



(12) 发明专利申请

(10) 申请公布号 CN 111737441 A

(43) 申请公布日 2020. 10. 02

(21) 申请号 202010786352.X

(22) 申请日 2020.08.07

(71) 申请人 北京百度网讯科技有限公司

地址 100085 北京市海淀区上地十街10号
百度大厦2层

(72) 发明人 吴华 王海峰 刘占一

(74) 专利代理机构 北京市汉坤律师事务所
11602

代理人 姜浩然 吴丽丽

(51) Int. Cl.

G06F 16/332 (2019.01)

G06F 40/30 (2020.01)

G06N 3/08 (2006.01)

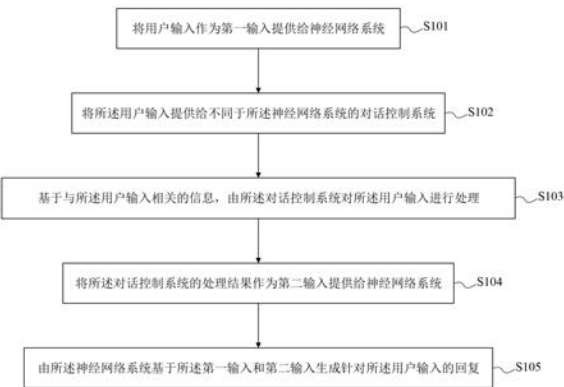
权利要求书5页 说明书24页 附图4页

(54) 发明名称

基于神经网络的人机交互方法、装置和介质

(57) 摘要

本公开提供一种基于神经网络的人机交互方法、装置和介质,涉及人工智能技术领域,具体为自然语言处理和知识图谱技术领域。方法包括:将用户输入作为第一输入提供给神经网络系统;将用户输入提供给不同于神经网络系统的对话控制系统;基于与所述用户输入相关的信息,由对话控制系统对用户输入进行处理;将对话控制系统的处理结果作为第二输入提供给神经网络系统;以及由神经网络系统基于所述第一输入和第二输入生成针对用户输入的回答。



1. 一种基于神经网络的人机交互方法,包括:
将用户输入作为第一输入提供给神经网络系统;
将所述用户输入提供给不同于所述神经网络系统的对话控制系统;
基于与所述用户输入相关的信息,由所述对话控制系统对所述用户输入进行处理;
将所述对话控制系统的处理结果作为第二输入提供给所述神经网络系统;以及
由所述神经网络系统基于所述第一输入和第二输入生成针对所述用户输入的回复。
2. 如权利要求1所述的方法,其中,所述与所述用户输入相关的信息包括长期记忆信息和仅在当前人机交互期间有效的工作记忆信息。
3. 如权利要求2所述的方法,其中,所述长期记忆信息包括形式为包含节点和有向边的第一有向图的知识信息,并且其中,所述第一有向图中的节点是包含语义内容和逻辑控制信息的结构化数据,所述第一有向图中的有向边表示相关节点之间的关联属性。
4. 如权利要求3所述的方法,其中,所述逻辑控制信息包括能够用于筛选与当前人机交互相关的节点的信息。
5. 如权利要求3所述的方法,其中,所述逻辑控制信息包括能够用于确定在当前人机交互中各节点之间的相关程度的信息。
6. 如权利要求3至5中任一项所述的方法,其中,所述第一有向图中的各节点包括第一类型节点和第二类型节点,所述第二类型节点的语义内容是与所述第二类型节点相关联的第一类型节点的语义内容的一部分,并且所述第二类型节点的逻辑控制信息包括下列中的至少一者:所述第二类型节点在与所述第二类型节点相关联的第一类型节点下的热度、所述第二类型节点与至少一个其他第二类型节点的关联跳转关系、所述第二类型节点的子类型。
7. 如权利要求6所述的方法,其中所述第一有向图中的各节点包括第三类型节点,所述第三类型节点的语义内容支持多模态内容,并且所述第三类型节点的逻辑控制信息包括下列中的至少一者:与所述第三类型节点相关联的第二类型节点的信息、用于表征所述第三类型节点的语义内容的信息。
8. 如权利要求3至5中任一项所述的方法,其中,所述长期记忆信息包括形式为包含节点和有向边的第二有向图的对话库信息,所述第二有向图和所述第一有向图同构。
9. 如权利要求3所述的方法,其中,所述工作记忆信息包括形式为包含节点和有向边的第三有向图的信息,其中,所述第三有向图和所述第一有向图同构并且所述第三有向图是所述第一有向图的一部分。
10. 如权利要求9所述的方法,其中,所述工作记忆信息包括取自于所述第一有向图的与所述当前人机交互相关的所有节点的语义内容和逻辑控制信息。
11. 如权利要求9或10所述的方法,所述工作记忆信息包括用于标记在所述当前人机交互中已经被涉及的语义内容的第一信息。
12. 如权利要求11所述的方法,所述工作记忆信息包括用于指明首先提及所述已经被涉及的语义内容的对话方的第二信息。
13. 如权利要求9或10所述的方法,所述工作记忆信息包括在所述当前人机交互期间的交互记录的历史数据。
14. 如权利要求2至5中任一项所述的方法,其中,所述处理结果包含在当前人机交互的

情境下对所述用户输入进行回复的规划。

15. 如权利要求14所述的方法,其中,所述由所述对话控制系统对所述用户输入进行处理,包括:

分析所述用户输入的语义内容;以及

分析所述用户在当前人机交互中对应于所述用户输入的交际意图。

16. 如权利要求15所述的方法,所述分析所述用户输入的语义内容包括:

判断所述用户输入是否能够对应于所述工作记忆信息中的某一节点;以及

响应于所述用户输入能够对应于所述工作记忆信息中的某一节点,基于所述工作记忆信息对所述用户输入进行处理。

17. 如权利要求16所述的方法,所述对所述用户输入进行处理包括:

基于所述工作记忆信息中的所述某一节点的信息,为所述用户输入补充相关的内容。

18. 如权利要求16或17所述的方法,响应于所述用户输入不能够对应于所述工作记忆信息中的节点,从所述长期记忆信息中提取与所述用户输入相关的节点的信息并存储到所述工作记忆信息。

19. 如权利要求15所述的方法,所述分析所述用户输入的语义内容包括:

对所述用户输入进行消歧处理。

20. 如权利要求19所述的方法,其中,所述对所述用户输入进行消歧处理包括:

基于所述用户输入和所述工作记忆信息中与当前人机交互相关的节点信息,识别所述用户输入中的具有歧义的至少部分内容并确定所述至少部分内容在当前人机交互情境中的含义。

21. 如权利要求18所述的方法,所述由所述对话控制系统对所述用户输入进行处理进一步包括:

根据所述用户输入的语义内容和在当前人机交互中对应于所述用户输入的交际意图,从所述工作记忆信息中查询所述用户输入的相关节点的信息;以及

根据与所述用户输入的关联程度,对查询得到的所述用户输入的相关节点进行排序,其中所述排序是基于所述相关节点的逻辑控制信息进行的。

22. 如权利要求21所述的方法,其中,根据与所述用户输入的关联程度,为所述相关节点赋予不同的分值。

23. 如权利要求21所述的方法,其中,在当前人机交互的情境下对所述用户输入进行回复的规划包括:

根据所述排序的结果,规划对话目标并选择与所述用户输入的关联程度最高的节点信息作为规划的对话内容;以及,

整合所述规划的对话内容和所述对话目标,作为所述第二输入提供给所述神经网络系统。

24. 如权利要求23所述的方法,其中,在没有针对所述用户输入更新所述工作记忆信息的情况下,响应于所述关联程度最高的节点不能满足预定的标准,重新查询所述长期记忆信息以对所述工作记忆进行更新。

25. 如权利要求1至5中任一项所述的方法,其中,所述神经网络系统是端到端神经网络系统。

26. 如权利要求25所述的方法,其中,所述端到端神经网络系统具有编码器和解码器,所述编码器用于接收所述用户输入以及已存储的当前人机交互的历史交互信息,所述解码器用于接收所述第二输入并且生成所述针对所述用户输入的回复。

27. 一种基于神经网络的人机交互装置,包括:

神经网络系统,被配置为接收用户输入作为第一输入;和

不同于所述神经网络系统的对话控制系统,被配置为接收所述用户输入,

其中,所述对话控制系统被进一步配置为基于与所述用户输入相关的信息对所述用户输入进行处理,并将处理结果作为第二输入提供给所述神经网络系统;以及

所述神经网络系统被进一步配置为基于所述第一输入和第二输入生成针对所述用户输入的回复。

28. 如权利要求27所述的装置,还包括:存储计算系统,所述存储计算系统包括工作记忆模块和长期记忆模块,所述与所述用户输入相关的信息包括取自所述长期记忆模块的长期记忆信息和取自所述工作记忆模块的仅在当前人机交互期间有效的工作记忆信息。

29. 如权利要求28所述的装置,其中,所述长期记忆模块中存储有意图知识图谱,所述意图知识图谱包括形式为包含节点和有向边的第一有向图的知识信息,并且其中,所述第一有向图中的节点是包含语义内容和逻辑控制信息的结构化数据,所述第一有向图中的有向边表示相关节点之间的关联属性。

30. 如权利要求29所述的装置,其中,所述逻辑控制信息包括能够用于筛选与当前人机交互相关的节点的信息。

31. 如权利要求29所述的装置,其中,所述逻辑控制信息包括能够用于确定在当前人机交互中各节点之间的相关程度的信息。

32. 如权利要求29至31中任一项所述的装置,其中,所述第一有向图中的各节点包括第一类型节点和第二类型节点,所述第二类型节点的语义内容是与所述第二类型节点相关联的第一类型节点的语义内容的一部分,并且所述第二类型节点的逻辑控制信息包括下列中的至少一者:所述第二类型节点在与所述第二类型节点相关联的第一类型节点下的热度、所述第二类型节点与至少一个其他第二类型节点的关联跳转关系、所述第二类型节点的子类型。

33. 如权利要求32所述的装置,其中所述第一有向图中的各节点包括第三类型节点,所述第三类型节点的语义内容支持多模态内容,并且所述第三类型节点的逻辑控制信息包括下列中的至少一者:与所述第三类型节点相关联的第二类型节点的信息、用于表征所述第三类型节点的语义内容的信息。

34. 如权利要求29至31中任一项所述的装置,其中,所述长期记忆模块中存储有对话库,所述对话库包括形式为包含节点和有向边的第二有向图的对话库信息,所述第二有向图和所述第一有向图同构。

35. 如权利要求29所述的装置,其中,所述工作记忆模块中存储有工作记忆信息,其中所述工作记忆信息包括形式为包含节点和有向边的第三有向图的信息,其中,所述第三有向图和所述第一有向图同构并且所述第三有向图是所述第一有向图的一部分。

36. 如权利要求35所述的装置,其中,所述工作记忆信息包括取自于所述第一有向图与所述当前人机交互相关的所有节点的语义内容和逻辑控制信息。

37. 如权利要求35或36所述的装置,所述工作记忆信息包括用于标记在所述当前人机交互中已经被涉及的语义内容的第一信息。

38. 如权利要求37所述的装置,所述工作记忆信息包括用于指明首先提及所述已经被涉及的语义内容的对话方的第二信息。

39. 如权利要求35或36所述的装置,所述工作记忆信息包括在所述当前人机交互期间的交互记录的历史数据。

40. 如权利要求28-31中任一项所述的装置,其中,所述处理结果包含在当前人机交互的情境下对所述用户输入进行回复的规划。

41. 如权利要求40所述的装置,其中,所述对话控制系统包括对话理解模块和对话控制模块,并且其中,所述对话理解模块被配置用于:

分析所述用户输入的语义内容;以及

分析所述用户在当前人机交互中对应于所述用户输入的交际意图。

42. 如权利要求41所述的装置,其中,所述对话理解模块包括:

判定子模块,被配置用于判断所述用户输入是否能够对应于所述工作记忆信息中的某一节点;以及

处理子模块,被配置用于响应于所述用户输入能够对应于所述工作记忆信息中的某一节点,基于所述工作记忆信息对所述用户输入进行处理。

43. 如权利要求42所述的装置,其中,所述处理子模块被进一步配置用于基于所述工作记忆信息中的所述某一节点的信息,为所述用户输入补充相关的内容。

44. 如权利要求42或43所述的装置,其中,所述对话理解模块被进一步配置用于响应于所述用户输入不能够对应于所述工作记忆信息中的节点,从所述长期记忆模块中提取与所述用户输入相关的节点信息并存储到所述工作记忆模块。

45. 如权利要求41所述的装置,其中,所述对话理解模块包括:

消歧子模块,被配置用于对所述用户输入进行消歧处理。

46. 如权利要求45所述的装置,其中,所述消歧子模块被进一步配置用于基于所述用户输入和所述工作记忆信息中与当前人机交互相关的节点信息,识别所述用户输入中的具有歧义的至少部分内容并确定所述至少部分内容在当前人机交互情境中的含义。

47. 如权利要求44所述的装置,其中,所述对话理解模块包括:

查询子模块,被配置用于根据所述用户输入的语义内容和在当前人机交互中对应于所述用户输入的交际意图,从所述工作记忆信息中查询所述用户输入的相关节点的信息;以及

排序子模块,被配置用于根据与所述用户输入的关联程度,对查询得到的所述用户输入的相关节点进行排序,其中所述排序是基于所述相关节点的逻辑控制信息进行的。

48. 如权利要求47所述的装置,其中,所述对话理解模块被进一步配置用于根据与所述用户输入的关联程度,为所述相关节点赋予不同的分值。

49. 如权利要求47所述的装置,其中,所述对话控制模块被配置用于执行以下操作以在当前人机交互的情境下对所述用户输入进行回复的规划:

根据所述排序的结果,规划对话目标并选择与所述用户输入的关联程度最高的节点信息作为规划的对话内容;以及,

整合所述规划的对话内容和所述对话目标,作为所述第二输入提供给所述神经网络系统。

50.如权利要求49所述的装置,其中,所述对话理解模块被进一步配置用于在没有针对所述用户输入更新所述工作记忆信息的情况下,响应于所述关联程度最高的节点仍不能满足预定的标准,重新查询所述长期记忆信息对所述工作记忆进行更新。

