



(12) 发明专利

(10) 授权公告号 CN 112926326 B

(45) 授权公告日 2024.01.19

(21) 申请号 202110191377.X

(22) 申请日 2021.02.20

(65) 同一申请的已公布的文献号
申请公布号 CN 112926326 A

(43) 申请公布日 2021.06.08

(73) 专利权人 深圳追一科技有限公司
地址 518057 广东省深圳市南山区粤海街
道科技园社区科苑路8号讯美科技广
场3号楼23A、23B

(72) 发明人 崔健

(74) 专利代理机构 华进联合专利商标代理有限
公司 44224
专利代理师 纪婷婧

(51) Int. Cl.
G06F 40/295 (2020.01)

(56) 对比文件

CN 111898382 A, 2020.11.06

CN 110688467 A, 2020.01.14

CN 109918680 A, 2019.06.21

WO 2020139865 A1, 2020.07.02

审查员 徐千慧

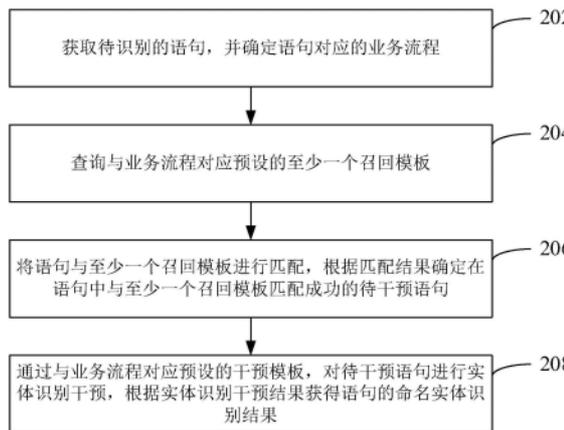
权利要求书2页 说明书16页 附图3页

(54) 发明名称

命名实体识别方法、装置、计算机设备和存储介质

(57) 摘要

本申请涉及一种命名实体识别方法、装置、计算机设备和存储介质。所述方法包括：获取待识别的语句，并确定语句对应的业务流程；查询与业务流程对应预设的至少一个召回模板；将语句与至少一个召回模板进行匹配，根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句；通过与业务流程对应预设的干预模板，对待干预语句进行实体识别干预，根据实体识别干预结果获得语句的命名实体识别结果。采用本方法能够提高命名实体识别的准确性。



1. 一种命名实体识别方法,其特征在于,所述方法包括:
 - 获取待识别的语句,并确定所述语句对应的业务流程;
 - 查询与所述业务流程对应预设的至少一个召回模板;
 - 将所述语句与所述至少一个召回模板进行匹配,根据匹配结果确定在所述语句中与所述至少一个召回模板匹配成功的待干预语句;
 - 通过与所述业务流程对应预设的干预模板,对所述待干预语句进行实体识别干预,根据实体识别干预结果获得所述语句的命名实体识别结果。
2. 根据权利要求1所述的方法,其特征在于,所述将所述语句与所述至少一个召回模板进行匹配,根据匹配结果确定在所述语句中与所述至少一个召回模板匹配成功的待干预语句,包括:
 - 确定所述至少一个召回模板中的召回特征字段;
 - 将所述召回特征字段在所述语句中进行字符匹配,当得到的字符匹配结果为匹配成功时,确定匹配成功的目标召回模板;
 - 基于所述目标召回模板从所述语句中确定待干预语句。
3. 根据权利要求2所述的方法,其特征在于,所述基于所述目标召回模板从所述语句中确定待干预语句,包括:
 - 确定所述目标召回模板中的待干预槽位;
 - 根据所述召回特征字段和所述待干预槽位将所述语句进行语句划分,得到语句划分结果;
 - 将所述语句划分结果中与所述待干预槽位对应的字段确定为待干预语句。
4. 根据权利要求1所述的方法,其特征在于,所述通过与所述业务流程对应预设的干预模板,对所述待干预语句进行实体识别干预,根据实体识别干预结果获得所述语句的命名实体识别结果,包括:
 - 查询与所述业务流程对应预设的正向干预模板;
 - 将所述正向干预模板与所述待干预语句的各字段进行正向匹配,得到正向匹配结果;
 - 当所述正向匹配结果为匹配成功时,将所述待干预语句中与所述正向干预模板匹配成功的字段,确定为所述语句的命名实体识别结果。
5. 根据权利要求1所述的方法,其特征在于,所述通过与所述业务流程对应预设的干预模板,对所述待干预语句进行实体识别干预,根据实体识别干预结果获得所述语句的命名实体识别结果,包括:
 - 查询与所述业务流程对应预设的负向干预模板;
 - 将所述负向干预模板与所述待干预语句的各字段进行负向匹配,得到负向匹配结果;
 - 当所述负向匹配结果为匹配成功时,确定所述语句的命名实体识别结果中不包括所述待干预语句中与所述负向干预模板匹配成功的字段。
6. 根据权利要求1所述的方法,其特征在于,所述通过与所述业务流程对应预设的干预模板,对所述待干预语句进行实体识别干预,根据实体识别干预结果获得所述语句的命名实体识别结果,包括:
 - 查询与所述业务流程对应预设的正向干预模板和负向干预模板;
 - 将所述负向干预模板和所述正向干预模板分别与所述待干预语句的各字段进行匹配,

