



(12) 发明专利申请

(10) 申请公布号 CN 115309874 A

(43) 申请公布日 2022. 11. 08

(21) 申请号 202210788319.X

(22) 申请日 2022.07.06

(71) 申请人 京东科技信息技术有限公司

地址 100000 北京市北京经济技术开发区
科创十一街18号院2号楼6层601

(72) 发明人 袁韶祖 王培英 陈蒙

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

专利代理师 周初冬

(51) Int. Cl.

G06F 16/332 (2019.01)

G06N 3/04 (2006.01)

G06N 3/08 (2006.01)

G10L 25/48 (2013.01)

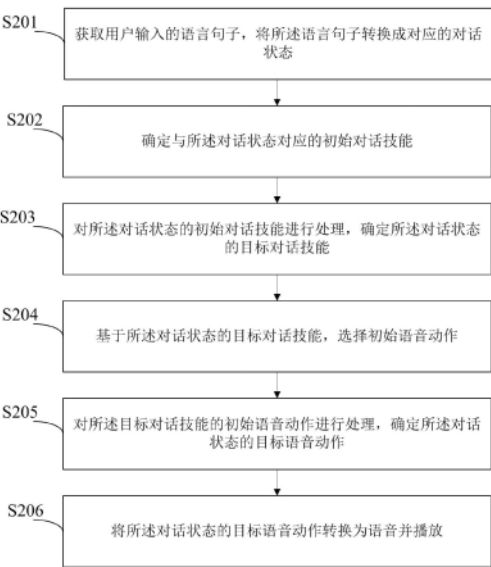
权利要求书2页 说明书10页 附图4页

(54) 发明名称

对话处理方法、对话系统、电子设备及计算机存储介质

(57) 摘要

本发明提供一种对话处理方法、对话系统、电子设备及计算机存储介质,包括将用户输入语音转换成对应的对话状态;确定对应的初始对话技能;对初始对话技能进行处理,确定目标对话技能;基于目标对话技能,选择初始语音动作;对初始语音动作进行处理,确定目标语音动作;将目标语音动作转换成对应的对话语音并播放。本方案提出了利用层次强化学习的方法,以构建顶层策略和底层策略的两层对话策略,通过顶层策略选择初始对话技能,进而处理得到目标对话技能;通过底层策略基于目标对话技能选择初始语音动作,进而处理得到目标语音动作。通过上述方法可以根据不同的对话场景调用相应的对话技能,并进行相应的话题引导,从而实现流畅对话的目的。



1. 一种对话处理方法,其特征在于,应用于对话系统,所述方法包括:
获取用户输入的语音,将所述语音转换成对应的对话状态;
从技能集合中确定与所述对话状态对应的初始对话技能,所述技能集合是基于对话类型对历史对话数据进行处理得到;
对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能;
基于所述对话状态的目标对话技能,选择初始语音动作;
对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作;
将所述对话状态的目标语音动作转换成对应的对话语音并播放。
2. 根据权利要求1所述的方法,其特征在于,还包括:
对所述语音状态进行跟踪,确定对应的历史语音状态;
对所述历史语音状态和所述语音状态进行分析处理,确定目标语音动作。
3. 根据权利要求1所述的方法,其特征在于,所述基于对话类型对历史对话数据进行处理得到技能集合,包括:
按照对话类型将对话系统中历史对话数据进行划分,分成不同对话类型的对话数据;
针对与每一对话类型,按照不同的动作类型对所述对话类型的对话数据进行划分,确定每一技能对应的对话数据。
4. 根据权利要求1所述的方法,其特征在于,所述对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能,包括:
将所述对话状态的初始对话技能作为第一强化学习DQN网络的输入;
第一DQN网络对输入的所述对话状态的初始对话技能进行计算,得到对应第一价值函数;
基于所述第一价值函数更新所述第一DQN网络的网络参数,以重新构建第一DQN网络;
利用更新后的所述第一DQN网络对所述对话状态的初始对话技能进行处理,输出目标对话技能。
5. 根据权利要求1所述的方法,其特征在于,所述对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作,包括:
将所述目标对话技能的初始语音动作输入第二强化学习DQN网络的输入;
第二DQN网络对输入的所述目标对话技能的初始语音动作进行计算,得到对应第二价值函数;
基于所述第二价值函数更新所述第二DQN网络的网络参数,以重新构建第二DQN网络;
利用更新后的所述第二DQN网络对所述目标对话技能的初始语音动作进行处理,输出目标语音动作。
6. 一种对话系统,所述对话系统包括:
语言理解模块,用于获取用户输入的语音,将所述语音转换成对应的对话状态;
对话管理模块,用于从技能集合中确定与所述对话状态对应的初始对话技能,所述技能集合是基于对话类型对历史对话数据进行处理得到;对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能;基于所述对话状态的目标对话技能,选择初始语音动作;对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作;

自然语言生成模块,用于将所述对话状态的目标语音动作转换成对应的对话语音并播放。

7.根据权利要求6所述的系统,其特征在于,所述对话管理模块,还用于:对所述语音状态进行跟踪,确定对应的历史语音状态;对所述历史语音状态和所述语音状态进行分析处理,确定目标语音动作。

8.根据权利要求6所述的系统,其特征在于,所述对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能的对话管理模块,具体用于:将所述对话状态的初始对话技能作为第一强化学习DQN网络的输入;

第一DQN网络对输入的所述对话状态的初始对话技能进行计算,得到对应第一价值函数;

基于所述第一价值函数更新所述第一DQN网络的网络参数,以重新构建第一DQN网络;