51.如权利要求27至31中任一项所述的装置,其中,所述神经网络系统是端到端神经网络系统。

52.如权利要求51所述的装置,其中,所述端到端神经网络系统具有编码器和解码器,所述编码器用于接收所述用户输入以及已存储的当前人机交互的历史交互信息,所述解码器用于接收所述第二输入并且生成所述针对所述用户输入的回复。

53.一种电子设备,包括:

处理器;以及

存储程序的存储器,所述程序包括指令,所述指令在由所述处理器执行时使所述处理器执行根据权利要求1-26中任一项所述的方法。

54.一种存储程序的计算机可读存储介质,所述程序包括指令,所述指令在由电子设备的处理器执行时,致使所述电子设备执行根据权利要求1-26中任一项所述的方法。

基于神经网络的人机交互方法、装置和介质

技术领域

[0001] 本公开涉及人工智能技术领域,具体为自然语言处理和知识图谱技术领域,特别涉及一种基于神经网络的人机交互方法、装置及介质。

背景技术

[0002] 开放域对话系统的目标是让机器像人一样,使用自然语言作为信息传递的媒介,机器通过回答问题、执行命令、聊天等来满足人们日常交互需求,聊天的主题和内容没有限制。

[0003] 目前,开放域对话系统主要采用基于语料库的检索方法或者生成方法。其中,语料库中存放的是在语言的实际使用中真实出现过的语言材料。在实际对话过程中用户输入存在意思相似但是字面表达差异很大、字面表达相似但是意思差异很大等问题,从而导致基于语料库的检索方法或者生成方法普遍存在对话内容空乏、逻辑不清晰、答非所问等问题。

[0004] 在此部分中描述的方法不一定是之前已经设想到或采用的方法。除非另有指明,否则不应假定此部分中描述的任何方法仅因其包括在此部分中就被认为是现有技术。类似地,除非另有指明,否则此部分中提及的问题不应认为在任何现有技术中已被公认。

发明内容

[0005] 根据本公开的一方面,提供一种基于神经网络的人机交互方法,包括:将用户输入作为第一输入提供给神经网络系统;将所述用户输入提供给不同于所述神经网络系统的对话控制系统;基于与所述用户输入相关的信息,由所述对话控制系统对所述用户输入进行处理;将所述对话控制系统的处理结果作为第二输入提供给所述神经网络系统;以及由所述神经网络系统基于所述第一输入和第二输入生成针对所述用户输入的回复。

[0006] 根据本公开的另一方面,提供一种基于神经网络的人机交互装置,包括:神经网络系统,被配置为接收用户输入作为第一输入;和不同于所述神经网络系统的对话控制系统,被配置为接收所述用户输入,其中,所述对话控制系统被进一步配置为基于所述与用户输入相关的信息对所述用户输入进行处理,并将处理结果作为第二输入提供给所述神经网络系统;以及所述神经网络系统被进一步配置为基于所述第一输入和第二输入生成针对所述用户输入的回复。

[0007] 根据本公开的另一方面,提供一种电子设备,包括:处理器;以及存储程序的存储器,所述程序包括指令,所述指令在由所述处理器执行时使所述处理器执行根据上述的方法。

[0008] 根据本公开的另一方面,提供一种存储程序的计算机可读存储介质,所述程序包括指令,所述指令在由电子设备的处理器执行时,致使所述电子设备执行根据上述的方法。

[0009] 根据本公开实施例的基于神经网络的人机交互方法,有助于改善人机交互过程中用户的聊天体验。

附图说明

[0010] 附图示例性地示出了实施例并且构成说明书的一部分,与说明书的文字描述一起用于讲解实施例的示例性实施方式。所示出的实施例仅出于例示的目的,并不限制权利要求的范围。在所有附图中,相同的附图标记指代类似但不一定相同的要素。

[0011] 图1是示出根据示例性实施例的基于神经网络的人机交互方法的流程图;

图2是示出根据示例性实施例的基于神经网络的人机交互装置的工作过程示意图;

图3和图4是示出根据示例性实施例的意图知识图谱的局部示意图;

图5是示出根据示例性实施例的对话控制系统的工作过程示意图;

图6是示出根据示例性实施例的对话理解模块的示意性组成框图;以及

图7是示出能够应用于示例性实施例的示例性计算设备的结构框图。

具体实施方式

[0012] 在本公开中,除非另有说明,否则使用术语“第一”、“第二”等来描述各种要素不意图限定这些要素的位置关系、时序关系或重要性关系,这种术语只是用于将一个元件与另一元件区分开。在一些示例中,第一要素和第二要素可以指向该要素的同一实例,而在某些情况下,基于上下文的描述,它们也可以指代不同实例。

[0013] 在本公开中对各种所述示例的描述中所使用的术语只是为了描述特定示例的目的,而并非旨在进行限制。除非上下文另外明确地表明,如果不特意限定要素的数量,则该要素可以是一个也可以是多个。此外,本公开中所使用的术语“和/或”涵盖所列出的项目中的任何一个以及全部可能的组合方式。

[0014] 开放域对话系统的聊天内容没有限制、主题任意,能够利用自然语言来进行回答问题、执行命令、聊天等。

[0015] 相关技术中,开放域对话系统获取用户的意图,根据意图将用户输入分发至多个交互子系统,接收多个交互子系统的返回结果,然后按照预设的排序策略来选择得分最高的结果并返回给用户。这种开放域对话系统存在以下问题:模块之间级联,容易发生错误传递;各个子系统相互独立,无法在子系统之间有效传递信息或者自然切换;不能有效地将知识融入基于深度学习的系统中,从而导致开放域对话系统存在对话内容空乏、逻辑不清晰、答非所问等问题。

[0016] 考虑到上述技术问题中的一个或多个,本公开提供一种基于神经网络的人机交互方法,该方法基于与用户输入相关的信息,由对话控制系统对用户输入进行处理。然后,将用户输入和对话控制系统的处理结果均作为输入提供给神经网络系统,由神经网络系统生成针对用户输入的回复,从而能够将与所述用户输入相关的信息融入到基于神经网络系统的对话系统中,充分利用相关信息,以解决相关技术中人机交互内容不理想的问题,使得人机交互的内容丰富、逻辑清晰。

[0017] 本公开的技术方案可以适用于所有使用对话系统的应用终端,例如,智能机器人、移动电话、计算机、个人数字助理、平板电脑等。

[0018] 以下将结合附图对基于神经网络的人机交互方法进行进一步描述。

[0019] 图1是示出根据本公开基于神经网络的人机交互方法的流程图。

[0020] 如图1所示,所述方法可以包括:步骤S101、将用户输入作为第一输入提供给神经

网络系统；步骤S102、将所述用户输入提供给不同于所述神经网络系统的对话控制系统；步骤S103、基于与所述用户输入相关的信息，由所述对话控制系统对所述用户输入进行处理；步骤S104、将所述对话控制系统的处理结果作为第二输入提供给所述神经网络系统；以及步骤S105、由所述神经网络系统基于所述第一输入和第二输入生成针对所述用户输入的回复。由此，通过将与所述用户输入相关的信息融入到基于神经网络系统的对话系统中，能够充分利用相关信息，使得人机交互的内容丰富、逻辑清晰。

[0021] 用户输入可以但不限于为文本信息或语音信息。可以对用户输入进行预处理然后再作为第一输入提供给神经网络系统和对话控制系统。所述预处理例如可以但不限于对语音信息进行语音识别，转换为相应的文本信息。

[0022] 参见图2所示，所述神经网络系统可以但不限于采用端到端神经网络系统101。所述端到端神经网络系统101可包括编码器1011和解码器1012。所述编码器1011可以对输入的文本内容进行隐式表示，生成向量；所述解码器1012可以根据给定的输入向量，生成流畅的自然语言文本。

[0023] 根据一些实施例，所述编码器1011可以用于接收所述用户输入以及已存储的当前人机交互的历史交互信息，并对用户输入以及已存储的当前人机交互的历史交互信息进行编码，生成隐式向量，所述隐式向量输入至编码器1012。所述解码器1012可以用于接收所述第二输入（即对话控制系统对用户输入进行处理得到的处理结果）以及编码器1011生成的隐式向量，并且生成所述针对所述用户输入的回复。由此，神经网络系统能够基于当前用户输入、已存储的当前人机交互的历史交互信息、以及对话控制系统基于与所述用户输入相关的信息对所述用户输入进行处理所得到的结果，来生成针对用户输入的回复，从而能够进一步保证机器回复内容符合当前人机交互的场景，对话逻辑清晰。

[0024] 所述端到端神经网络系统例如可以采用Transformer神经网络系统或UniLM神经网络系统。

[0025] 根据一些实施例，所述与所述用户输入相关的信息可以包括长期记忆信息和仅在当前人机交互期间有效的工作记忆信息。作为一个示例实施例，与所述用户输入相关的信息可以为预先存储的信息。在这种情况下，长期记忆信息可以是对话系统需要长时间存储的信息，可以包括各种知识信息，例如可包括下列中的至少其中之一：常识、领域知识、语言知识、问答库以及对话库。工作记忆信息可以为基于当前人机交互内容从长期记忆信息中获得。也就是说，工作记忆信息是与当前人机交互内容相关的知识信息。从而能够将当前人机交互内容相关的知识信息融入到基于神经网络系统的对话系统中，实现基于相关知识信息在当前人机交互的情境下对用户输入进行回复规划，充分利用知识信息，使得人机交互的内容丰富、逻辑清晰。可以理解的是，与所述用户输入相关的信息也可以包括从网络上实时抓取的信息，在此不作限定。

[0026] 根据一些实施例，长期记忆信息可以但不限于包括意图知识图谱、问答库和对话库。下面将首先描述意图知识图谱、问答知识和对话库的数据内容、数据组织形式等。

[0027] 意图知识图谱能够从对话场景的知识交互需求出发，不仅满足知识查询功能，更多的是满足多轮多场景交互中的联想、类比、预测等。意图知识图谱的节点的有序组织，便于文本计算，控制知识信息，并且通过知识信息计算，能够支持对话中的行为跳转（场景跳转、同一场景下内容跳转），具有很强的语义转移逻辑性。意图知识图谱整合了不同类型的

多场景信息,能够提供多视角理解语言的能力。

[0028] 根据一些实施例,意图知识图谱可以包括形式为包含节点和有向边的第一有向图的知识信息,并且其中,所述第一有向图中的节点是包含语义内容和逻辑控制信息的结构化数据。所述第一有向图中的有向边可表示相关节点之间的关联属性以及节点与相应的逻辑控制信息之间的关联属性。可以理解的是,其它知识信息也可以采用第一有向图的数据组织形式,并不局限于意图知识图谱,在此仅是以意图知识图谱为例来说明如何利用第一有向图来表示知识信息。

[0029] 根据一些实施例,意图知识图谱的所述逻辑控制信息可包括能够用于筛选与当前人机交互相关的节点的信息,例如,热度、时效性、情感等信息,用于筛选与当前人机交互内容相关的节点,从而能够实现在用户主动发起知识聊天的情况下去检索相关的知识信息,使得对话内容逻辑清晰。例如,可以设置第一热度阈值,从工作记忆信息中与当前人机交互内相关的节点中筛选出相应的逻辑控制信息中的热度大于所述第一热度阈值的节点。可以设置第一有效时间点,从工作记忆信息中与当前人机交互内容相关的节点中筛选出相应的逻辑控制信息中的时效信息在所述第一有效时间点之后的节点。可以设置第一预设情感类型,从与当前人机交互内容相关的节点中筛选出相应的逻辑控制信息中的情感类型为所述第一预设情感类型的节点。

[0030] 根据一些实施例,意图知识图谱的所述逻辑控制信息还可包括能够用于确定在当前人机交互中各节点之间的相关程度的信息,例如,热度、节点间的关联关系等信息,用于扩展与当前人机交互内容相关的节点,从而能够实现机器主动切换、触发或推荐知识聊天,使得对话内容丰富,避免尬聊。例如,可以设置第二热度阈值,从长期记忆信息中用户输入的各相关节点中获取相应的逻辑控制信息中的热度大于所述第二热度阈值的节点。可以根据关联关系,从当前的节点扩展到与该当前节点关联程度最高的节点。

[0031] 意图知识图谱的所述节点可以包括多个不同类型的节点。根据一些实施例,所述第一有向图中的各节点可包括第一类型节点和第二类型节点。所述第二类型节点的语义内容可以是与所述第二类型节点相关联的第一类型节点的语义内容的一部分,并且所述第二类型节点的逻辑控制信息包括下列中的至少一者:所述第二类型节点在与所述第二类型节点相关联的第一类型节点下的热度、所述第二类型节点与至少一个其他第二类型节点的关联跳转关系、所述第二类型节点的子类型。由此,通过查询第一类型节点即可获得与第一类型节点语义相关联的第二类型节点的知识信息,便于文本计算,控制知识信息。

[0032] 所述第一类型节点例如可以为下表中的核心节点,所述第二类型节点例如可以为下表中的标签节点。所述有向边可以表示核心节点之间、核心节点与标签节点之间的关联属性,以及各节点与相应的逻辑控制信息之间的关联属性。核心节点和标签节点可为结构化数据,从而能够实现语义内容的理解和控制。核心节点可是语义完整的基本单元,可以包含实体、概念、事件、指令,例如可以为人物、物品、结构、产品、建筑、地点、组织、事件、艺术作品、科学技术、科学定理等。核心节点的逻辑控制信息可以包括热度、时效性、用于召回标签节点的所有标签、任务API等。每一个核心节点可以包括多个相关联的标签节点。标签节点的语义内容可以是与标签节点相关联的核心节点语义内容的一部分,与核心节点是部分与整体的关系。

[0033] 意图知识图谱的节点表示:

核心节点	定义：语义完整的基本模块，包含实体、概念、事件、指令
	逻辑控制信息：热度、时效性、用于召回节点的所有标签、任务API 等
标签节点	定义：核心节点的语义内容的一部分，与核心节点是部分与整体的关系，是内容节点的主题或概括
	逻辑控制信息：在核心节点下的热度、标签节点间的关联跳转关系、标签节点的类型等
内容节点	定义：对话内容，具有多模态（词、句子、图片、视频等）、多样性、细粒度等特点
	逻辑控制信息：核心标签、关键词、核心标签在对话内容中的重要程度、对话内容的概括性短语、类别、情感极性、得分等