得到负向匹配结果和正向匹配结果；

当所述负向匹配结果和所述正向匹配结果均为匹配成功、且所述待干预语句中与所述负向干预模板匹配成功的字段和所述待干预语句中与所述正向干预模板匹配成功的字段为相同字段时，确定所述语句的命名实体识别结果中不包括所述相同字段。

7. 根据权利要求1至6任意一项所述的方法，其特征在于，所述方法还包括：

确定所述业务流程对应的待填充词槽；

将所述命名实体识别结果填充至所述待填充词槽中，得到所述业务流程对应的业务词槽；

当所述业务词槽满足流程执行条件时，基于所述业务词槽执行所述业务流程。

8. 一种命名实体识别装置，其特征在于，所述装置包括：

语句获取模块，用于获取待识别的语句，并确定所述语句对应的业务流程；

召回模板查询模块，用于查询与所述业务流程对应预设的至少一个召回模板；

待干预语句确定模块，用于将所述语句与所述至少一个召回模板进行匹配，根据匹配结果确定在所述语句中与所述至少一个召回模板匹配成功的待干预语句；

识别干预模块，用于通过与所述业务流程对应预设的干预模板，对所述待干预语句进行实体识别干预，根据实体识别干预结果获得所述语句的命名实体识别结果。

9. 一种计算机设备，包括存储器和处理器，所述存储器存储有计算机程序，其特征在于，所述处理器执行所述计算机程序时实现权利要求1至7中任一项所述的方法的步骤。

10. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，所述计算机程序被处理器执行时实现权利要求1至7中任一项所述的方法的步骤。

命名实体识别方法、装置、计算机设备和存储介质

技术领域

[0001] 本申请涉及人工智能技术领域,特别是涉及一种命名实体识别方法、装置、计算机设备和存储介质。

背景技术

[0002] 随着人工智能(Artificial Intelligence, AI)技术的发展,自然语言处理(Natural Language Processing, NLP)技术在语音识别、语音翻译、理解完整的句子、理解匹配词的同义词,以及生成语法正确完整句子和段落等方面得到广泛的应用。作为自然语言处理的一个基础任务,命名实体识别(Named Entities Recognition, NER)的目的在于识别语料中人名、地名、组织机构名等有具体含义的实体,如从语句中识别出人名、地名、机构名、时间、日期等命名实体。

[0003] 目前,在针对有歧义的语句进行识别时,如对于“南京都”进行命名实体识别,传统的命名实体识别方法可能识别到“南京”或“京都”,命名实体识别的准确性较低。

发明内容

[0004] 基于此,有必要针对上述技术问题,提供一种能够提高命名实体识别准确性的命名实体识别方法、装置、计算机设备和存储介质。

[0005] 一种命名实体识别方法,所述方法包括:

[0006] 获取待识别的语句,并确定语句对应的业务流程;

[0007] 查询与业务流程对应预设的至少一个召回模板;

[0008] 将语句与至少一个召回模板进行匹配,根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句;

[0009] 通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果。

[0010] 在其中一个实施例中,将语句与至少一个召回模板进行匹配,根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句,包括:确定至少一个召回模板中的召回特征字段;将召回特征字段在语句中进行字符匹配,当得到的字符匹配结果为匹配成功时,确定匹配成功的目标召回模板;基于目标召回模板从语句中确定待干预语句。