利用更新后的所述第一DQN网络对所述对话状态的初始对话技能进行处理,输出目标对话技能。

9.一种电子设备,其特征在于,所述电子设备用于运行程序,其中,所述程序运行时执行如权利要求1-5中任一所述的对话处理方法。

10.一种计算机存储介质,其特征在于,所述存储介质包括存储程序,其中,在所述程序运行时控制所述存储介质所在设备执行如权利要求1-5中任一所述的对话处理方法。

对话处理方法、对话系统、电子设备及计算机存储介质

技术领域

[0001] 本发明涉及语音处理技术领域,尤其涉及一种对话处理方法、对话系统、电子设备及计算机存储介质。

背景技术

[0002] 通常,对话可以被理解为一个序列决策过程,机器需要在对话过程中,通过理解用户语句更新维护内部的对话状态,再根据当前的对话状态选择下一步的最优动作(例如确认需求,询问限制条件,提供结果等等),从而完成任务。

[0003] 现有对话系统是基于某种单一对话类型的对话,没有考虑如何同时处理多个对话类型,例如闲聊、问答、任务、和推荐等。但在实际的人人对话中,通常覆盖了多种不同类型的对话片段;若直接将多个类型的对话数据简单组合,会导致在对话过程中,容易出现对话不流畅的问题。

[0004] 有鉴于此,如何提供一种对话处理方法对多个类型的对话数据进行管理,以实现对话流畅是本发明亟待解决的问题。

发明内容

[0005] 有鉴于此,本发明实施例提供一种对话处理方法、对话系统、电子设备及计算机存储介质,以解决现有技术存在的对话不流畅的问题。

[0006] 为实现上述目的,本发明实施例提供如下技术方案:

[0007] 本发明实施例第一方面示出了一种对话处理方法,应用于对话系统,所述方法包括:

[0008] 获取用户输入的语音,将所述语音转换成对应的对话状态;

[0009] 从技能集合中确定与所述对话状态对应的初始对话技能,所述技能集合是基于对话类型对历史对话数据进行处理得到;

[0010] 对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能;

[0011] 基于所述对话状态的目标对话技能,选择初始语音动作;

[0012] 对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作;

[0013] 将所述对话状态的目标语音动作转换成对应的对话语音并播放。

[0014] 可选的,还包括:

[0015] 对所述语音状态进行跟踪,确定对应的历史语音状态;

[0016] 对所述历史语音状态和所述语音状态进行分析处理,确定目标语音动作。

[0017] 可选的,所述基于对话类型对历史对话数据进行处理得到技能集合,包括:

[0018] 按照对话类型将对话系统中历史对话数据进行划分,分成不同对话类型的对话数据;

[0019] 针对与每一对话类型,按照不同的动作类型对所述对话类型的对话数据进行划

分,确定每一技能对应的对话数据。

[0020] 可选的,所述对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能,包括:

[0021] 将所述对话状态的初始对话技能作为第一强化学习DQN网络的输入;

[0022] 第一DQN网络对输入的所述对话状态的初始对话技能进行计算,得到对应第一价值函数;

[0023] 基于所述第一价值函数更新所述第一DQN网络的网络参数,以重新构建第一DQN网络;

[0024] 利用更新后的所述第一DQN网络对所述对话状态的初始对话技能进行处理,输出目标对话技能。

[0025] 可选的,所述对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作,包括:

[0026] 将所述目标对话技能的初始语音动作输入第二强化学习DQN网络的输入;

[0027] 第二DQN网络对输入的所述目标对话技能的初始语音动作进行计算,得到对应第二价值函数;

[0028] 基于所述第二价值函数更新所述第二DQN网络的网络参数,以重新构建第二DQN网络;

[0029] 利用更新后的所述第二DQN网络对所述目标对话技能的初始语音动作进行处理,输出目标语音动作。

[0030] 本发明实施例第二方面示出了一种对话系统,所述对话系统包括:

[0031] 语言理解模块,用于获取用户输入的语音,将所述语音转换成对应的对话状态;

[0032] 对话管理模块,用于从技能集合中确定与所述对话状态对应的初始对话技能,所述技能集合是基于对话类型对历史对话数据进行处理得到;对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能;基于所述对话状态的目标对话技能,选择初始语音动作;对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作;

[0033] 自然语言生成模块,用于将所述对话状态的目标语音动作转换成对应的对话语音并播放。

[0034] 可选的,所述对话管理模块,还用于:对所述语音状态进行跟踪,确定对应的历史语音状态;对所述历史语音状态和所述语音状态进行分析处理,确定目标语音动作。

[0035] 可选的,所述对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能的对话管理模块,具体用于:将所述对话状态的初始对话技能作为第一强化学习DQN网络的输入;

[0036] 第一DQN网络对输入的所述对话状态的初始对话技能进行计算,得到对应第一价值函数;

[0037] 基于所述第一价值函数更新所述第一DQN网络的网络参数,以重新构建第一DQN网络;

[0038] 利用更新后的所述第一DQN网络对所述对话状态的初始对话技能进行处理,输出目标对话技能。

[0039] 本发明实施例第三方面示出了一种电子设备,所述电子设备用于运行程序,其中,所述程序运行时执行如本发明实施例第一发明示出的对话处理方法。

[0040] 本发明实施例第四方面示出了一种计算机存储介质,所述存储介质包括存储程序,其中,在所述程序运行时控制所述存储介质所在设备执行如本发明实施例第一发明示出的对话处理方法。

[0041] 基于上述本发明实施例提供的一种对话处理方法、对话系统、电子设备及计算机存储介质,该方法包括:获取用户输入的语音,将所述语音转换成对应的对话状态;确定与所述对话状态对应的初始对话技能,所述技能集合是基于对话类型对历史对话数据进行处理得到;对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能;基于所述对话状态的目标对话技能,选择初始语音动作;对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作;将所述对话状态的目标语音动作转换成对应的对话语音并播放。在本发明实施例提出了利用层次强化学习的方法,以构建顶层策略和底层策略的两层对话策略,通过顶层策略基于转换得到的对话状态选择初始对话技能,接着对初始对话技能进行处理,确定所述对话状态的目标对话技能;通过底层策略基于目标对话技能选择初始语音动作,以对初始语音动作进行处理,确定所述对话状态的目标语音动作;最后将目标语音动作转换成对应的对话语音并播放。通过上述示出的方法可以根据不同的对话场景调用相应的对话技能,并进行相应的话题引导,从而实现流畅对话的目的。