根据一些实施例，与当前人机交互有关的信息可以包括从第一有向图中获取的与用户输入相关的节点信息。可以将用户输入映射到所述第一有向图的核心节点上，映射得到的核心节点以及与映射得到的核心节点相关的标签节点可以均作为与用户输入相关的知识信息。若用户输入无法映射到所述第一有向图的核心节点上，可以将当前人机交互内容的历史用户输入映射得到的核心节点作为与当前用户输入对应的核心节点。例如，若当前用户输入为“主角是谁？”，则当前用户输入在所述第一有向图中无对应的核心节点。这种情况下，可以将当前人机交互中上一次在所述第一有向图中对应的核心节点作为当前用户输入的核心节点，以获取与当前用户输入相关的知识信息。其中，当前人机交互内容可包括当前用户输入以及当前人机交互的历史交互信息。

[0034] 如图3所示，实线圆形（“电影A”、“电影B”、“赵六”）示意的是核心节点，实线椭圆形示意的是标签节点，虚线圆形示意的是逻辑控制信息。每一个虚线椭圆可围住一个节点单元，作为与用户输入相关的信息单元。实线线段表示各节点之间的有向边，虚线线段表示节点和相应的逻辑控制信息之间的有向边。也就是说，当用户输入映射到一个节点单元（如图3中的节点单元100）的核心节点上时，节点单元的所有节点信息均认为与用户输入相关的知识信息，并将其添加到工作记忆信息中。需要说明的是，根据系统可用计算资源的大小，也可以将与映射得到的一个核心节点相关联的至少一个其它核心节点所在的节点单元也认为与用户输入相关，也添加到工作记忆信息中，在此不作限定。以下内容中将以映射得到的核心节点所在的节点单元作为与用户输入相关的知识信息为例，来具体描述本公开的技术方案。

[0035] 以第一类型节点（核心节点）为电影实体“电影A”为例，“电影A”用户召回标签节点的标签可以包括演员、角色、导演、场景等。与第一类型节点相关联的标签节点（第二类型节点）可以包括“赵六”（假设为演员）、“角色A”、“角色B”、“李四”（假设为电视导演）、“知名场景”。其中，标签节点“赵六”对应相关联的第一类型节点“电影A”的演员标签，“角色A”和“角

色B”对应相关联的第一类型节点“电影A”的角色标签,“李四”对应相关联的第一类型节点“电影A”的导演标签,“知名场景”对应相关联的第一类型节点“电影A”的场景标签。与核心节点“电影A”相关联的核心节点可包括“电影B”,与标签节点“赵六”相关联的核心节点可包括“赵六”。当用户输入映射到核心节点“电影A”,则可以将核心节点“电影A”以及与标签节点“赵六”、“角色A”、“角色B”、“李四”、“知名场景”均作为与用户输入相关的信息。

[0036] 根据一些实施例,所述第一有向图中的各节点还可包括第三类型节点,所述第三类型节点的语义内容支持多模态内容,并且所述第三类型节点的逻辑控制信息可包括下列中的至少一者:与所述第三类型节点相关联的第二类型节点的信息、用于表征所述第三类型节点的语义内容的信息。由此,通过设置第三类型节点,能够支持多模态的语义内容,更加丰富对话内容。

[0037] 所述第三类型节点例如可以为上表中的内容节点。所述有向边还可以表示标签节点(第二类型节点)和内容节点之间的关联属性。内容节点可为非结构化数据,能够支持丰富的多模态内容。每一个核心节点(第一类型节点)可包括多个内容节点,与内容节点相关联的标签节点可以是该内容节点的主题或概括。内容节点可包含对话内容,具有多模态(可以为词、句子、图片或视频等)、多样性、细粒度等特点。内容节点的逻辑控制信息例如可包括核心标签、关键词、核心标签在内容节点的语义内容中的重要程度、内容节点的语义内容的概括性短语、与内容节点相关联的标签节点的类别、与内容节点相关联的标签节点的情感极性、与内容节点相关联的标签节点的得分等。

[0038] 根据一些实施例,与当前人机交互内容有关的信息可以包括从第一有向图中获取的与用户输入相关的节点信息。可以将用户输入映射到所述第一有向图的核心节点上,映射得到的核心节点、与映射得到的核心节点相关联的标签节点、以及与所得到的标签节点相关联的内容标签可以均作为与用户输入相关的信息。

[0039] 如图3所示,矩形框示意的是内容节点。以第一类型节点(核心节点)为电影实体“电影A”为例,与第一类型节点相关联的标签节点(第二类型节点)可以包括“赵六”、“角色A”、“角色B”、“李四”、“知名场景”。其中,与标签节点“赵六”和标签节点“角色A”相关联的内容节点可包括“角色A剧照.jpg”(假设为电影A中角色A的知名剧照),与标签节点“角色A”相关联的内容节点还可包括“角色A具有极其完全的生命力和生命意志”,与标签节点“李四”相关联的内容节点可包括:《电影A》是李四导演的武侠巅峰之作。当用户输入映射到核心节点“电影A”,则可以将核心节点“电影A”、与标签节点“赵六”、“角色A”、“角色B”、“李四”、“知名场景”、以及与标签节点“赵六”、“角色A”、“李四”相关联的内容节点均作为与用户输入相关的信息。

[0040] 两个节点相关联可以是指这两个节点之间可通过包括至少一条有向边的有向路径关联起来。不同节点之间可以通过有向边连接,并示意被连接的节点之间的关联属性。有向边例如可以包括核心节点到核心节点之间的关联边、核心节点到标签节点的关联边、标签节点到核心节点的关联边、标签节点到内容节点的关联边。有向边的属性可以包括语义关系(如:导演、作品、妻子等)、逻辑关系(时序、因果等)、相关性强度、语义上下位关系等多种类型。

[0041] 例如,如图3中,核心节点“电影A”和核心节点“电影B”之间的有向边的属性可为相关性强度,标签节点“赵六”和核心节点“赵六”之间的有向边的属性可为相关性强度。核心

节点“电影A”与标签节点“李四”、“赵六”、“角色A”、“知名场景”之间的有向边的属性为语义关系。标签节点“赵六”与内容节点“角色A剧照.jpg”之间的有向边的属性可为语义关系。

[0042] 根据一些实施例,对话库可以包括包含节点和有向边的第二有向图的知识信息,用于记录人机交互过程中的语义信息和其特征,为在当前人机交互的情境下对所述用户输入进行回复的规划提供参考。由于基于对话库能够实现利用大数据来获取用户更加倾向的意图,因此,能够为针对用户输入的回复规划提供合理引导。所述第二有向图可以和上述的第一有向图(例如,意图知识图谱)同构,参见图3所示,在此不再详述。由此,通过设置对话库和意图知识图谱为同构的有向图,能够有效实现对话库和意图知识图谱进行融合,便于控制知识信息。可以理解的是,其它知识信息也可以采用第二有向图的数据组织形式,并不局限于对话库,在此仅是以对话库例来说明如何利用第二有向图来表示知识信息。通过设置不同知识信息采用同构的有向图,能够有效实现不同知识信息的融合,便于控制知识信息。

[0043] 根据一些实施例,问答库可以是形式为问题-答案的问答知识信息。问答库的功能为:针对用户的问题,查询问答库,返回与问题相匹配的答案,满足用户的信息需求。例如,可以在用户输入为问答的情况下,优先从问答库中查询是否有与用户输入匹配的答案,能够快速实现回复。

[0044] 问答库的形式可如下表所示:

Key (问题)	Value (答案)
礼泉一中占地面积	150 亩
张仲景出生时间	公元 150 年
马到成功拼音	mǎ dào chéng gōng
青华镇面积	91. 59 平方公里
李世民是哪个朝代的皇帝	唐朝
联合国成立时间	1945 年
stm 发明时间	1981 年
淑均拼音	shū jūn
维您是哪里的品牌	美国洛杉矶
月光斩作者	莫言
唐明皇是谁	李隆基
掌声雷动拼音	zhǎng shēng léi dòng

根据一些实施例,所述长期记忆信息可以包括意图知识图谱、对话库和问答库。上述通过示例描述了长期记忆信息的意图知识图谱、对话库和问答库的数据内容和数据组织形式,仅是举例说明,并不作为一种限定。当然,所述长期记忆信息也可以为其它与当前人机交互有关的知识信息组合,在此不作限定。

[0045] 长期记忆信息还可进行语言计算与信息抽取。语言计算可以包括比较、归纳、演绎、推理等；信息抽取例如可以包括概念抽取、实体抽取、事件抽取、指令抽取等，从而能够基于用户输入从长期记忆信息中获取与当前人机交互内容相关的工作记忆信息。当前人机交互内容可包括当前用户输入以及当前用户输入之前的历史交互信息。工作记忆信息还可以包括当前人机交互内容，从而能够基于当前人机交互历史和从长期记忆信息中获取的与用户输入相关的知识信息，获取在当前人机交互的情境下对所述用户输入进行回复的规划，将在以下内容中详细介绍。

[0046] 根据一些实施例，所述工作记忆信息可包括形式为包含节点和有向边的第三有向图的信息，所述第三有向图可以和上述的第一有向图（例如，意图知识图谱）同构。由此，通过设置工作记忆信息包括与长期记忆信息的知识信息同构的信息，能够方便知识信息的调用和融合。优选地，所述第三有向图可以为所述第一有向图的与当前人机交互相关的一部分，更加方便知识信息的调用和融合。也就是说，所述第三有向图可包括核心节点和标签节点，从而所有用户意图和系统回复（意图）也都能够尽可能地映射到工作记忆信息中的核心节点和相关联的标签节点上，便于各模块使用。另外，由于只需要从长期记忆信息中提取与当前人机交互相关的部分节点信息，可以减少占用内存和提高回复效率。所述第三有向图还可包括内容节点，支持多模态的语义内容，从而能够基于工作记忆信息也获得丰富的对话内容。可以理解的是，所述第三有向图也可以和上述的第一有向图不同构。

[0047] 根据一些实施例，所述工作记忆信息还可包括取自于所述第一有向图的与所述当前人机交互相关的所有节点的语义内容和逻辑控制信息。也就是说，所述第三有向图的核心节点包括对应于第一有向图的所述第一类型节点的语义内容和逻辑控制信息，所述标签节点包括对应于第一有向图的所述第二类型节点的语义内容和逻辑控制信息，所述内容节点包括对应于第一有向图的所述第三类型节点的语义内容和逻辑控制信息。由此，工作记忆信息能够基于当前人机交互从长期记忆信息中尽可能地获取所有可聊的话题，以便于基于工作记忆信息即能够实现针对用户输入进行回复规划。由于工作记忆信息中的数据量远小于长期记忆信息中的数据量，从而能够提高回复速度，提升用户体验。

[0048] 根据一些实施例，当工作记忆信息中没有与用户输入对应的节点信息时，可以基于用户输入从长期记忆信息中获取与用户输入相关的知识信息，对工作记忆信息进行更新；然后基于更新后的工作记忆信息，在当前人机交互的情境下对所述用户输入进行回复的规划。根据一些实施例，可以基于用户输入从第一有向图中获取与用户输入相关的子图，将获取的子图融合到工作记忆信息中的第三有向图中，对工作记忆信息进行更新。

[0049] 对于工作记忆信息，与当前人机交互内容中的历史交互信息对应的各节点中，可以仅保留核心节点的语义内容和逻辑控制信息，而不保留与核心节点相关联的标签节点和内容节点，从而能够减少计算资源需求。因为已聊过的话题很可能不会再次涉及，因此，仅保留当前人机交互中的历史交互信息对应的核心节点的语义内容和逻辑控制信息，对人机交互的影响很小。

[0050] 根据一些实施例，所述工作记忆信息还可包括用于标记在所述当前人机交互中已经被涉及的语义内容的第一信息，从而能够区分聊过和没聊过的消息，避免重复。根据一些实施例，在第三有向图中，对于在所述当前人机交互中已经被涉及的语义内容相关的所有节点（包括目标节点、标签节点和内容节点）还可包括所述第一信息，以指示该节点已被聊

过。

[0051] 根据一些实施例,所述工作记忆信息还可包括用于指明首先提及所述已经被涉及的语义内容的对话方的第二信息,从而能够准确区分聊过相关内容的话题,更精确的避免对话方的对话重复。根据一些实施例,在第三有向图中,对于在所述当前人机交互中已经被涉及的语义内容相关的所有节点(包括目标节点、标签节点和内容节点)还可包括所述第二信息,以指示该节点被哪个对话方聊过。

[0052] 根据一些实施例,所述工作记忆信息还可包括在所述当前人机交互期间的交互记录的历史数据,从而能够获取当前人机交互的场景,以便于为多轮策略提供决策特征。

[0053] 所述工作记忆信息还可包括其它信息,例如,所述对话控制系统的各工作模块的分析结果,方便各模块的使用。例如,除了上述的基于对用户输入从长期记忆信息中获取的与用户输入相关的知识信息,还可包括从长期记忆信息中获取的与用户输入相关的知识信息进行排序的结果、回复决策结果。

[0054] 根据一些实施例,步骤S103、基于与用户输入相关的信息,对话控制系统对所述用户输入进行处理,所得到的处理结果可以包含在当前人机交互的情境下对所述用户输入进行回复的规划。从而能够充分利用相关信息,并且基于相关信息在当前人机交互的情境下对用户输入进行回复规划,进一步使得人机交互的内容丰富、逻辑清晰。

[0055] 根据一些实施例,对话控制系统可以包括对话理解模块和对话控制模块。在一个实施例中,可以利用对话理解模块首先基于用户输入从长期记忆信息中获取相关的知识信息,对工作记忆信息进行更新;然后对话控制模块基于更新后的工作记忆信息,在当前人机交互的情境下对所述用户输入进行回复的规划。