[0011] 在其中一个实施例中,基于目标召回模板从语句中确定待干预语句,包括:确定目标召回模板中的待干预槽位;根据召回特征字段和待干预槽位将语句进行语句划分,得到语句划分结果;将语句划分结果中与待干预槽位对应的字段确定为待干预语句。

[0012] 在其中一个实施例中,通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果,包括:查询与业务流程对应预设的正向干预模板;将正向干预模板与待干预语句的各字段进行正向匹配,得到正向匹配结果;当正向匹配结果为匹配成功时,将待干预语句中与正向干预模板匹配成功的字段,确定为语句的命名实体识别结果。

[0013] 在其中一个实施例中,通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果,包括:查询与业务流程对应预设的负向干预模板;将负向干预模板与待干预语句的各字段进行负向匹配,得到负向匹配结果;当负向匹配结果为匹配成功时,确定语句的命名实体识别结果中不包括待干预语句中与负向干预模板匹配成功的字段。

[0014] 在其中一个实施例中,通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果,包括:查询与业务流程对应预设的正向干预模板和负向干预模板;将负向干预模板和正向干预模板分别与待干预语句的各字段进行匹配,得到负向匹配结果和正向匹配结果;当负向匹配结果和正向匹配结果均为匹配成功、且待干预语句中与负向干预模板匹配成功的字段和待干预语句中与正向干预模板匹配成功的字段为相同字段时,确定语句的命名实体识别结果中不包括相同字段。

[0015] 在其中一个实施例中,所述方法还包括:确定业务流程对应的待填充词槽;将命名实体识别结果填充至待填充词槽中,得到业务流程对应的业务词槽;当业务词槽满足流程执行条件时,基于业务词槽执行业务流程。

[0016] 一种命名实体识别装置,所述装置包括:

[0017] 语句获取模块,用于获取待识别的语句,并确定语句对应的业务流程;

[0018] 召回模板查询模块,用于查询与业务流程对应预设的至少一个召回模板;

[0019] 待干预语句确定模块,用于将语句与至少一个召回模板进行匹配,根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句;

[0020] 识别干预模块,用于通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果。

[0021] 在其中一个实施例中,待干预语句确定模块包括特征字段确定模块、字符匹配模块和目标召回模板模块;其中:特征字段确定模块,用于确定至少一个召回模板中的召回特征字段;字符匹配模块,用于将召回特征字段在语句中进行字符匹配,当得到的字符匹配结果为匹配成功时,确定匹配成功的目标召回模板;目标召回模板模块,用于基于目标召回模板从语句中确定待干预语句。

[0022] 在其中一个实施例中,目标召回模板模块包括待干预槽位确定模块、语句划分模块和划分结果处理模块;其中:待干预槽位确定模块,用于确定目标召回模板中的待干预槽位;语句划分模块,用于根据召回特征字段和待干预槽位将语句进行语句划分,得到语句划分结果;划分结果处理模块,用于将语句划分结果中与待干预槽位对应的字段确定为待干预语句。

[0023] 在其中一个实施例中,识别干预模块包括正向模板查询模块、正向匹配模块和正向干预模块;其中:正向模板查询模块,用于查询与业务流程对应预设的正向干预模板;正向匹配模块,用于将正向干预模板与待干预语句的各字段进行正向匹配,得到正向匹配结果;正向干预模块,用于当正向匹配结果为匹配成功时,将待干预语句中与正向干预模板匹配成功的字段,确定为语句的命名实体识别结果。

[0024] 在其中一个实施例中,识别干预模块包括负向模板查询模块、负向匹配模块和负向干预模块;其中:负向模板查询模块,用于查询与业务流程对应预设的负向干预模板;负

向匹配模块,用于将负向干预模板与待干预语句的各字段进行负向匹配,得到负向匹配结果;负向干预模块,用于当负向匹配结果为匹配成功时,确定语句的命名实体识别结果中不包括待干预语句中与负向干预模板匹配成功的字段。

[0025] 在其中一个实施例中,识别干预模块包括干预模块查询模块、干预模板匹配模块和干预处理模块;其中:干预模块查询模块,用于查询与业务流程对应预设的正向干预模板和负向干预模板;干预模板匹配模块,用于将负向干预模板和正向干预模板分别与待干预语句的各字段进行匹配,得到负向匹配结果和正向匹配结果;干预处理模块,用于当负向匹配结果和正向匹配结果均为匹配成功、且待干预语句中与负向干预模板匹配成功的字段和待干预语句中与正向干预模板匹配成功的字段为相同字段时,确定语句的命名实体识别结果中不包括相同字段。