附图说明

[0042] 为了更清楚地说明本发明实施例或现有技术中的技术方案,下面将对实施例或现有技术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发明的实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据提供的附图获得其他的附图。

[0043] 图1为本发明实施例示出的一种对话系统的结构示意图;

[0044] 图2为本发明实施例示出的一种对话处理方法的流程示意图;

[0045] 图3为本发明实施例示出的确定目标对话技能的流程示意图;

[0046] 图4为本发明实施例示出的确定目标语音动作的流程示意图;

[0047] 图5为本发明实施例示出的另一种对话处理方法的流程示意图。

具体实施方式

[0048] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0049] 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”和“具有”以及他们的任何变形,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。

[0050] 需要说明的是,在本发明中涉及“第一”、“第二”等的描述仅用于描述目的,而不能理解为指示或暗示其相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。另外,各个实施例之间的技术方案可以相互结合,但是必须是以本领域普通技术人员能够实现为基础,当技术方案的结合出现相互矛盾或无法实现时应当认为这种技术方案的结合不存在,也不在本发明要求的保护范围之内。

[0051] 在本申请中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0052] 在本发明实施例中,常见的对话系统可分为三类:聊天型,任务导向型和问答型。聊天型对话的目标是要产生有趣且富有信息量的自然回复使得人机对话可以持续进行下去;问答型对话多指一问一答,用户提出一个问题,系统通过对问题进行解析和知识库查找以返回正确答案;任务导向型对话则是指由任务驱动的多轮对话,机器需要通过理解、主动询问、澄清等方式来确定用户的目标,调用相应的API查询后,返回正确结果,完成用户需求。

[0053] 参见图1,为本发明实施例示出的一种对话系统的结构示意图。

[0054] 在本发明实施例示出的对话系统可部署于机器人等应用设备内,该对话系统为融合了多种对话技能的机器人。

[0055] 所述对话系统包括输入端10、语言理解模块NLU20、自然语言生成模块NLG30、和对话管理模块DM40。

[0056] 对话管理模块DM40中存储有由于多个技能组成的技能集合。具体的,所述技能集合的构建过程为:按照对话类型将对话系统中历史对话数据进行划分,分成不同对话类型的数据;针对与每一对话类型,按照不同的动作类型对所述对话类型的对话数据进行划分,确定每一技能对应的对话数据,即确定多个技能,

[0057] 需要说明的是,每一技能为每一大类基本动作的集合;将每一对话类型下的技能组成技能集合G,每一对话类型下的技能数量为多个。

[0058] G可表示为{任务导向型Task-Domain1,Task-Domain2,...,聊天型Chat-Topic1,Chat-Topic2,...,问答型QA1,QA2...},其中,任务导向型Task-Domain1,Task-Domain2等为任务导向型的技能;聊天型Chat-Topic1,Chat-Topic2等为聊天型的技能;问答型QA1,QA2等为问答型的技能。

[0059] 进一步,需要说明的是,对话类型至少包括任务导向型、聊天型和问答型等。

[0060] 构建不同的对话状态以及对话类型之间的映射关系。

[0061] 所述对话管理模块DM40还包括全局对话状态追踪模块global DST、顶层策略top-levelpolicy、及底层策略low-levelpolicy。

- [0062] 其中,全局对话状态追踪global DST模块,用于对全局对话状态进行跟踪与记录。
- [0063] 顶层策略top-levelpolicy,用于根据当前对话状态选择应答的技能。
- [0064] 底层策略low-levelpolicy,用于完成某技能的某个具体对话动作。
- [0065] 基于上述本发明实施例示出的对话系统具体实现对话处理方法的过程包括:
- [0066] 语言理解模块NLU20,用于获取用户输入的语音,将所述语音转换成对应的对话状态;
- [0067] 对话管理模块DM40,用于基于顶层策略top-levelpolicy从技能集合中确定与所述对话状态对应的初始对话技能,所述技能集合是基于对话类型对历史对话数据进行处理得到,并对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能;基于底层策略low-levelpolicy从所述对话状态的目标对话技能中选择初始语音动作,并对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作。
- [0068] 自然语言生成模块NLG30,用于将所述对话状态的目标语音动作转换成对应的对话语音并播放。
- [0069] 在具体实现中,将输出的所述对话状态的目标语音动作转换为语音并向用户播放,以形成对话。
- [0070] 在本发明实施例中,利用层次强化学习的方法,以构建顶层策略和底层策略的两层对话策略,通过顶层策略基于转换得到的对话状态选择初始对话技能,接着对初始对话技能进行处理,确定所述对话状态的目标对话技能;通过底层策略基于目标对话技能选择初始语音动作,以对初始语音动作进行处理,确定所述对话状态的目标语音动作;最后将目标语音动作转换成对应的对话语音并播放。通过上述示出的方法可以根据不同的对话场景调用相应的对话技能,并进行相应的话题引导,从而实现流畅对话的目的。
- [0071] 可选的,基于上述本发明实施例示出的对话系统,对话管理模块DM40,还用于:对所述语音状态进行跟踪,确定对应的历史语音状态;对所述历史语音状态和所述语音状态进行分析处理,确定目标语音动作。
- [0072] 在具体实现中,利用全局对话追踪对语音状态进行跟踪,以确定与所述语音状态相关的历史语音状态,并对所述历史语音状态和所述语音状态进行处理,以确定每一历史语音状态和语音状态的对话意图、槽填充状态、对话话题、以及调用的对话模块等。基于每一历史语音状态和语音状态的对话意图、槽填充状态、对话话题、以及调用的对话模块的语义,选择对应对话策略,即目标语音动作,并输出。
- [0073] 在本发明实施例中,基于全局对话状态追踪的方式对语音状态进行跟踪,以确定与所述语音状态相关的历史语音状态,进而对其进行分析,确定对话意图、槽填充状态、对话话题、以及调用的对话模块等,选择对应对话策略,即目标语音动作,并输出。通过上述示出的方法可以根据不同的对话场景调用相应的对话技能,并进行相应的话题引导,从而实现流畅对话的目的。
- [0074] 可选的,基于上述本发明实施例示出的对话系统,对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能的对话管理模块DM40,具体用于:
- [0075] 将所述对话状态的初始对话技能作为第一强化学习DQN网络的输入;第一DQN网络对输入的所述对话状态的初始对话技能进行计算,得到对应第一价值函数;基于所述第一价值函数更新所述第一DQN网络的网络参数,以重新构建第一DQN网络;利用更新后的所述