[0056] 基于此,步骤S103、基于与所述用户输入相关的信息,由所述对话控制系统对所述用户输入进行处理可以包括:分析所述用户输入的语义内容;以及分析所述用户在当前人机交互中对应于所述用户输入的交际意图。也就是说,用户输入的理解结果可以包括语义内容和交际意图。交际意图例如可以为意图体系中的其中之一,例如,提问、澄清、建议、拒绝、鼓励或安慰等等。

[0057] 作为一个示例性实施例,意图体系的列表可为如下:

1	主动告知	22	安慰
2	提问	23	澄清问
3	具体回答	24	祝福
4	指令	25	道别
5	建议	26	接受承诺
6	肯定回答	27	道歉
7	同意	28	疑问式回答
8	承诺	29	部分接受行动请求
9	确认收到新型	30	拒绝承诺
10	否定回答	31	自谦
11	不同意	32	自我介绍
12	表情符	33	回应致谢
13	感叹	34	恭喜
14	恭维	35	抛弃
15	接受行动请求	36	自我纠错
16	拒绝行动请求	37	回答并提问
17	致谢	38	回应道歉
18	纠正	39	部分接受承诺
19	问候	40	撤回
20	避不回答	41	稍等
21	鼓励		

根据一些实施例,可以基于意图知识图谱来对所述用户输入进行理解。例如,结合图4所示,接收到的第一用户输入为:你知道电影C的主演是谁吗?第一用户输入的理解结果的语义内容为电影C,交际意图为问答。接收到的第二用户输入为:我很喜欢张三(假设为演员)。第二用户输入的理解结果的语义内容为张三,交际意图为闲聊。

[0058] 用户输入的理解结果还可包括描述用户状态的状态意图,例如,用户的心情状态、用户是否喜欢当前聊天。从而能够结合用户的状态意图来进行对话决策,规划回复内容。

[0059] 根据一些实施例,可以基于训练完成的意图神经网络模型来理解用户输入的交际意图。可获取第一用户输入样本集,采用人工标注第一用户输入样本集中的常用用户输入样本的交际意图。利用所述第一用户输入样本集对意图神经网络模型进行训练。例如可以基于日志数据(例如,搜索引擎日志)来获取所述第一用户输入样本集。还可获取低频用户输入(例如“我不知道你说的啥”)并人工标记低频用户输入的交际意图,生成语料库。对于意图神经网络模型无法识别交际意图的用户输入,即意图体系无对应的交际意图,则可以在所述语料库中查找与用户输入语义相似度最高的低频用户输入,并将查找得到的低频用户输入相应的交际意图作为用户输入的交际意图,从而能够保证理解用户输入的交际意图。

[0060] 以下将以意图知识图谱来详细描述对所述用户输入进行理解的过程。

[0061] 根据一些实施例,步骤S103中的所述分析所述用户输入的语义内容可以包括:判断所述用户输入是否能够对应于所述工作记忆信息中的某一节点;响应于所述用户输入能够对应于所述工作记忆信息中的某一节点,基于所述工作记忆信息对所述用户输入进行处理,从而能够基于工作记忆信息来理解用户输入的语义内容,实现在当前人机交互的情景下对用户输入的理解,提高对话理解的准确性和效率。所述某一节点例如可以为第三有向

图的节点,如上面内容所描述的,所述第三有向图可以与所述第一有向图(意图知识图谱)同构并且为所述第一有向图的一部分。

[0062] 根据一些实施例,所述对所述用户输入进行处理可以包括:基于所述工作记忆信息中的所述某一节点的信息,为所述用户输入补充相关的内容。例如,用户输入为“主角是谁”,可以基于从工作记忆信息中找到的对应于所述用户输入的某一核心节点“电影A”,据此可将用户输入补全为“电影A的主角是谁”。根据一些实施例,可以查找当前人机交互内容对应于工作记忆信息中的上一核心节点,并进行以下判定:用户输入是否被该上一核心节点的逻辑控制信息中的标签覆盖,若被覆盖,则根据该上一节点为所述用户输入补充相关的内容。例如,该上一核心节点“电影A”的逻辑控制信息中的标签包括:演员、角色、导演、场景。由于“主角”和“演员”的语义相同,确定用户输入被核心节点“电影A”的标签覆盖,根据核心节点“电影A”,将用户输入补全为“电影A的主角是谁”。

[0063] 可以基于补全后的用户输入来进一步分析所述用户输入的语义内容,以提高对话理解的准确性。

[0064] 根据一些实施例,可响应于所述用户输入不能够对应于所述工作记忆信息中的节点,从所述长期记忆信息中提取与所述用户输入相关的节点信息存储到所述工作记忆信息。从而能够在用户输入未被工作记忆信息中的知识信息覆盖的情况下,扩大知识范围(例如,基于整个意图图谱),尝试基于知识信息来理解用户输入。

[0065] 根据一些实施例,步骤S103中的所述分析所述用户输入的语义内容还可包括:对所述用户输入进行消歧处理。例如,用户输入“我爱读水浒传”,其中经分词处理后得到的“水浒传”存在歧义,其可能是电视剧,也可能是小说。因此,需要对用户输入进行消歧处理,确定“水浒传”的类型,以能够准确用户输入的语义内容。

[0066] 可以基于消歧处理的结果来进一步分析所述用户输入的语义内容,以提高对话理解的准确性。

[0067] 根据一些实施例,对所述用户输入进行消歧处理可包括:基于所述用户输入和所述工作记忆信息中与当前人机交互相关的节点信息,识别所述用户输入中的具有歧义的至少部分内容,并确定所述至少部分内容在当前人机交互情境中的含义,从而能够基于当前人机交互的情景来对用户输入进行消歧处理。例如,用户输入“我爱读水浒传”,由于“水浒传”既可以指小说、也可以指电视剧,因而具有歧义,在此情形下,系统通过用户输入中的“读”,确定出“水浒传”在当前语境中的真实含义应当是指小说,而不是电视剧。作为一个示例性实施例,可以基于用户输入和工作记忆信息中与当前人机交互对应的上一核心节点(可为工作记忆信息中最新更新的核心节点,包括语义内容和逻辑控制信息),来对所述用户输入进行消歧处理。例如,可将用户输入和工作记忆信息中与当前人机交互内容对应的上一核心节点输入消歧神经网络模型,以获取消歧神经网络模型输出的用户输入中的具有歧义的至少部分内容的类型。可以利用类型语料库来对消歧神经网络模型进行度量训练,以能够实现用户输入和所述工作记忆信息中与当前人机交互相关的节点的信息的组合,与类型语料库中相应的类型距离更近,以能够输出用户输入中的具有歧义的至少部分内容以及具有歧义的所述至少部分内容的类型。在此仅是举例说明如何确定用户输入中的具有歧义的至少部分内容以及具有歧义的所述至少部分内容的类型,并不是一种限定。

[0068] 可以理解的是,在此并不限定仅能根据工作记忆信息来进行消歧处理,例如,还可

基于长期工作信息中的对话库来对用户输入进行消歧处理。例如,在对话库中,对于输入“我爱水浒传”更加倾向于阅读的意图,则可确定“水浒传”的类型为小说。

[0069] 根据一些实施例,步骤S103中的所述分析所述用户输入的语义内容可包括:消歧处理和信息补全处理。可以基于消歧处理的结果和补全后的用户输入来进一步分析所述用户输入的语义内容,以提高对话理解的准确性。

[0070] 可以基于交际意图来决定后续的操作。例如,交际意图为查询,可以基于消歧处理的结果和补全后的用户输入和交际意图生成意图查询表达式,以查找相关的知识信息。交际意图为根据意图表达式道别,则不用查找相关知识信息。在需要查找相关知识信息的情况下,可以首先在工作记忆信息中查找是否有与用户输入相关的知识信息,若无,则在长期记忆信息中继续查找是否有与用户输入相关的知识信息。

[0071] 根据一些实施例,步骤S103进一步可包括:根据所述用户输入的语义内容和在当前人机交互中对应于所述用户输入的交际意图,从工作记忆信息中查询所述用户输入的相关节点的信息;根据与所述用户输入的关联程度,对查询得到的所述用户输入的相关节点进行排序,其中所述排序是基于所述相关节点的逻辑控制信息进行的。例如,可以基于热度或时效性等进行打分,来确定相关节点与所述用户输入的关联程度,从而能够根据相关节点与所述用户输入的关联程度来进行对话决策,实现对话系统生成的回复与用户输入的关联性。

[0072] 根据一些实施例,根据与所述用户输入的关联程度,为所述相关节点赋予不同的分值,从而能够为对话决策提供参考。例如,可以在工作记忆信息中的第三有向图的核心节点的逻辑控制信息中增加与所述用户输入的关联分数。

[0073] 分析得到的用户输入的语义内容例如可以为第三有向图中的与用户输入相关的核心节点。

[0074] 根据一些实施例,在根据与所述用户输入的关联程度,为所述相关节点赋予不同的分值的情况下,在当前人机交互的情境下对所述用户输入进行回复的规划可包括:根据所述排序的结果,规划对话目标并选择与所述用户输入的关联程度最高的节点信息作为规划的对话内容;以及,整合所述规划的对话内容和所述对话目标,作为所述第二输入提供给所述神经网络系统,从而对话系统能够融入知识信息,根据用户输入进行答复规划,使得对话逻辑清晰。

[0075] 可以理解的是,可以基于意图知识图谱中节点的逻辑控制信息来获取所述相关节点与用户输入的关联程度,也可以基于对话库获取所述相关节点与用户输入的关联程度,还可以基于用户喜好来获取所述相关节点与用户输入的关联程度,在此不作限定,只要能够从知识信息中获取所述相关节点与用户输入的关联程度即可。可以基于用户的当前人机交互内容和历史人机交互内容来获取该用户的喜好,例如,该用户在多次人机交互中均涉及阅读,可以确定该用户喜好阅读,在对话决策的过程中可以根据用户的喜好来规划对话内容。

[0076] 根据一些实施例,在工作记忆信息中没有对应用户输入的知识信息的情况下,可以基于用户输入的语义内容从长期记忆信息中获取相关的知识信息,对工作记忆信息进行更新,然后基于用户输入的语义内容和交际意图,以及更新后的工作记忆信息在当前人机交互的情境下对所述用户输入进行回复的规划。

[0077] 例如,结合图4所示,接收到的第一用户输入为:你知道电影C的主演是谁吗? 第一用户输入的理解结果的语义内容为电影C,交际意图为问答。可以根据对第一用户输入的理解结果从长期记忆信息中获取与电影C相关的信息,并添加到工作记忆信息中,对工作记忆信息进行更新。也就是说,将图4中,核心节点“电影C”所在的节点单元添加到工作记忆信息中,对工作记忆信息进行更新。然后根据交际意图和工作记忆信息中与核心节点“电影C”相关的信息对所述第一用户输入进行回复的规划。例如,可以规划第一对话目标为问答,规划第一对话内容为“张三”是主演。神经网络系统基于第一用户输入,以及第一对话目标规划和第一对话内容规划的整合结果生成“张三”作为回答。

[0078] 接下来,接收到的第二用户输入:我很喜欢张三。第二用户输入的理解结果的语义内容为张三,交际意图为闲聊。如果判断此时工作记忆中没有存储关于“张三”的相关信息,可以根据对第二用户输入的理解结果从长期记忆信息中获取与“张三”相关的信息并添加到工作记忆信息中,对工作记忆信息进行更新。也就是说,将图4中,核心节点“张三”所在的节点单元添加到工作记忆信息中,对工作记忆信息进行更新;然后根据交际意图和工作记忆信息中与“张三”相关的信息对所述第二用户输入进行回复的规划。例如,可以规划第二对话目标为闲聊,规划第二对话内容为“有爱心”。神经网络系统基于第二用户输入,以及第二对话目标和第二对话内容规划的整合结果生成“她很有爱心”作为回答。

[0079] 若用户输入在长期记忆信息中没有相关的知识信息,则回复规划为空。神经网络系统基于用户输入生成回答。

[0080] 在经过设定轮数(例如两轮或三轮)的闲聊后,可以规划对话目标为推荐,基于工作记忆信息中与当前人机交互内容对应的上一核心节点推荐关联程度较高的其它节点的知识信息,从而能够在多次闲聊之后主动切换知识点,避免尬聊。

[0081] 例如,在上面的例子中,接收到的第三用户输入为:除了有爱心,她还很有才华。交际意图为闲聊。然后根据交际意图和工作记忆信息中与当前人机交互内容对应的上一核心节点“张三”相关程度较高的其它节点对所述第三用户输入进行回复的规划。例如,可以规划第三对话目标为推荐,然后可以根据工作记忆信息中与当前人机交互内容对应的上一核心节点“张三”从长期记忆信息中获取与核心节点“张三”相关程度较高的其它节点的信息,例如,可以获取热度较高的核心节点“电影D”。基于此,可以规划第三对话内容为“电影D”和“很法式的一部短片”。神经网络系统基于第三用户输入,以及第三对话目标和第三对话内容规划的整合结果生成“给您推荐一部张三主演的电影D,很法式的一部短片”作为回答。

[0082] 基于此,根据一些实施例,在没有针对所述用户输入更新所述工作记忆信息的情况下,响应于所述关联程度最高的节点仍不能满足预定的标准(例如,各候选回复内容的分值均未达到预定的阈值),则重新查询长期记忆信息以对所述工作记忆进行更新,从而能够主动推荐或切换聊天知识点,避免尬聊。

[0083] 下面根据一个示例性实施例来描述本公开中基于神经网络的人机交互方法。图5中示意了该示例性实施例中对话控制系统的工作过程示意图,其中箭头示意信号流方向,①②……⑩示意方法的步骤。

[0084] 参见图5所示,接收到当前用户输入后,对当前用户输入进行对话理解,得到当前用户输入的交际意图和语义内容,根据交际意图和语义内容从长期记忆信息中获取在当前人机交互的情景下当前用户输入的相关节点的信息;可以根据关联程度对获取的所述相关