[0026] 在其中一个实施例中,所述装置还包括词槽确定模块、词槽填充模块和业务处理模块;其中:词槽确定模块,用于确定业务流程对应的待填充词槽;词槽填充模块,用于将命名实体识别结果填充至待填充词槽中,得到业务流程对应的业务词槽;业务处理模块,用于当业务词槽满足流程执行条件时,基于业务词槽执行业务流程。

[0027] 一种计算机设备,包括存储器和处理器,所述存储器存储有计算机程序,所述处理器执行所述计算机程序时实现以下步骤:

[0028] 获取待识别的语句,并确定语句对应的业务流程;

[0029] 查询与业务流程对应预设的至少一个召回模板;

[0030] 将语句与至少一个召回模板进行匹配,根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句;

[0031] 通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果。

[0032] 在其中一个实施例中,处理器执行计算机程序时还实现以下步骤:确定至少一个召回模板中的召回特征字段;将召回特征字段在语句中进行字符匹配,当得到的字符匹配结果为匹配成功时,确定匹配成功的目标召回模板;基于目标召回模板从语句中确定待干预语句。

[0033] 在其中一个实施例中,处理器执行计算机程序时还实现以下步骤:确定目标召回模板中的待干预槽位;根据召回特征字段和待干预槽位将语句进行语句划分,得到语句划分结果;将语句划分结果中与待干预槽位对应的字段确定为待干预语句。

[0034] 在其中一个实施例中,处理器执行计算机程序时还实现以下步骤:查询与业务流程对应预设的正向干预模板;将正向干预模板与待干预语句的各字段进行正向匹配,得到正向匹配结果;当正向匹配结果为匹配成功时,将待干预语句中与正向干预模板匹配成功的字段,确定为语句的命名实体识别结果。

[0035] 在其中一个实施例中,处理器执行计算机程序时还实现以下步骤:查询与业务流程对应预设的负向干预模板;将负向干预模板与待干预语句的各字段进行负向匹配,得到负向匹配结果;当负向匹配结果为匹配成功时,确定语句的命名实体识别结果中不包括待干预语句中与负向干预模板匹配成功的字段。

[0036] 在其中一个实施例中,处理器执行计算机程序时还实现以下步骤:查询与业务流程对应预设的正向干预模板和负向干预模板;将负向干预模板和正向干预模板分别与待干

预语句的各字段进行匹配,得到负向匹配结果和正向匹配结果;当负向匹配结果和正向匹配结果均为匹配成功、且待干预语句中与负向干预模板匹配成功的字段和待干预语句中与正向干预模板匹配成功的字段为相同字段时,确定语句的命名实体识别结果中不包括相同字段。

[0037] 在其中一个实施例中,处理器执行计算机程序时还实现以下步骤:确定业务流程对应的待填充词槽;将命名实体识别结果填充至待填充词槽中,得到业务流程对应的业务词槽;当业务词槽满足流程执行条件时,基于业务词槽执行业务流程。

[0038] 一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现以下步骤:

[0039] 获取待识别的语句,并确定语句对应的业务流程;

[0040] 查询与业务流程对应预设的至少一个召回模板;

[0041] 将语句与至少一个召回模板进行匹配,根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句;

[0042] 通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果。

[0043] 在其中一个实施例中,计算机程序被处理器执行时还实现以下步骤:确定至少一个召回模板中的召回特征字段;将召回特征字段在语句中进行字符匹配,当得到的字符匹配结果为匹配成功时,确定匹配成功的目标召回模板;基于目标召回模板从语句中确定待干预语句。

[0044] 在其中一个实施例中,计算机程序被处理器执行时还实现以下步骤:确定目标召回模板中的待干预槽位;根据召回特征字段和待干预槽位将语句进行语句划分,得到语句划分结果;将语句划分结果中与待干预槽位对应的字段确定为待干预语句。

[0045] 在其中一个实施例中,计算机程序被处理器执行时还实现以下步骤:查询与业务流程对应预设的正向干预模板;将正向干预模板与待干预语句的各字段进行正向匹配,得到正向匹配结果;当正向匹配结果为匹配成功时,将待干预语句中与正向干预模板匹配成功的字段,确定为语句的命名实体识别结果。