第一DQN网络对所述对话状态的初始对话技能进行处理,输出目标对话技能。

[0076] 需要说明的是,针对不同类型的对话技能,根据多次经验或实验设置不同的奖励信号,即不同的奖励的折扣因子和外部奖励。

[0077] 在具体实现中,通过顶层策略将所述对话状态的初始对话技能输入第一强化学习DQN网络的输入。获取所述初始对话技能对应的奖励的折扣因子和外部奖励,将所述初始对话技能对应的奖励的折扣因子和外部奖励,所述对话状态的初始对话技能,以及记录的上一对话状态的历史目标对话技能 g' 代入公式(1)进行计算,得到第一价值函数 $Q_1(S, g)$,也就是说,在对话状态 S 下执行技能 g 时奖励函数的期望,即最大化实际得到的外部奖励之和。利用计算得到第一价值函数作为每一步骤的反馈对奖励函数进行更新,即对第一DQN网络的网络参数进行更新,从而调整第一DQN网络,进而基于更新后的网络参数重新训练第一DQN网络。执行重新训练得到的第一DQN网络,以对所述对话状态的初始对话技能进行处理,输出目标对话技能。

[0078] 公式(1):

$$[0079] \quad Q_1(s, g) = E \left[\sum_{k=0}^{N-1} \gamma^k r_{t+k}^e + \gamma^N \cdot \max_g Q_1(s_{t+N}, g') | s_t = s, g_t = g \right] \quad (1)$$

[0080] 其中, γ 为奖励的折扣因子, r^e 为外部奖励, N 指的是技能完成或者结束需要的时间步,也就是说这里将 N 步获得的外在环境激励作为一个整体,作为顶层策略的奖励reward,为记录的上一对话状态的历史目标对话技能, S 为对话状态, S_t 为 t 时刻的对话状态, E 为期望。

[0081] 进一步的,当前时刻的对话状态 S 受上一时刻的对话状态的影响。

[0082] 在本发明实施例中,利用层次强化学习的方法,以构建顶层策略和底层策略的两层对话策略,顶层策略利用第一强化学习DQN网络调整初始对话技能,以确定对应的目标对话技能;通过底层策略基于目标对话技能选择初始语音动作,以对初始语音动作进行处理,确定所述对话状态的目标语音动作;最后将目标语音动作转换成对应的对话语音并播放。通过上述示出的方法可以根据不同的对话场景调用相应的对话技能,并进行相应的话题引导,从而实现流畅对话的目的。

[0083] 可选的,基于上述本发明实施例示出的对话系统,所述对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作,的对话管理模块DM40,具体用于:

[0084] 将所述目标对话技能的初始语音动作输入第二强化学习DQN网络的输入;第二DQN网络对输入的所述目标对话技能的初始语音动作进行计算,得到对应第二价值函数;基于所述第二价值函数更新所述第二DQN网络的网络参数,以重新构建第二DQN网络;利用更新后的所述第二DQN网络对所述目标对话技能的初始语音动作进行处理,输出目标语音动作。

[0085] 在具体实现中,通过底层策略将所述目标对话技能的初始语音动作输入第二强化学习DQN网络的输入。获取所述目标对话技能对应的奖励的折扣因子和目标对话技能当前时刻 t 的奖励,将所述初始对话技能对应的奖励的折扣因子和目标对话技能当前时刻 t 的奖励,所述初始语音动作和目标对话技能 g ,以该对话状态在 $t+1$ 时刻的对话状态 S_{t+1} ,初始语音动作 a 在 $t+1$ 时刻的语音动作 a_{t+1} 代入公式(2)进行计算,得到第二价值函数 $Q_2(S, a, g)$,也

就是说,在对话状态S下执行目标技能g的初始语音动作时奖励函数的期望,即最大化实际得到的外部奖励之和。利用计算得到第二价值函数作为每一步骤的反馈对奖励函数进行更新,即对第二DQN网络的网络参数进行更新,从而调整第二DQN网络,进而基于更新后的网络参数重新训练第二DQN网络。执行重新训练得到的第二DQN网络,以对所述对话状态的初始语音动作进行处理,输出目标语音动作。

[0086] 公式(2):

$$[0087] \quad Q_2(s, a, g) = E \left[\sum_{k=0}^{N-1} r_t^i + \gamma \cdot \max_{a_{t+1}} Q_2(s_{t+1}, a_{t+1}, g) | s_t = s, g_t = g \right] \quad (2)$$

[0088] 其中, γ 为奖励的折扣因子, r_t^i 为目标对话技能当前时刻t的奖励, N指的是动作完成或者结束需要的时间步, s_t 为记录的上一对话状态的历史目标对话技能, S为对话状态, s_{t+1} 为该对话状态在t+1时刻的对话状态, a_{t+1} 为初始语音动作a在t+1时刻的语音动作, g为目标对话技能, E为期望。

[0089] 可选的,还包括:将当前的对话状态,该对话状态对应的目标对话技能,以及该目标对话技能中的目标语音动作进行记录。

[0090] 在本发明实施例中,利用层次强化学习的方法,以构建顶层策略和底层策略的两层对话策略,顶层策略利用第一强化学习DQN网络调整初始对话技能,以确定对应的目标对话技能;通过底层策略利用第二DQN调整初始语音动作,确定所述对话状态的目标语音动作;最后将目标语音动作转换成对应的对话语音并播放。通过上述示出的方法可以根据不同的对话场景调用相应的对话技能,并进行相应的话题引导,从而实现流畅对话的目的。