节点进行关联打分,然后基于关联分数进行排序,将关联分数添加到相关节点的逻辑控制信息中并且融合到工作记忆信息中,更新工作记忆信息;从工作记忆信息中获取当前人机交互的历史交互数据、以及当前用户输入相关节点的信息,进行对话控制,所述对话控制包括对话目标规划和对话内容规划,若规划的对话目标例如为主动推荐,可以从长期记忆信息中获取与当前用户输入关联程度较高的其它节点的信息,以实现主动推荐知识聊天;整合规划的对话目标和对话内容提供给神经网络系统的解码器;解码器根据规划的对话目标和对话内容的整合以及对当前用户输入和已存储的当前人机交互的历史交互信息进行编码得到的隐式向量,生成针对当前用户输入的回复。

[0085] 根据本公开的另一方面,如图2所示,还提供一种基于神经网络的人机交互装置,可包括:神经网络系统,被配置为接收用户输入作为第一输入;和不同于所述神经网络系统的对话控制系统102,被配置为接收所述用户输入,其中,所述对话控制系统102被进一步配置为基于所述与用户输入相关的信息对所述用户输入进行处理,并将处理结果作为第二输入提供给所述神经网络系统;以及所述神经网络系统被进一步配置为基于所述第一输入和第二输入生成针对所述用户输入的回复。

[0086] 所述神经网络系统可以但不限于采用端到端神经网络系统101。所述端到端神经网络系统101可包括编码器1011和解码器1012。所述编码器1011可以对输入的文本内容进行隐式表示,生成向量;所述解码器1012可以根据给定的输入向量,生成流畅的自然语言文本。

[0087] 根据一些实施例,所述编码器1011可以用于接收所述用户输入以及已存储的当前人机交互的历史交互信息,并对用户输入以及已存储的当前人机交互的历史交互信息进行编码,生成隐式向量,所述隐式向量输入至编码器1012。所述解码器1012可以用于接收所述第二输入(即对话控制系统对用户输入进行处理得到的处理结果)以及编码器1011生成的隐式向量,并且生成所述针对所述用户输入的回复。由此,神经网络系统能够基于当前用户输入、已存储的当前人机交互的历史交互信息、以及对话控制系统基于与用户输入相关的信息对用户输入进行处理所得到的结果,来生成针对用户输入的回复,从而能够进一步保证机器回复内容符合当前人机交互的场景,对话逻辑清晰。

[0088] 所述端到端神经网络系统例如可以采用Transformer神经网络系统或UniLM神经网络系统。

[0089] 根据一些实施例,所述装置还可以包括存储计算系统103。存储计算系统103可以包括长期记忆模块1031和工作记忆模块1032。在这种情况下,所述与所述用户输入相关的信息可以包括取自长期记忆模块的长期记忆信息和取自所述工作记忆模块的仅在当前人机交互期间有效的工作记忆信息。长期记忆信息可以是对话系统需要长时间存储的信息,可以包括各种知识信息,例如可包括下列中的至少其中之一:常识、领域知识、语言知识、问答库以及对话库。工作记忆信息可以为基于当前人机交互内容从长期记忆信息中而获得。也就是说,工作记忆信息是与当前人机交互内容相关的知识信息。从而能够将与当前人机交互内容相关的知识信息融入到基于神经网络系统的对话系统中,实现基于相关知识信息在当前人机交互的情境下对用户输入进行回复规划,充分利用知识信息,使得人机交互的内容丰富、逻辑清晰。可以理解的是,与所述用户输入相关的信息也可以包括从网络上实时抓取的信息,在此不作限定。

[0090] 根据一些实施例,长期记忆信息可以但不限于包括意图知识图谱、问答库和对话库。下面将首先描述意图知识图谱、问答知识和对话库的数据内容、数据组织形式等。

[0091] 意图知识图谱能够从对话场景的知识交互需求出发,不仅满足知识查询功能,更多的是满足多轮多场景交互中的联想、类比、预测等。意图知识图谱的节点的有序组织,便于文本计算,控制知识信息,并且通过知识信息计算,能够支持对话中的行为跳转(场景跳转、同一场景下内容跳转),具有很强的语义转移逻辑性。意图知识图谱整合了不同类型的多场景信息,能够提供多视角理解语言的能力。

[0092] 根据一些实施例,所述长期记忆模块1031中存储有意图知识图谱,意图知识图谱可以包括形式为包含节点和有向边的第一有向图的知识信息,并且其中,所述第一有向图中的节点是包含语义内容和逻辑控制信息的结构化数据。所述第一有向图中的有向边表示相关节点之间的关联属性以及节点与相应的逻辑控制信息之间的关联属性。可以理解的是,其它知识信息也可以采用第一有向图的数据组织形式,并不局限于意图知识图谱,在此仅是以意图知识图谱为例来说明如何利用第一有向图来表示知识信息。

[0093] 根据一些实施例,意图知识图谱的所述逻辑控制信息可包括能够用于筛选与当前人机交互相关的节点的信息,例如,热度、时效性、情感等信息,用于筛选与当前人机交互内容相关的节点,从而能够实现在用户主动发起知识聊天的情况下去检索相关的知识信息,使得对话内容逻辑清晰。例如,可以设置第一热度阈值,从工作记忆信息中与当前人机交互内容相关的节点中筛选出相应的逻辑控制信息中的热度大于所述第一热度阈值的节点。可以设置第一有效时间点,从工作记忆信息中与当前人机交互内容相关的节点中筛选出相应的逻辑控制信息中的时效信息在所述第一有效时间点之后的节点。可以设置第一预设情感类型,从与当前人机交互内容相关的节点中筛选出相应的逻辑控制信息中的情感类型为所述第一预设情感类型的节点。

[0094] 根据一些实施例,意图知识图谱的所述逻辑控制信息还可包括能够用于确定在当前人机交互中各节点之间的相关程度的信息,例如,热度、节点间的关联关系等信息,用于扩展与当前人机交互内容相关的节点,从而能够实现机器主动切换、触发或推荐知识聊天,使得对话内容丰富,避免尬聊。例如,可以设置第二热度阈值,从长期记忆信息中用户输入的各相关节点中获取相应的逻辑控制信息中的热度大于所述第二热度阈值的节点。可以根据关联关系,从当前的节点扩展到与该当前节点关联程度最高的节点。

[0095] 意图知识图谱的所述节点可以包括多个不同类型的节点。根据一些实施例,所述第一有向图中的各节点可包括第一类型节点和第二类型节点。所述第二类型节点的语义内容可以是与所述第二类型节点相关联的第一类型节点的语义内容的一部分,并且所述第二类型节点的逻辑控制信息包括下列中的至少一者:所述第二类型节点在与所述第二类型节点相关联的第一类型节点下的热度、所述第二类型节点与至少一个其他第二类型节点的关联跳转关系、所述第二类型节点的子类型。由此,通过查询第一类型节点即可获得与第一类型节点语义相关联的第二类型节点的知识信息,便于文本计算,控制知识信息。

[0096] 所述第一类型节点例如可以为核心节点,所述第二类型节点例如可以为标签节点。所述有向边可以表示核心节点之间、核心节点与标签节点之间的关联属性。核心节点和标签节点可为结构化数据,从而能够实现语义内容的理解和控制。核心节点可是语义完整的基本单元,可以包含实体、概念、事件、指令,例如可以为例如,人物、物品、结构、产品、建

筑、地点、组织、事件、艺术作品、科学技术、科学定理等。核心节点的逻辑控制信息可以包括热度、时效性、用于召回标签节点的所有标签、任务API等。每一个核心节点可以包括多个相关联的标签节点。标签节点的语义内容可以是与标签节点相关联的核心节点语义内容的一部分，与核心节点是部分与整体的关系。

[0097] 根据一些实施例，与当前人机交互有关的信息可以包括从第一有向图中获取的与用户输入相关的节点信息。可以将用户输入映射到所述第一有向图的核心节点上，映射得到的核心节点以及与映射得到的核心节点相关的标签节点可以均作为与用户输入相关的知识信息。若用户输入无法映射到所述第一有向图的核心节点上，可以将当前人机交互的历史用户输入映射得到的核心节点作为与当前用户输入对应的核心节点。例如，若当前用户输入为“主角是谁？”，则当前用户输入在所述第一有向图中无对应的核心节点。这种情况下，可以将当前人机交互中上一次在所述第一有向图中对应的核心节点作为当前用户输入的核心节点，以获取与当前用户输入相关的知识信息。其中，当前人机交互内容可包括当前用户输入以及当前人机交互的历史交互信息。

[0098] 如图3所示，实线圆形（“电影A”、“电影B”、“赵六”）示意的是核心节点，实线椭圆形示意的是标签节点，虚线圆形示意的是逻辑控制信息。每一个虚线椭圆可围住一个节点单元，作为与用户输入相关的信息单元。也就是说，当用户输入映射到一个节点单元（如图3中的节点单元100）的核心节点上时，节点单元的所有节点信息均认为与用户输入相关的知识信息，并将其添加到工作记忆信息中。需要说明的是，根据系统可用计算资源的大小，也可以将与映射得到的一个核心节点相关联的至少一个其它核心节点所在的节点单元也认为与用户输入相关，也添加到工作记忆信息中，在此不作限定。以下内容中将以映射得到的核心节点所在的节点单元作为与用户输入相关的知识信息为例，来具体描述本公开的技术方案。

[0099] 以第一类型节点（核心节点）为电影实体“电影A”为例，“电影A”用户召回标签节点的标签可以包括演员、角色、导演、场景等。与第一类型节点相关联的标签节点（第二类型节点）可以包括“赵六”、“角色A”、“角色B”、“李四”、“知名场景”。其中，标签节点“赵六”对应相关联的第一类型节点“电影A”的演员标签，“角色A”和“角色B”对应相关联的第一类型节点“电影A”的角色标签，“李四”对应相关联的第一类型节点“电影A”的导演标签，“知名场景”对应相关联的第一类型节点“电影A”的场景标签。与核心节点“电影A”相关联的核心节点可包括“电影B”，与标签节点“赵六”相关联的核心节点可包括“赵六”。当用户输入映射到核心节点“电影A”，则可以将核心节点“电影A”以及与标签节点“赵六”、“角色A”、“角色B”、“李四”、“知名场景”均作为与用户输入相关的信息。

[0100] 根据一些实施例，所述第一有向图中的各节点还可包括第三类型节点，所述第三类型节点的语义内容支持多模态内容，并且所述第三类型节点的逻辑控制信息可包括下列中的至少一者：与所述第三类型节点相关联的第二类型节点的信息、用于表征所述第三类型节点的语义内容的信息。由此，通过设置第三类型节点，能够支持多模态的语义内容，更加丰富对话内容。

[0101] 所述第三类型节点例如可以为内容节点。所述有向边还可以表示标签节点（第二类型节点）和内容节点之间的关联属性。内容节点可为非结构化数据，能够支持丰富的多模态内容。每一个核心节点（第一类型节点）可包括多个内容节点，与内容节点相关联的标签

节点可以是该内容节点的主题或概括。内容节点可包含对话内容,具有多模态(可以为词、句子、图片或视频等)、多样性、细粒度等特点。内容节点的逻辑控制信息例如可包括核心标签、关键词、核心标签在内容节点的语义内容中的重要程度、内容节点的语义内容的概括性短语、与内容节点相关联的标签节点的类别、与内容节点相关联的标签节点的情感极性、与内容节点相关联的标签节点的得分等。

[0102] 根据一些实施例,与当前人机交互内容有关的信息可以包括从第一有向图中获取的与用户输入相关的节点信息。可以将用户输入映射到所述第一有向图的核心节点上,映射得到的核心节点、与映射得到的核心节点相关联的标签节点、以及与所得到的标签节点相关联的内容标签可以均作为与用户输入相关的信息。

[0103] 如图3所示,矩形框示意的是内容节点。以第一类型节点(核心节点)为电影实体“电影A”为例,与第一类型节点相关联的标签节点(第二类型节点)可以包括“赵六”、“角色A”、“角色B”、“李四”、“知名场景”。其中,与标签节点“赵六”和标签节点“角色A”相关联的内容节点可包括“角色A剧照.jpg”,与标签节点“角色A”相关联的内容节点还可包括“角色A具有极其完全的生命力和生命意志”,与标签节点“李四”相关联的内容节点可包括:《电影A》是李四导演的武侠巅峰之作。当用户输入映射到核心节点“电影A”,则可以将核心节点“电影A”、与标签节点“赵六”、“角色A”、“角色B”、“李四”、“知名场景”、以及与标签节点“赵六”、“角色A”、“李四”相关联的内容节点均作为与用户输入相关的信息。

[0104] 两个节点相关联可以是指这两个节点之间可通过包括至少一条有向边的有向路径关联起来。不同节点之间可以通过有向边连接,并示意被连接的节点之间的关联属性。有向边例如可以包括核心节点到核心节点之间的关联边、核心节点到标签节点的关联边、标签节点到核心节点的关联边、标签节点到内容节点的关联边。有向边的属性可以包括语义关系(如:导演、作品、妻子等)、逻辑关系(时序、因果等)、相关性强度、语义上下位关系等多种类型。

[0105] 例如,如图3中,核心节点“电影A”和核心节点“电影B”之间的有向边的属性可为相关性强度,标签节点“赵六”和核心节点“赵六”之间的有向边的属性可为相关性强度。核心节点“电影A”与标签节点“李四”、“赵六”、“角色A”、“知名场景”之间的有向边的属性为语义关系。标签节点“赵六”与内容节点“角色A剧照.jpg”之间的有向边的属性可为语义关系。

[0106] 根据一些实施例,所述长期记忆模块1031中可存储有对话库,对话库可包括形式为包含节点和有向边的第二有向图,用于记录人机交互过程中的语义信息和其特征,为在当前人机交互的情境下对所述用户输入进行回复的规划提供参考。由于基于对话库能够实现利用大数据来获取用户更加倾向的意图,因此,能够为针对用户输入的回复规划提供合理引导。所述第二有向图可以和上述的第一有向图(例如,意图知识图谱)同构,参见图3所示,在此不再详述。由此,通过设置对话库和意图知识图谱为同构的有向图,能够有效实现对话库和意图知识图谱进行融合,便于控制知识信息。可以理解的是,其它知识信息也可以采用第二有向图的数据组织形式,并不局限于对话库,在此仅是以对话库例来说明如何利用第二有向图来表示知识信息。通过设置不同知识信息采用同构的有向图,能够有效实现不同知识信息的融合,便于控制知识信息。

