342 A (上海奔影网络科技有限公司) 2017年 8月 1日 (2017 - 08 - 01) 说明书第[0020]-[0042]段, 第[0094]-[0097]段	1-20
Y	CN 107918634 A (上海壹账通金融科技有限公司) 2018年 4月 17日 (2018 - 04 - 17) 说明书第[0092]-[0100]段	1-20
A	CN 105068661 A (百度在线网络技术北京有限公司) 2015年 11月 18日 (2015 - 11 - 18) 全文	1-20
A	US 2018129646 A1 (MICROSOFT TECHNOLOGY LICENSING, LLC) 2018年 5月 10日 (2018 - 05 - 10) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 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 19日

国际检索报告邮寄日期

2020年 2月 28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

受权官员

任兴超

电话号码 86-(10)-53961375

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/120595

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	109857848	A	2019年 6月 7日	无			
CN	106997342	A	2017年 8月 1日	无			
CN	107918634	A	2018年 4月 17日	HK	1250541	A0	2018年 12月 21日
CN	105068661	A	2015年 11月 18日	WO	2017041372	A1	2017年 3月 16日
US	2018129646	A1	2018年 5月 10日	US	2012253788	A1	2012年 10月 4日