[0091] 参见图2,为本发明实施例示出的一种对话处理方法的流程示意图,该方法包括:

[0092] 步骤S201:获取用户输入的语言句子,将所述语言句子转换成对应的对话状态。

[0093] 在具体实现步骤S201的过程中,获取用户在与机器人对话时,所产生的语句,通过语言理解模块对所述语句对应的语言句子进行识别,以确定该句子的用户意图,以及提取槽信息,进而确定对应的对话状态。

[0094] 需要说明的是,通过该对话状态能够确定对应的对话类型。

[0095] 步骤S202:确定与所述对话状态对应的初始对话技能。

[0096] 在具体实现步骤S202的过程中,遍历所述技能集合,查找与所述对话状态对应的对话类型,进而确定与所述对话类型对应动作类型;基于所述动作类型从所述技能集合中选择一技能为初始对话技能。

[0097] 需要说明的是,初始对话技能g属于G。

[0098] 步骤S203:对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能。

[0099] 需要说明的是,具体实现步骤S203对所述对话状态的初始对话技能进行处理,确定所述对话状态的目标对话技能的过程,如图3所示,包括以下步骤:

[0100] 步骤S301:将所述对话状态的初始对话技能作为第一强化学习DQN网络的输入。

[0101] 在具体实现步骤S301的过程中,将所述对话状态的初始对话技能输入第一强化学习DQN网络的输入。

[0102] 步骤S302:第一DQN网络对输入的所述对话状态的初始对话技能进行计算,得到对

应第一价值函数。

[0103] 需要说明的是,针对不同类型的对话技能,根据多次经验或实验设置不同的奖励信号,即不同的奖励的折扣因子和外部奖励。

[0104] 在具体实现步骤S302的过程中,获取所述初始对话技能对应的奖励的折扣因子和外部奖励,将所述初始对话技能对应的奖励的折扣因子和外部奖励,所述对话状态的初始对话技能,以及记录的上一对话状态的历史目标对话技能 g' 代入公式(1)进行计算,得到第一价值函数 $Q_1(S, g)$,也就是说,在对话状态S下执行技能g时奖励函数的期望,即最大化实际得到的外部奖励之和。

[0105] 进一步的,当前时刻的对话状态S受上一时刻的对话状态的影响。

[0106] 可选的,还包括:将当前的对话状态,以及该对话状态对应的目标对话技能进行记录。

[0107] 步骤S303:基于所述第一价值函数更新所述第一DQN网络的网络参数,以重新构建第一DQN网络。

[0108] 在具体实现步骤S303的过程中,利用计算得到第一价值函数作为每一步骤的反馈对奖励函数进行更新,即对第一DQN网络的网络参数进行更新,从而调整第一DQN网络,进而基于更新后的网络参数重新训练第一DQN网络。

[0109] 步骤S304:利用更新后的所述第一DQN网络对所述对话状态的初始对话技能进行处理,输出目标对话技能。

[0110] 在具体实现步骤S304的过程中,执行重新训练得到的第一DQN网络,以对所述对话状态的初始对话技能进行处理,输出目标对话技能。

[0111] 在本发明实施例中,对于用户兴趣一致性以及用户提及新话题时,机器人bot中的对话系统应该相应调整对应的话题。

[0112] 基于上述步骤S301至步骤S304的过程,以基于顶层策略指定对话主题,即对话状态进行一系列对话。

[0113] 步骤S204:基于所述对话状态的目标对话技能,选择初始语音动作。

[0114] 在具体实现步骤S204的过程中,接收上一层传达的目标对话技能g以及当前的状态s,然后选择一个可能的系统动作a,即初始语音动作。

[0115] 步骤S205:对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作。

[0116] 需要说明的是,具体实现步骤S205对所述目标对话技能的初始语音动作进行处理,确定所述对话状态的目标语音动作的过程中,如图4所示,包括以下步骤:

[0117] 步骤S401:将所述目标对话技能的初始语音动作输入第二强化学习DQN网络的输入。

[0118] 需要说明的是,具体实现步骤S401的过程与上述步骤S301的过程相同,可相互参见。

[0119] 步骤S402:第二DQN网络对输入的所述目标对话技能和所述初始语音动作进行计算,得到对应第二价值函数。

[0120] 在具体实现步骤S402的过程中,获取所述目标对话技能对应的奖励的折扣因子和目标对话技能当前时刻t的奖励,将所述初始对话技能对应的奖励的折扣因子和目标对话

技能当前时刻 t 的奖励,所述初始语音动作和目标对话技能 g ,以该对话状态在 $t+1$ 时刻的对话状态 S_{t+1} ,初始语音动作 a 在 $t+1$ 时刻的语音动作 a_{t+1} 代入公式(2)进行计算,得到第二价值函数 $Q_2(S,a,g)$,也就是说,在对话状态 S 下执行目标技能 g 的初始语音动作时奖励函数的期望,即最大化实际得到的外部奖励之和。

[0121] 可选的,还包括:将当前的对话状态,该对话状态对应的目标对话技能,以及该目标对话技能中的目标语音动作进行记录。

[0122] 步骤S403:基于所述第二价值函数更新所述第二DQN网络的网络参数,以重新构建第二DQN网络。

[0123] 步骤S404:利用更新后的所述第二DQN网络对所述目标对话技能的初始语音动作进行处理,输出目标语音动作。

[0124] 需要说明的是,步骤S403至步骤S404的具体实现过程与上述步骤S303至步骤S304的过程相同。

[0125] 可选的,若目标语音动作为知识型问答的动作则直接对应其应答模块。

[0126] 步骤S206:将所述对话状态的目标语音动作转换为语音并播放。

[0127] 在具体实现步骤S206的过程中,将输出的所述对话状态的目标语音动作转换为语音并向用户播放,以形成对话。

[0128] 在本发明实施例中,利用层次强化学习的方法,以构建顶层策略和底层策略的两层对话策略,通过顶层策略基于转换得到的对话状态选择初始对话技能,接着对初始对话技能进行处理,确定所述对话状态的目标对话技能;通过底层策略基于目标对话技能选择初始语音动作,以对初始语音动作进行处理,确定所述对话状态的目标语音动作;最后将目标语音动作转换成对应的对话语音并播放。通过上述示出的方法可以根据不同的对话场景调用相应的对话技能,并进行相应的话题引导,从而实现流畅对话的目的。