[0107] 根据一些实施例,问答库可以是形式为问题-答案的问答知识信息。问答库的功能为:针对用户的问题,查询问答库,返回与问题相匹配的答案,满足用户的信息需求。例如,

可以在用户输入为问答的情况下,优先从问答库中查询是否有与用户输入匹配的答案,能够快速实现回复。

[0108] 根据一些实施例,所述长期记忆模块中存储的长期记忆信息可以包括意图知识图谱、对话库和问答库。上述通过示例描述了长期记忆信息的意图知识图谱、对话库和问答库的数据内容和数据组织形式,仅是举例说明,并不作为一种限定。当然,所述长期记忆信息也可以为其它与当前人机交互有关的知识信息组合,在此不作限定。

[0109] 长期记忆信息还可进行语言计算与信息抽取。语言计算可以包括比较、归纳、演绎、推理等;信息抽取例如可以包括概念抽取、实体抽取、事件抽取、指令抽取等,从而能够基于用户输入从长期记忆信息中获取与当前人机交互内容相关的工作记忆信息。当前人机交互内容可包括当前用户输入以及当前用户输入之前的历史交互信息。工作记忆信息还可以包括当前人机交互内容,从而能够基于当前人机交互历史和从长期记忆信息中获取的与用户输入相关的知识信息,获取在当前人机交互的情境下对所述用户输入进行回复的规划,将在以下内容中详细介绍。

[0110] 根据一些实施例,所述工作记忆模块1032中可存储有工作记忆信息。所述工作记忆信息包括形式为包含节点和有向边的第三有向图的信息,所述第三有向图可以和上述的第一有向图(例如,意图知识图谱)同构。由此,通过设置工作记忆信息包括与长期记忆信息的知识信息同构的信息,能够方便知识信息的调用和融合。优选地,所述第三有向图可以为所述第一有向图的与当前人机交互相关的一部分,更加方便知识信息的调用和融合。也就是说,所述第三有向图可包括核心节点和标签节点,从而所有用户意图和系统回复(意图)也都能够尽可能地映射到工作记忆信息中的核心节点和相关联的标签节点上,便于各模块使用。另外,由于只需要从长期记忆信息中提取与当前人机交互相关的部分节点信息,可以减少占用内存和提高回复效率。所述第三有向图还可包括内容节点,支持多模态的语义内容,从而能够基于工作记忆信息也获得丰富的对话内容。可以理解的是,所述第三有向图也可以和上述的第一有向图不同构。

[0111] 根据一些实施例,所述工作记忆信息还可包括取自于所述第一有向图的与所述当前人机交互相关的所有节点的语义内容和逻辑控制信息。也就是说,所述第三有向图的核心节点包括对应于第一有向图的所述第一类型节点的语义内容和逻辑控制信息,所述标签节点包括对应于第一有向图的所述第二类型节点的语义内容和逻辑控制信息,所述内容节点包括对应于第一有向图的所述第三类型节点的语义内容和逻辑控制信息。由此,工作记忆信息能够基于当前人机交互从长期记忆信息中尽可能地获取所有可聊的话题,以便于基于工作记忆信息即能够实现针对用户输入进行回复规划。由于工作记忆信息中的数据量远小于长期记忆信息中的数据量,从而能够提高回复速度,提升用户体验。

[0112] 根据一些实施例,当工作记忆信息中没有与用户输入对应的节点信息时,可以基于用户输入从长期记忆信息中获取与用户输入相关的知识信息,对工作记忆信息进行更新;然后基于更新后的工作记忆信息,在当前人机交互的情境下对所述用户输入进行回复的规划。根据一些实施例,可以基于用户输入从第一有向图中获取与用户输入相关的子图,将获取的子图融合到工作记忆信息中的第三有向图中,对工作记忆信息进行更新。

[0113] 对于工作记忆信息,与当前人机交互内容中的历史交互信息对应的各节点中,可以仅保留核心节点的语义内容和逻辑控制信息,而不保留与核心节点相关联的标签节点和

内容节点,从而能够减少计算资源需求。因为已聊过的话题很可能不会再次涉及,因此,仅保留当前人机交互中的历史交互信息对应的核心节点的语义内容和逻辑控制信息,对人机交互的影响很小。

[0114] 根据一些实施例,所述工作记忆信息还可包括用于标记在所述当前人机交互中已经被涉及的语义内容的第一信息,从而能够区分聊过和没聊过的消息,避免重复。根据一些实施例,在第三有向图中,对于在所述当前人机交互中已经被涉及的语义内容相关的所有节点(包括目标节点、标签节点和内容节点)还可包括所述第一信息,以指示该节点已被聊过。

[0115] 根据一些实施例,所述工作记忆信息还可包括用于指明首先提及所述已经被涉及的语义内容的对话方的第二信息,从而能够准确区分聊过相关内容的话题,更精确的避免对话方的对话重复。根据一些实施例,在第三有向图中,对于在所述当前人机交互中已经被涉及的语义内容相关的所有节点(包括目标节点、标签节点和内容节点)还可包括所述第二信息,以指示该节点被哪个对话方聊过。

[0116] 根据一些实施例,所述工作记忆信息还可包括在所述当前人机交互期间的交互记录的历史数据,从而能够获取当前人机交互的场景,以便于为多轮策略提供决策特征。

[0117] 所述工作记忆信息还可包括其它信息,例如,所述对话控制系统的各工作模块的分析结果,方便各模块的使用。例如,除了上述的基于对用户输入从长期记忆信息中获取的与用户输入相关的知识信息,还可包括从长期记忆信息中获取的与用户输入相关的知识信息进行排序的结果、回复决策结果。

[0118] 根据一些实施例,所述对话控制系统可被配置为执行以下步骤以对所述用户输入进行处理:在当前人机交互的情境下对所述用户输入进行回复规划。从而能够充分利用相关信息,并且基于相关信息在当前人机交互的情境下对用户输入进行回复规划,进一步使得人机交互的内容丰富、逻辑清晰。

[0119] 根据一些实施例,对话控制系统102可以包括对话理解模块1021和对话控制模块1022。在一个实施例中,可以利用对话理解模块1021首先基于用户输入从长期记忆信息中获取相关的知识信息,对工作记忆信息进行更新;然后对话控制模块1022基于更新后的工作记忆信息,在当前人机交互的情境下对所述用户输入进行回复的规划。

[0120] 所述对话理解模块1021可被配置用于:分析所述用户输入的语义内容;以及分析所述用户在当前人机交互中对应于所述用户输入的交际意图。也就是说,用户输入的理解结果可以包括语义内容和交际意图。交际意图例如可以选择意图体系中的其中之一,例如,提问、澄清、建议、拒绝、鼓励或安慰等等。

[0121] 根据一些实施例,可以基于意图知识图谱来对所述用户输入进行理解。例如,结合图4所示,接收到的第一用户输入为:你知道电影C的主演是谁吗? 第一用户输入的理解结果的语义内容为电影C,交际意图为问答。接收到的第二用户输入:我很喜欢张三。第二用户输入的理解结果的语义内容为张三,交际意图为闲聊。

[0122] 用户输入的理解结果还可包括描述用户状态的状态意图,例如,用户的心情状态、用户是否喜欢当前聊天。从而能够结合用户的状态意图来进行对话决策,规划回复内容。

[0123] 根据一些实施例,可以基于训练完成的意图神经网络模型来理解用户输入的交际意图。可获取第一用户输入样本集,采用人工标注第一用户输入样本集中的常用用户输入

样本的交际意图。利用所述第一用户输入样本集对意图神经网络模型进行训练。例如可以基于日志数据(例如,搜索引擎日志)来获取所述第一用户输入样本集。还可获取低频用户输入(例如“我不知道你说的啥”)并人工标记低频用户输入的交际意图,生成语料库。对于意图神经网络模型无法识别交际意图的用户输入,即意图体系无对应的交际意图,则可以在所述语料库中查找与用户输入语义相似度最高的低频用户输入,并将查找得到的低频用户输入相应的交际意图作为用户输入的交际意图,从而能够保证理解用户输入的交际意图。

[0124] 以下将以意图知识图谱来详细描述对所述用户输入进行理解的过程。

[0125] 根据一些实施例,如图6所示,所述对话理解模块1021可包括:判定子模块10211,被配置用于判断所述用户输入是否能够对应于所述工作记忆信息中的某一节点;处理子模块10212,被配置用于响应于所述用户输入能够对应于所述工作记忆信息中的某一节点,基于所述工作记忆信息对所述用户输入进行处理,从而能够基于工作记忆信息来理解用户输入的语义内容,实现在当前人机交互的情景下对用户输入的理解,提高对话理解的准确性和效率。所述某一节点例如可以为第三有向图的节点,如上面内容所描述的,所述第三有向图可以与所述第一有向图(意图知识图谱)同构并且为所述第一有向图的一部分。

[0126] 根据一些实施例,所述处理子模块10212被进一步配置用于基于所述工作记忆信息中的所述某一节点的信息,为所述用户输入补充相关的内容。例如,用户输入为“主角是谁”,可以基于从工作记忆信息中找到的对应于所述用户输入的某一核心节点“电影A”,据此可将用户输入补全为“电影A的主角是谁”。根据一些实施例,可以查找当前人机交互内容对应于工作记忆信息中的上一核心节点,并进行以下判定:用户输入是否被该上一核心节点的逻辑控制信息中的标签覆盖,若被覆盖,则根据该上一节点为所述用户输入补充相关的内容。例如,该上一核心节点“电影A”的逻辑控制信息中的标签包括:演员、角色、导演、场景。由于“主角”和“演员”的语义相同,确定用户输入被核心节点“电影A”的标签覆盖,根据核心节点“电影A”,将用户输入补全为“电影A的主角是谁”。

[0127] 可以基于补全后的用户输入来进一步分析所述用户输入的语义内容,以提高对话理解的准确性。

[0128] 根据一些实施例,所述对话理解模块可被进一步配置用于响应于所述用户输入不能够对应于所述工作记忆信息中的节点,从所述长期记忆模块中提取与所述用户输入相关的节点信息存储到所述工作记忆模块。从而能够在用户输入未被工作记忆信息中的知识信息覆盖的情况下,扩大知识范围(例如,基于整个意图图谱),尝试基于知识信息来理解用户输入。

[0129] 根据一些实施例,所述对话理解模块1021还可包括:消歧子模块10213,被配置用于对所述用户输入进行消歧处理。例如,用户输入“我爱读水浒传”,其中经分词处理后得到的“水浒传”存在歧义,其可能是电视剧,也可能是小说。因此,需要对用户输入进行消歧处理,确定“水浒传”的类型,以能够准确用户输入的语义内容。

[0130] 可以基于消歧处理的结果来进一步分析所述用户输入的语义内容,以提高对话理解的准确性。

[0131] 根据一些实施例,所述消歧子模块10213可被进一步配置用于基于用户输入和所述工作记忆信息中与当前人机交互相关的节点信息,识别所述用户输入中的具有歧义的至

少部分内容,并确定所述至少部分内容在当前人机交互情境中的含义。从而能够基于当前人机交互的情景来对用户输入进行消歧处理。例如,用户输入“我爱读水浒传”,由于“水浒传”既可以指小说、也可以指电视剧,因而具有歧义,在此情形下,系统通过用户输入中的“读”,确定出“水浒传”在当前语境中的真实含义应当是指小说,而不是电视剧。作为一个示例性实施例,可以基于用户输入和工作记忆信息中与当前人机交互对应的上一核心节点(可为工作记忆信息中最新更新的核心节点,包括语义内容和逻辑控制信息),来对所述用户输入进行消歧处理。例如,可将用户输入和工作记忆信息中与当前人机交互内容对应的上一核心节点输入消歧神经网络模型,以获取消歧神经网络模型输出的用户输入中的具有歧义的至少部分内容的类型。可以利用类型语料库来对消歧神经网络模型进行度量训练,以能够实现用户输入和所述工作记忆信息中与当前人机交互相关的节点的信息的组合,与类型语料库中相应的类型距离更近,以能够输出用户输入中的具有歧义的至少部分内容以及具有歧义的所述至少部分内容的类型。在此仅是举例说明如何确定用户输入中的具有歧义的至少部分内容以及具有歧义的所述至少部分内容的类型,并不是一种限定。

[0132] 可以理解的是,在此并不限定仅能根据工作记忆信息来进行消歧处理,例如,还可基于长期工作信息中的对话库来对用户输入进行消歧处理。例如,在对话库中,对于输入“我爱水浒传”更加倾向于阅读的意图,则可确定“水浒传”的类型为小说。

[0133] 根据一些实施例,所述对话理解模块1021可被配置用于执行消歧处理和信息补全处理。可以基于消歧处理的结果和补全后的用户输入来进一步分析所述用户输入的语义内容,以提高对话理解的准确性。

[0134] 可以基于交际意图来决定后续的操作。例如,交际意图为查询,可以基于消歧处理的结果和补全后的用户输入和交际意图生成意图查询表达式,以查找相关的知识信息。交际意图为根据意图表达式道别,则不用查找相关知识信息。在需要查找相关知识信息的情况下,可以首先在工作记忆信息中查找是否有与用户输入相关的知识信息,若无,则在长期记忆信息中继续查找是否有与用户输入相关的知识信息。

[0135] 根据一些实施例,所述对话理解模块1021还可包括:查询子模块10214,被配置用于根据用户输入的语义内容和在当前人机交互中对应于所述用户输入的交际意图,从工作记忆信息中查询用户输入的相关节点的信息;排序子模块10215,被配置用于根据与所述用户输入的关联程度,对查询得到的所述用户输入的相关节点进行排序,其中所述排序是基于所述相关节点的逻辑控制信息进行的。例如,可以基于热度或时效性等进行打分,来确定相关节点与所述用户输入的关联程度,从而能够根据相关节点与所述用户输入的关联程度来进行对话决策,实现对话系统生成的回复与用户输入的关联性。

[0136] 根据一些实施例,所述对话理解模块1021被进一步配置用于根据与所述用户输入的关联程度,为所述相关节点赋予不同的分值,从而能够为对话决策提供参考。例如,可以在工作记忆信息中的第三有向图的核心节点的逻辑控制信息中增加与所述用户输入的关联分数。