[0046] 在其中一个实施例中,计算机程序被处理器执行时还实现以下步骤:查询与业务流程对应预设的负向干预模板;将负向干预模板与待干预语句的各字段进行负向匹配,得到负向匹配结果;当负向匹配结果为匹配成功时,确定语句的命名实体识别结果中不包括待干预语句中与负向干预模板匹配成功的字段。

[0047] 在其中一个实施例中,计算机程序被处理器执行时还实现以下步骤:查询与业务流程对应预设的正向干预模板和负向干预模板;将负向干预模板和正向干预模板分别与待干预语句的各字段进行匹配,得到负向匹配结果和正向匹配结果;当负向匹配结果和正向匹配结果均为匹配成功、且待干预语句中与负向干预模板匹配成功的字段和待干预语句中与正向干预模板匹配成功的字段为相同字段时,确定语句的命名实体识别结果中不包括相同字段。

[0048] 在其中一个实施例中,计算机程序被处理器执行时还实现以下步骤:确定业务流程对应的待填充词槽;将命名实体识别结果填充至待填充词槽中,得到业务流程对应的业务词槽;当业务词槽满足流程执行条件时,基于业务词槽执行业务流程。

[0049] 上述命名实体识别方法、装置、计算机设备和存储介质,通过与待识别的语句对应的业务流程预设的至少一个召回模板,与待识别的语句进行匹配,确定在语句中与至少一个召回模板匹配成功的待干预语句,通过与业务流程对应预设的干预模板对待干预语句进行实体识别干预,并根据实体识别干预结果获得语句的命名实体识别结果。在命名实体识别的处理过程中,通过与业务流程对应预设的召回模板确定待识别的语句中的待干预语句,并通过与业务流程对应预设的干预模板对待干预语句进行实体识别干预,根据不同业务流程配置对应的召回模板和干预模板对语句进行命名实体识别,提高了在不同业务流程中命名实体识别的针对性,从而提高了针对不同业务流程的待识别语句进行命名实体识别处理的准确性。

附图说明

- [0050] 图1为一个实施例中命名实体识别方法的应用环境图;
- [0051] 图2为一个实施例中命名实体识别方法的流程示意图;
- [0052] 图3为一个实施例中确定待干预语句的流程示意图;
- [0053] 图4为一个实施例中实体识别干预的流程示意图;
- [0054] 图5为一个实施例中命名实体识别装置的结构框图;
- [0055] 图6为一个实施例中计算机设备的内部结构图。

具体实施方式

[0056] 为了使本申请的目的、技术方案及优点更加清楚明白,以下结合附图及实施例,对本申请进行进一步详细说明。应当理解,此处描述的具体实施例仅仅用以解释本申请,并不用于限定本申请。

[0057] 本申请提供的命名实体识别方法,可以应用于如图1所示的应用环境中。其中,终端102通过网络与服务器104进行通信。用户在终端102输入触发业务流程处理的语音数据或文本数据,如用户在终端102通过输入的语音数据或文本数据触发订机票的业务流程,终端102获得用户输入的待识别的语句,并将待识别的语句发送至服务器104,服务器104通过与待识别的语句对应的业务流程预设的至少一个召回模板,与待识别的语句进行匹配,确定在语句中与至少一个召回模板匹配成功的待干预语句,通过与业务流程对应预设的干预模板对待干预语句进行实体识别干预,并根据实体识别干预结果获得语句的命名实体识别结果,服务器104可以将命名实体识别结果反馈至终端102,服务器104也可以基于命名实体识别结果执行相应的业务流程,如为用户预订机票,并反馈预订结果至终端102。其中,终端102可以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,服务器104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。

[0058] 在一个实施例中,如图2所示,提供了一种命名实体识别方法,以该方法应用于图1中的服务器为例进行说明,包括以下步骤:

[0059] 步骤202,获取待识别的语句,并确定语句对应的业务流程。

[0060] 其中,语句是需要进行命名实体识别的语料,如可以为文本语句,语句由一个或多个字符组成。例如,对于语句“我明天想坐飞机去北京玩儿几天呢。”语句还可以是基于语音识别技术获得的语音文本,如终端可以通过麦克风采集用户的语音数据,基于语音识别技术