[0129] 基于上述示出的对话处理方法,本发明实施例还对应公开了另一对话处理方法,如图5所示,为本发明实施例示出的一种对话处理方法的流程示意图,该方法包括:

[0130] 步骤S501:获取用户输入的语言句子,将所述语言句子转换成对应的对话状态。

[0131] 需要说明的是,具体实现步骤S501的过程与上述步骤S201的具体实现过程

[0132] 步骤S502:对所述语音状态进行跟踪,确定对应的历史语音状态。

[0133] 在具体实现步骤S502的过程中,利用全局对话追踪对语音状态进行跟踪,以确定与所述语音状态相关的历史语音状态。

[0134] 步骤S503:对所述历史语音状态和所述语音状态进行分析处理,确定目标语音动作。

[0135] 在具体实现步骤S503的过程中,利用全局对话追踪对所述历史语音状态和所述语音状态进行处理,以确定每一历史语音状态和语音状态的对话意图、槽填充状态、对话话题、以及调用的对话模块等。基于每一历史语音状态和语音状态的对话意图、槽填充状态、对话话题、以及调用的对话模块的语义,选择对应对话策略,即目标语音动作,并输出。

[0136] 可选的,针对对话话题序列的连贯性,本发明实施例利用额外的知识图谱等信息进行语音衡量。

[0137] 步骤S504:将所述对话状态的目标语音动作转换为语音并播放。

[0138] 需要说明的是,具体实现步骤S504的过程与上述具体实现步骤S206的具体实现过

程相同,可相互参见。

[0139] 在本发明实施例中,基于全局对话状态追踪的方式对语音状态进行跟踪,以确定与所述语音状态相关的历史语音状态,进而对其进行分析,确定对话意图、槽填充状态、对话话题、以及调用的对话模块等,选择对应对话策略,即目标语音动作,并输出。通过上述示出的方法可以根据不同的对话场景调用相应的对话技能,并进行相应的话题引导,从而实现流畅对话的目的。

[0140] 本发明实施例还公开了一种电子设备,该电子设备用于运行数据库存储过程,其中,所述运行数据库存储过程时执行上述图2公开的数据处理方法。

[0141] 本发明实施例还公开了一种计算机存储介质,所述存储介质包括存储数据库存储过程,其中,在所述数据库存储过程运行时控制所述存储介质所在设备执行上述图2公开的数据处理方法。

[0142] 在本公开的上下文中,计算机存储介质可以是有形的介质,其可以包含或存储以供指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合地使用的程序。机器可读介质可以是机器可读信号介质或机器可读储存介质。机器可读介质可以包括但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置或设备,或者上述内容的任何合适组合。机器可读存储介质的更具体示例会包括基于一个或多个线的电气连接、便携式计算机盘、硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM或快闪存储器)、光纤、便捷式紧凑盘只读存储器(CD-ROM)、光学储存设备、磁储存设备、或上述内容的任何合适组合。

[0143] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于系统或系统实施例而言,由于其基本相似于方法实施例,所以描述得比较简单,相关之处参见方法实施例的部分说明即可。以上所描述的系统及系统实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0144] 专业人员还可以进一步意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、计算机软件或者二者的结合来实现,为了清楚地说明硬件和软件的可互换性,在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本发明的范围。

[0145] 对所公开的实施例的上述说明,使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的,本文中所定义的一般原理可以在不脱离本发明的精神或范围的情况下,在其它实施例中实现。因此,本发明将不会被限制于本文所示的这些实施例,而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