[0137] 分析得到的用户输入的语义内容例如可以为第三有向图中的与用户输入相关的核心节点。

[0138] 根据一些实施例,在根据与所述用户输入的关联程度,为所述相关节点赋予不同的分值的情况下,所述对话控制模块被配置用于执行以下操作以在当前人机交互的情境下

对所述用户输入进行回复的规划:根据所述排序的结果,规划对话目标并选择与所述用户输入的关联程度最高的节点信息作为规划的对话内容;以及,整合所述规划的对话内容和所述对话目标,作为所述第二输入提供给所述神经网络系统,从而对话系统能够融入知识信息,根据用户输入进行答复规划,使得对话逻辑清晰。

[0139] 可以理解的是,可以基于意图知识图谱中节点的逻辑控制信息来获取所述相关节点与用户输入的关联程度,也可以基于对话库获取所述相关节点与用户输入的关联程度,还可以基于用户喜好来获取所述相关节点与用户输入的关联程度,在此不作限定,只要能够从知识信息中获取所述相关节点与用户输入的关联程度即可。可以基于用户的当前人机交互内容和历史人机交互内容来获取该用户的喜好,例如,该用户在多次人机交互中均涉及阅读,可以确定该用户喜好阅读,在对话决策的过程中可以根据用户的喜好来规划对话内容。

[0140] 根据一些实施例,在工作记忆信息中没有对应用户输入的知识信息的情况下,可以基于用户输入的语义内容从长期记忆信息中获取相关的知识信息,对工作记忆信息进行更新,然后基于用户输入的语义内容和交际意图,以及更新后的工作记忆信息在当前人机交互的情境下对所述用户输入进行回复的规划。

[0141] 例如,结合图4所示,接收到的第一用户输入为:你知道电影C的主演是谁吗? 第一用户输入的理解结果的语义内容为电影C,交际意图为问答。可以根据对第一用户输入的理解结果从长期记忆信息中获取与电影C相关的信息,并添加到工作记忆信息中,对工作记忆信息进行更新。也就是说,将图4中,核心节点“电影C”所在的节点单元添加到工作记忆信息中,对工作记忆信息进行更新。然后根据交际意图和工作记忆信息中与核心节点“电影C”相关的信息对所述第一用户输入进行回复的规划。例如,可以规划第一对话目标为问答,规划第一对话内容为“张三”是主演。神经网络系统基于第一用户输入,以及第一对话目标规划和第一对话内容规划的整合结果生成“张三”作为回答。

[0142] 接下来,接收到的第二用户输入:我很喜欢张三。第二用户输入的理解结果的语义内容为张三,交际意图为闲聊。如果判断此时工作记忆中没有存储关于“张三”的相关信息,可以根据对第二用户输入的理解结果从长期记忆信息中获取与“张三”相关的信息并添加到工作记忆信息中,对工作记忆信息进行更新。也就是说,将图4中,核心节点“张三”所在的节点单元添加到工作记忆信息中,对工作记忆信息进行更新;然后根据交际意图和工作记忆信息中与“张三”相关的信息对所述第二用户输入进行回复的规划。例如,可以规划第二对话目标为闲聊,规划第二对话内容为“有爱心”。神经网络系统基于第二用户输入,以及第二对话目标和第二对话内容规划的整合结果生成“她很有爱心”作为回答。

[0143] 若用户输入在长期记忆信息中没有相关的知识信息,则回复规划为空。神经网络系统基于用户输入生成回答。

[0144] 在经过设定轮数(例如两轮或三轮)的闲聊后,可以规划对话目标为推荐,基于工作记忆信息中与当前人机交互内容对应的上一核心节点推荐关联程度较高的其它节点的知识信息,从而能够在多次闲聊之后主动切换知识点,避免尬聊。

[0145] 例如,在上面的例子中,接收到的第三用户输入为:除了有爱心,她还很有才华。交际意图为闲聊。然后根据交际意图和工作记忆信息中与当前人机交互内容对应的上一核心节点“张三”相关程度较高的其它节点对所述第三用户输入进行回复的规划。例如,可以规

划第三对话目标为推荐,然后可以根据工作记忆信息中与当前人机交互内容对应的上一核心节点“张三”从长期记忆信息中获取与核心节点“张三”相关程度较高的其它节点的信息,例如,可以获取热度较高的核心节点“电影D”。基于此,可以规划第三对话内容为“电影D”和“很法式的一部短片”。神经网络系统基于第三用户输入,以及第三对话目标和第三对话内容规划的整合结果生成“给您推荐一部张三主演的电影D,很法式的一部短片”作为回答。

[0146] 基于此,根据一些实施例,在没有针对所述用户输入更新所述工作记忆信息的情况下,响应于所述关联程度最高的节点不能满足预定的标准(例如,各候选回复内容的分值均未达到预定的阈值),则重新查询长期记忆信息对所述工作记忆进行更新,从而能够主动推荐或切换聊天知识点,避免尬聊。

[0147] 根据本公开的另一方面,还提供一种电子设备,包括:处理器;以及存储程序的存储器,所述程序包括指令,所述指令在由所述处理器执行时使所述处理器执行根据上述的方法。

[0148] 根据本公开的另一方面,还提供一种存储程序的计算机可读存储介质,所述程序包括指令,所述指令在由电子设备的处理器执行时,致使所述电子设备执行根据上述的方法。

[0149] 参见图7所示,现将描述计算设备2000,其是可以应用于本公开的各方面的硬件设备(电子设备)的示例。计算设备2000可以是被配置为执行处理和/或计算的任何机器,可以是但不限于工作站、服务器、台式计算机、膝上型计算机、平板计算机、个人数字助理、机器人、智能电话、车载计算机或其任何组合。上述方法可以全部或至少部分地由计算设备2000或类似设备或系统实现。

[0150] 计算设备2000可以包括(可能经由一个或多个接口)与总线2002连接或与总线2002通信的元件。例如,计算设备2000可以包括总线2002、一个或多个处理器2004、一个或多个输入设备2006以及一个或多个输出设备2008。一个或多个处理器2004可以是任何类型的处理器,并且可以包括但不限于一个或多个通用处理器和/或一个或多个专用处理器(例如特殊处理芯片)。输入设备2006可以是能向计算设备2000输入信息的任何类型的设备,并且可以包括但不限于鼠标、键盘、触摸屏、麦克风和/或遥控器。输出设备2008可以是能呈现信息的任何类型的设备,并且可以包括但不限于显示器、扬声器、视频/音频输出终端、振动器和/或打印机。计算设备2000还可以包括非暂时性存储设备2010,非暂时性存储设备可以是非暂时性的并且可以实现数据存储的任何存储设备,并且可以包括但不限于磁盘驱动器、光学存储设备、固态存储器、软盘、柔性盘、硬盘、磁带或任何其他磁介质,光盘或任何其他光学介质、ROM(只读存储器)、RAM(随机存取存储器)、高速缓冲存储器和/或任何其他存储器芯片或盒、和/或计算机可从其读取数据、指令和/或代码的任何其他介质。非暂时性存储设备2010可以从接口拆卸。非暂时性存储设备2010可以具有用于实现上述方法和步骤的数据/程序(包括指令)/代码。计算设备2000还可以包括通信设备2012。通信设备2012可以是使得能够与外部设备和/或与网络通信的任何类型的设备或系统,并且可以包括但不限于调制解调器、网卡、红外通信设备、无线通信设备和/或芯片组,例如蓝牙™设备、1302.11设备、WiFi设备、WiMax设备、蜂窝通信设备和/或类似物。

[0151] 计算设备2000还可以包括工作存储器2014,其是可以存储对处理器2004的工作有用的程序(包括指令)和/或数据的任何类型的工作存储器,并且可以包括但不限于随

机存取存储器和/或只读存储器设备。

[0152] 软件要素(程序)可以位于工作存储器2014中,包括但不限于操作系统2016、一个或多个应用程序2018、驱动程序和/或其他数据和代码。用于执行上述方法和步骤的指令可以被包括在一个或多个应用程序2018中,并且上述构建方法可以通过由处理器2004读取和执行一个或多个应用程序2018的指令来实现。更具体地,上述方法中,步骤S101~步骤S105可以例如通过处理器2004执行具有步骤S101~步骤S105的指令的应用程序2018而实现。此外,上述方法中的其它步骤可以例如通过处理器2004执行具有执行相应步骤中的指令的应用程序2018而实现。软件要素(程序)的指令的可执行代码或源代码可以存储在非暂时性计算机可读存储介质(例如上述存储设备2010)中,并且在执行时可以被存入工作存储器2014中(可能被编译和/或安装)。软件要素(程序)的指令的可执行代码或源代码也可以从远程位置下载。

[0153] 还应该理解,可以根据具体要求而进行各种变型。例如,也可以使用定制硬件,和/或可以用硬件、软件、固件、中间件、微代码,硬件描述语言或其任何组合来实现特定元件。例如,所公开的方法和设备中的一些或全部可以通过使用根据本公开的逻辑和算法,用汇编语言或硬件编程语言(诸如VERILOG,VHDL,C++)对硬件(例如,包括现场可编程门阵列(FPGA)和/或可编程逻辑阵列(PLA)的可编程逻辑电路)进行编程来实现。

[0154] 还应该理解,前述方法可以通过服务器-客户端模式来实现。例如,客户端可以接收用户输入的数据并将所述数据发送到服务器。客户端也可以接收用户输入的数据,进行前述方法中的一部分处理,并将处理所得到的数据发送到服务器。服务器可以接收来自客户端的数据,并且执行前述方法或前述方法中的另一部分,并将执行结果返回给客户端。客户端可以从服务器接收到方法的执行结果,并例如可以通过输出设备呈现给用户。

[0155] 还应该理解,计算设备2000的组件可以分布在例如云平台的网络上。例如,可以使用一个处理器执行一些处理,而同时可以由远离该一个处理器的另一个处理器执行其他处理。计算系统2000的其他组件也可以类似地分布。这样,计算设备2000可以被解释为在多个位置执行处理的分布式计算系统。

[0156] 虽然已经参照附图描述了本公开的实施例或示例,但应理解,上述的方法、系统和设备仅仅是示例性的实施例或示例,本发明的范围并不由这些实施例或示例限制,而是仅由授权后的权利要求书及其等同范围来限定。实施例或示例中的各种要素可以被省略或者可由其等同要素替代。此外,可以通过不同于本公开中描述的次序来执行各步骤。进一步地,可以以各种方式组合实施例或示例中的各种要素。重要的是随着技术的演进,在此描述的很多要素可以由本公开之后出现的等同要素进行替换。