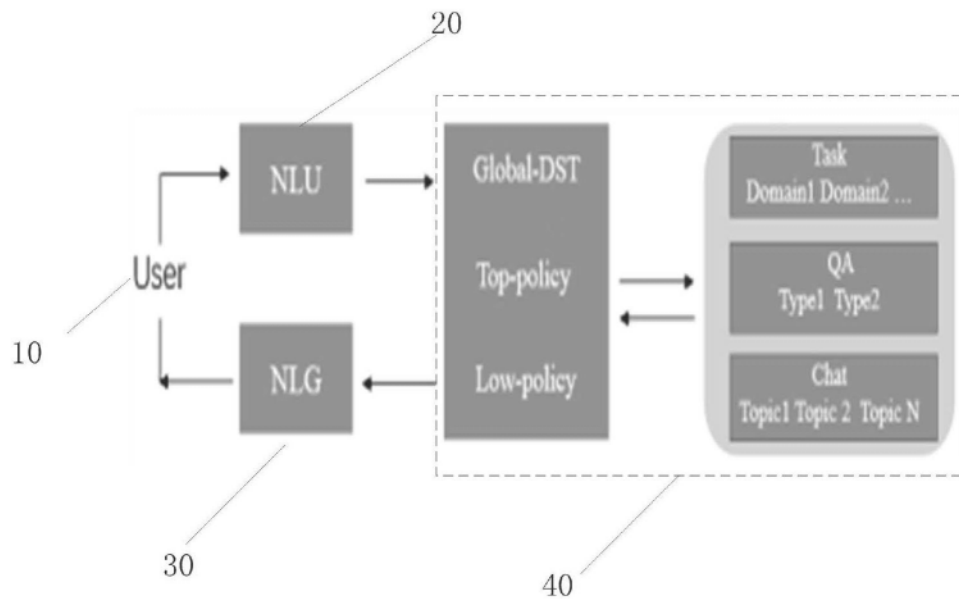


图1

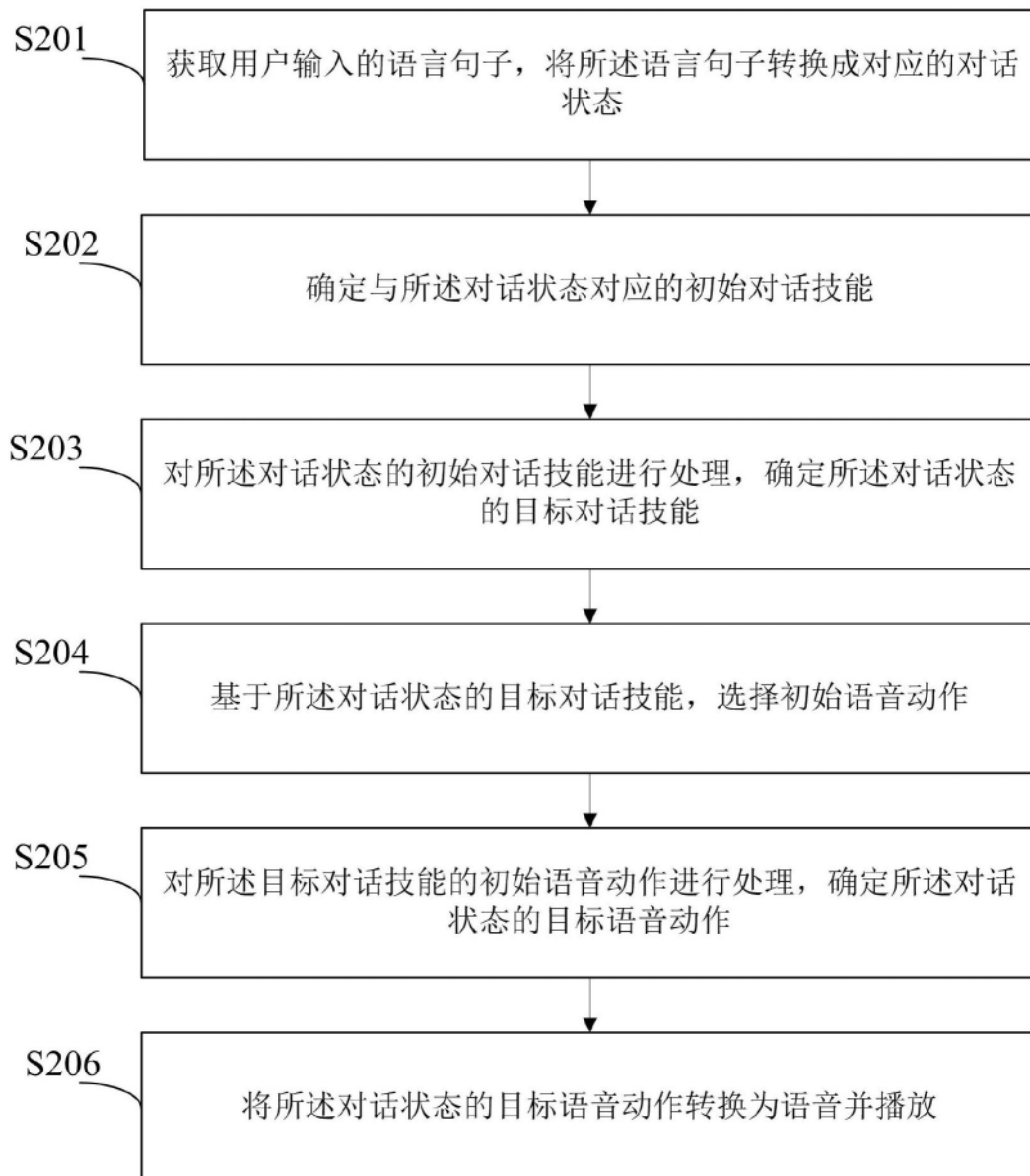


图2

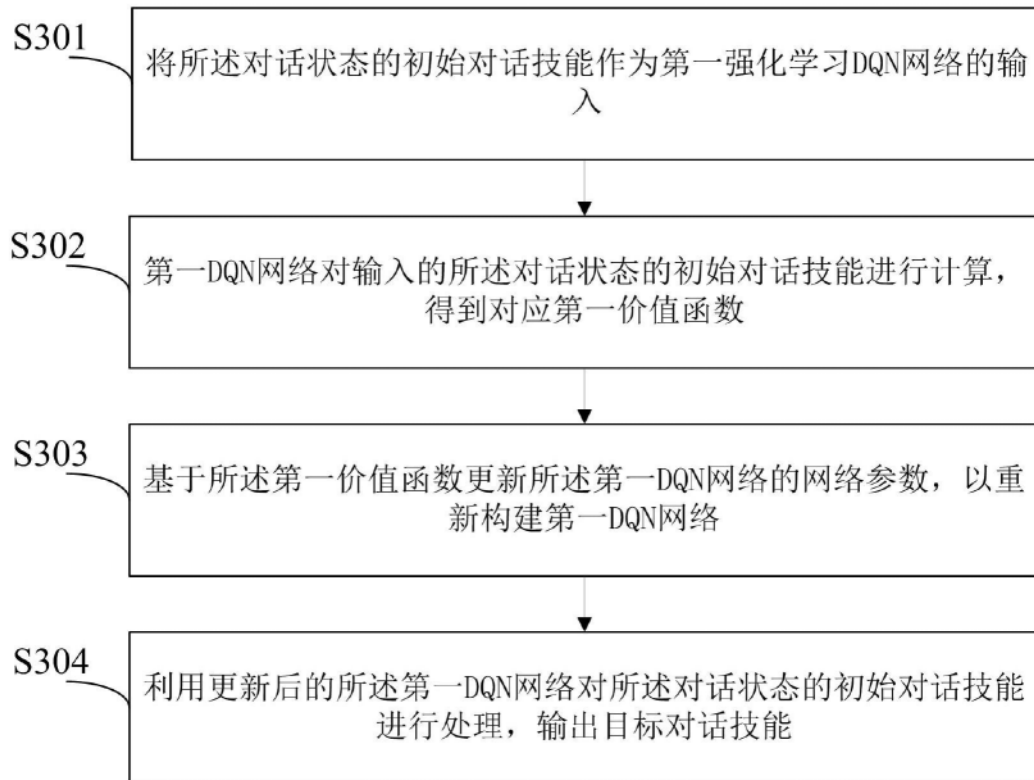


图3

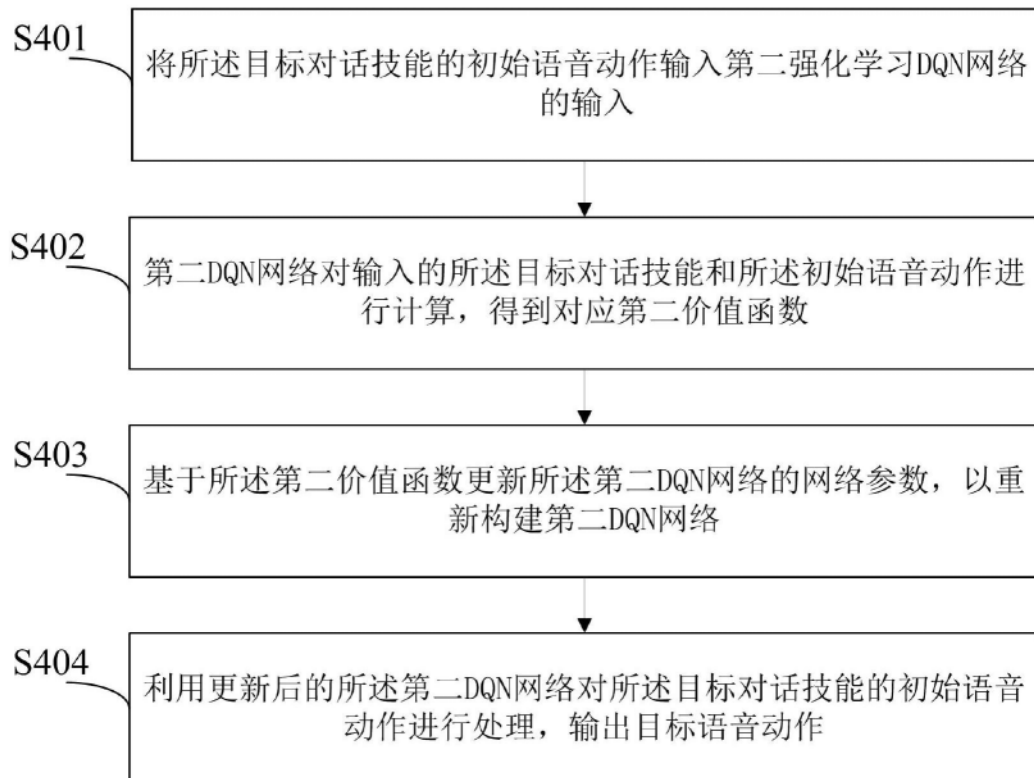


图4

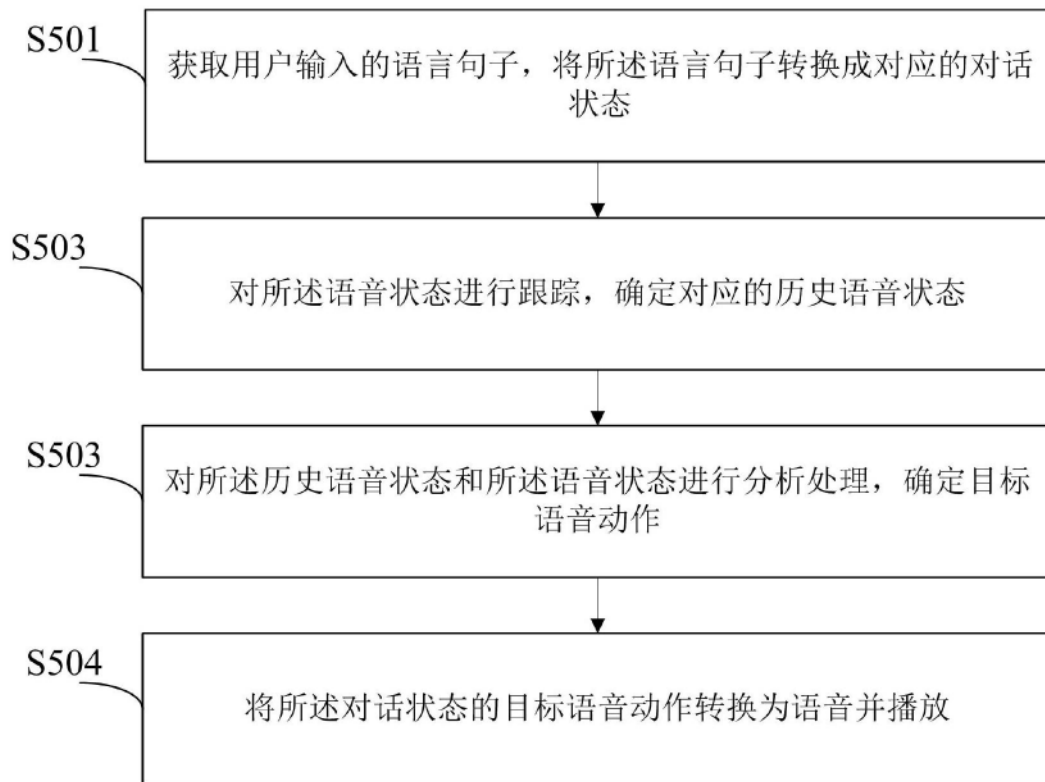


图5