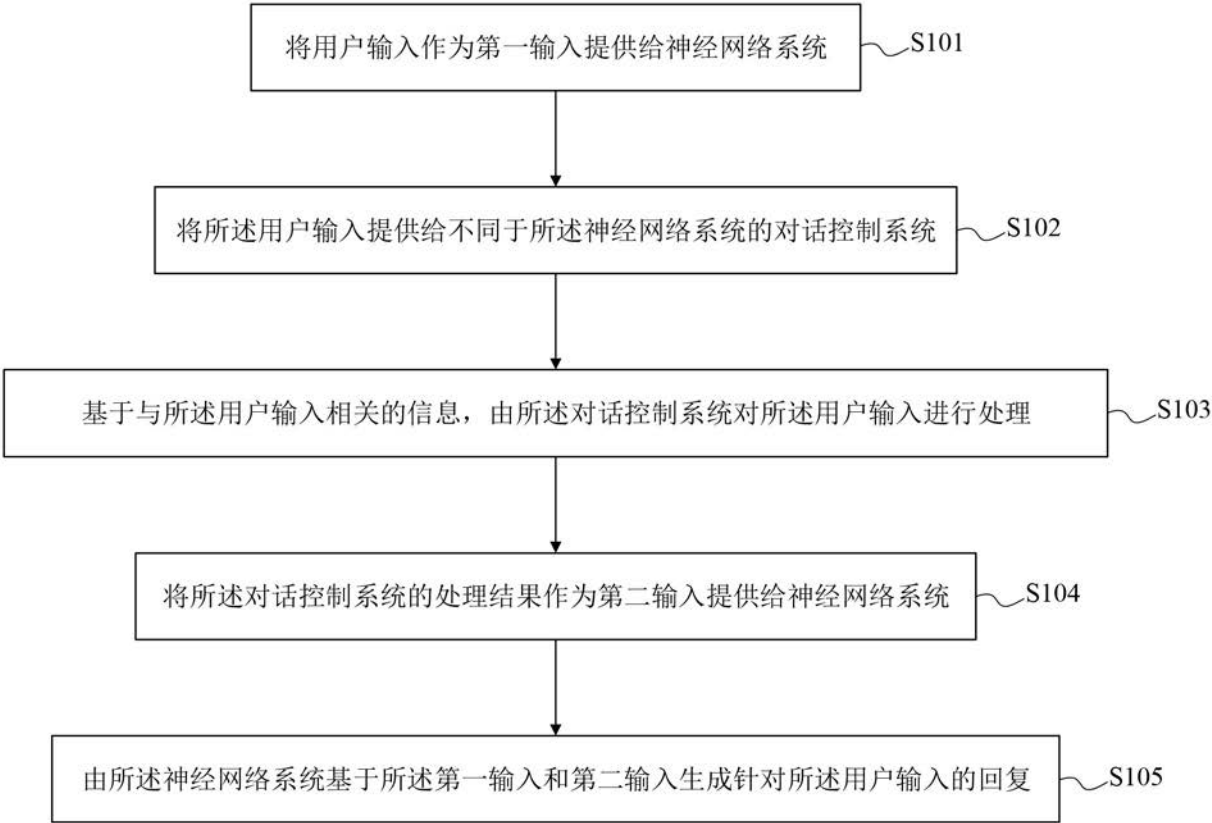


图 1

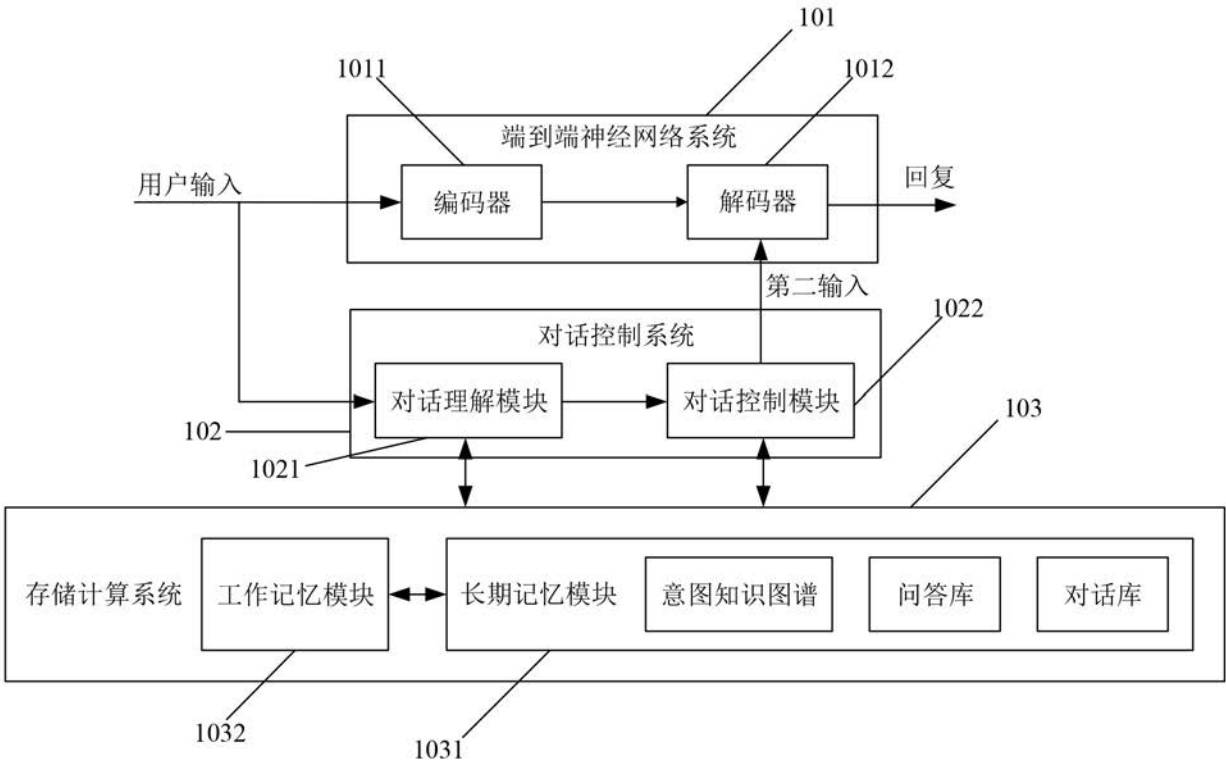


图 2

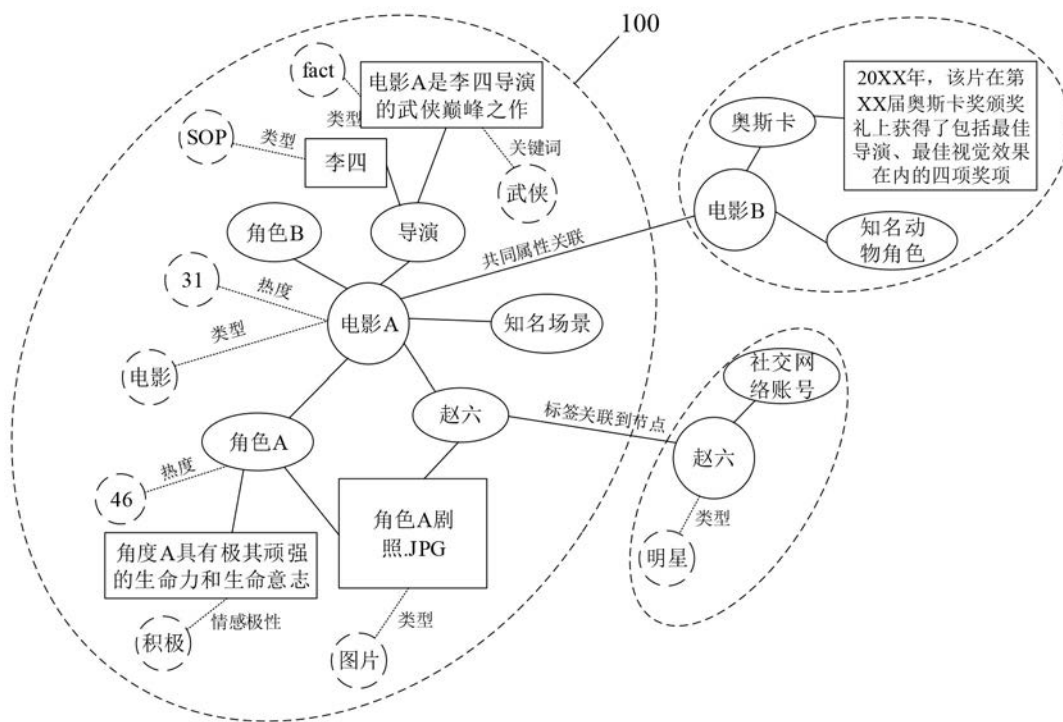


图 3

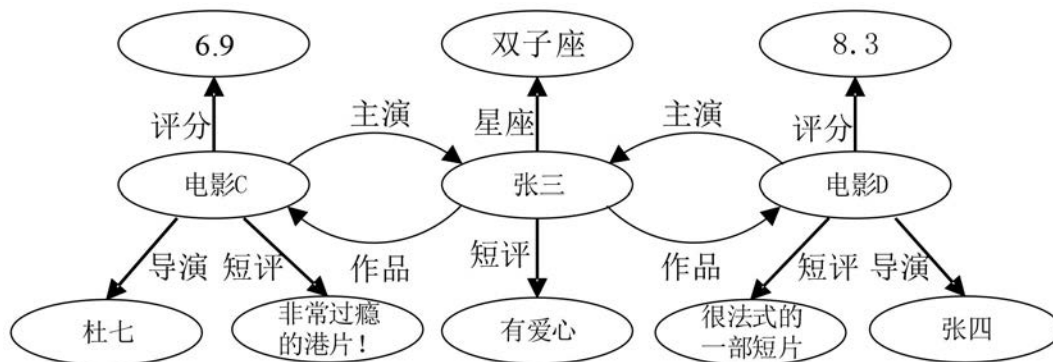


图 4

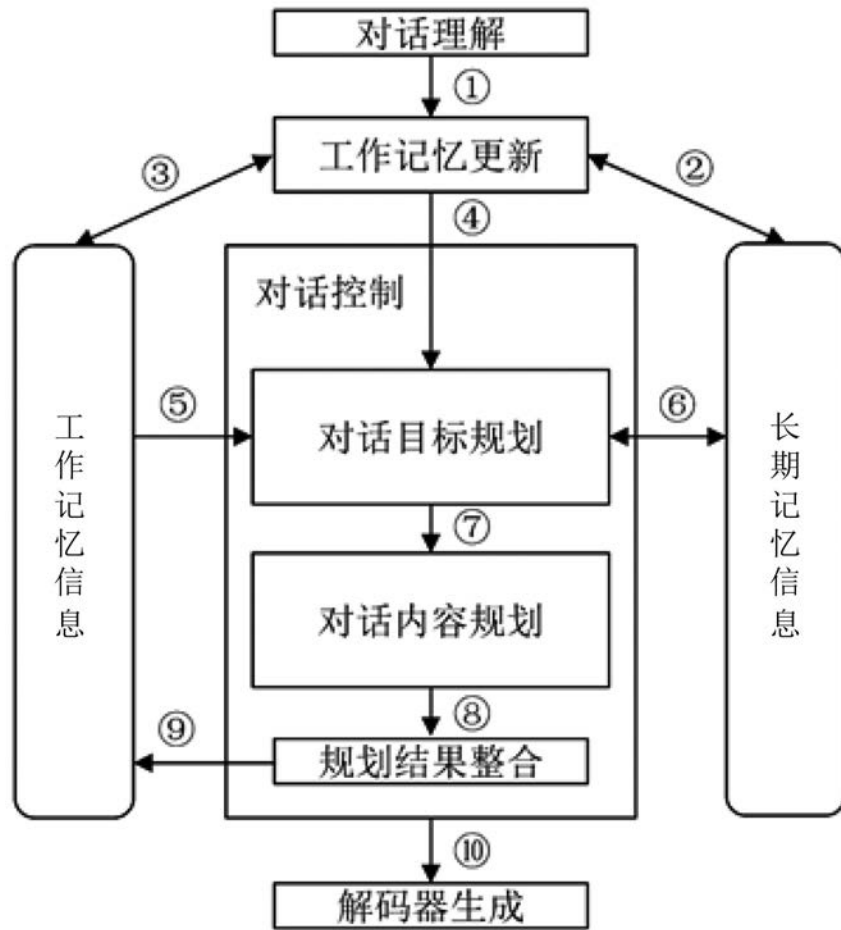


图 5

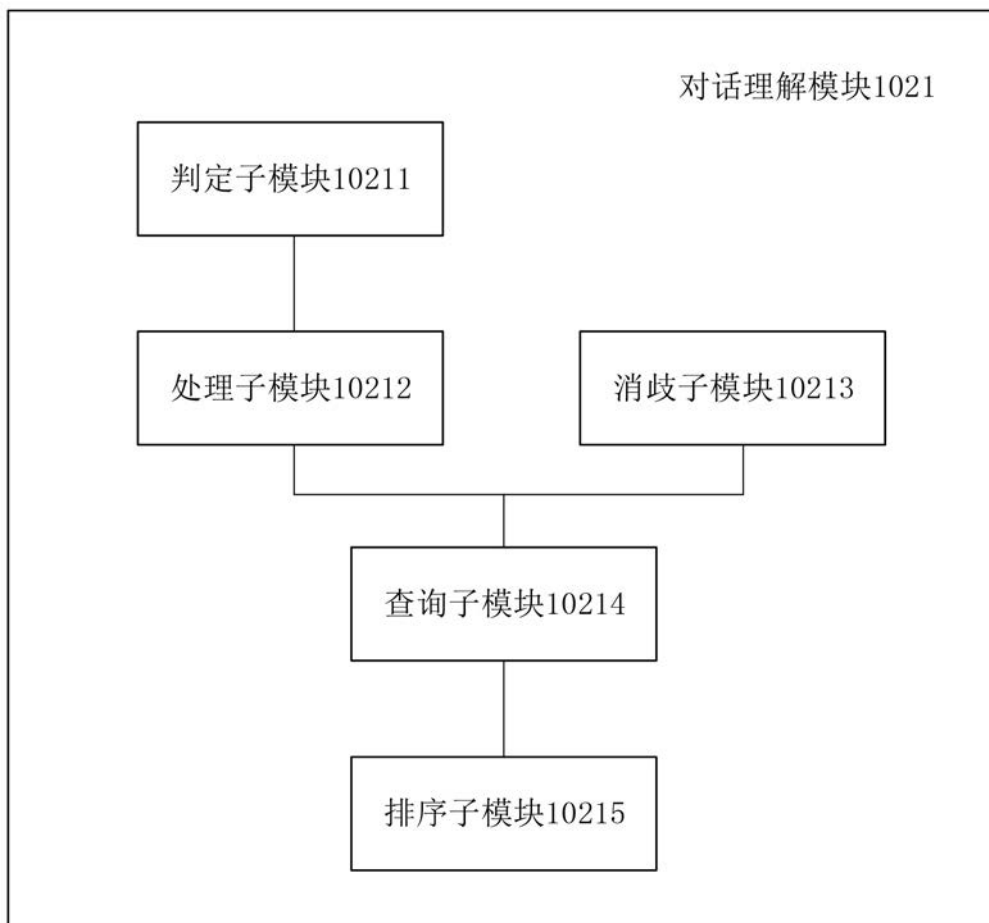


图 6

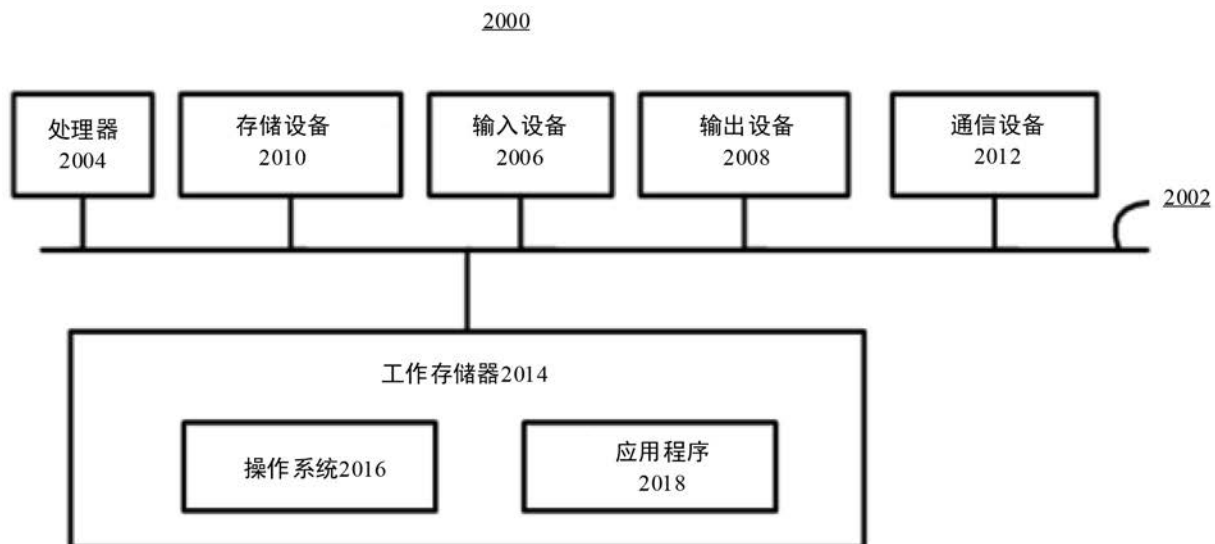


图 7