对采集的语音数据进行自然语言理解处理后得到文本语句。命名实体是指语句中具有特定意义的实体,如人名、地名、组织机构、日期、时间、百分数或货币等,命名实体识别即需要从语句中识别出命名实体,如从待识别的语句中识别出人名、地名、机构名、专有名词、时间、日期等。业务流程可以为通过语句执行的业务相应的流程,如用户通过语句进行出租车预订处理时,业务流程即为预订出租车的处理流程。一般地,不同的业务场景对应于不同的业务流程,在不同的业务流程的交互过程中,涉及不同的语句。例如,在预订机票的业务流程中,涉及的语句交互包括出发地点、到达地点和出发时间等;而在预订酒店的业务流程中,涉及的语句交互包括入住时间、离店时间、酒店名称等。

[0061] 具体地,在触发进行命名实体识别处理时,如接收到用户通过终端发送与业务流程相关的语句时,获取待识别的语句,并确定语句对应的业务流程。在具体应用时,可以对待识别的语句进行意图识别,如通过词表穷举法、规则模板解析法、深度学习意图识别模型等,对待识别的语句进行意图识别,根据意图识别结果确定语句对应的业务流程。在具体实现时,若待识别的语句并非触发业务流程的语句,即待识别的语句为触发业务流程后的交互语句,如用户在触发预订飞机票的业务流程后,待识别的语句为用户在人机交互过程中关于出发时间的描述时,可以根据待识别的语句的上下文确定待识别的语句对应的业务流程。

[0062] 步骤204,查询与业务流程对应预设的至少一个召回模板。

[0063] 其中,召回模板用于判定语句中是否存在需要进行实体识别干预的字段,例如可以判定语句中是否存在有歧义的字段,从而需要对语句的命名实体识别处理进行干预,以获得准确的实体识别结果。例如,召回模板可以为“*ABC*”,其中“*”为通配符,用于指代一个或若干个字符,“A”、“B”和“C”为三种字符,该召回模板表示在语句中存在“ABC”时,需要对语句中的“ABC”字段的实体识别进行干预,如干预后可以识别成“A”、“B”、“C”“AB”或“BC”等,具体识别结果根据实际需要进行配置。召回模板根据业务流程对应预先设置,即不同的业务流程可以配置不同的召回模板,从而可以根据不同的业务流程,对相同的语句通过不同的召回模板对是否需要进行实体识别干预进行准确判定,以获得准确的命名实体识别结果。

[0064] 具体地,得到待识别的语句,并确定语句对应的业务流程后,服务器查询与业务流程对应预设的至少一个召回模板。具体应用时,可以针对不同的业务流程预先构建相应的召回模板库,召回模板库中存储有与该业务流程关联的至少一个召回模板。服务器在确定语句对应的业务流程后,可以查询该业务流程对应的召回模板库,从而从召回模板库中查询得到与业务流程对应预设的至少一个召回模板。

[0065] 步骤206,将语句与至少一个召回模板进行匹配,根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句。

[0066] 其中,待干预语句为待识别的语句中需要进行实体识别干预处理的字段,待干预语句与召回模板匹配成功,即语句中的待干预语句命中了召回模板,需要对待干预语句进行实体识别干预处理。

[0067] 具体地,服务器查询得到与业务流程对应预设的至少一个召回模板后,将语句与获得的召回模板进行匹配,从而判定语句是否命中了召回模板,若是则需要进行实体识别干预处理,否则不需要。服务器得到匹配结果后,根据匹配结果确定在语句中与至少一个召

回模板匹配成功的待干预语句,如可以将语句中与召回模板匹配成功的字段确定为待干预语句。

[0068] 步骤208,通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果。

[0069] 其中,干预模板用于对语句进行实体识别干预,以得到语句的命名实体识别结果。干预模板根据业务流程对应预先设置,即不同的业务流程可以配置不同的干预模板,从而可以根据不同的业务流程,对相同的语句通过不同的干预模板对语句进行不同的实体识别干预,以获得准确的命名实体识别结果。干预模板可以包括不同干预类型的模板,如可以包括正向干预模板和负向干预模板。例如,正向干预模板可以为“*AB*”,即语句中存在“AB”的字段时,与正向干预模板匹配成功,即语句命中正向干预模板,则可以识别得到语句中的命名实体识别结果包括“AB”。负向干预模板可以为“*BC*”,语句中存在“*BC*”的字段时,与负向干预模板匹配成功,即语句命中负向干预模板,则可以识别得到语句中的命名实体识别结果不包括“BC”,即不识别出语句中的“BC”字段,而识别为其他字段,如语句中为“*ABC*”,且命中“*BC*”的负向干预模板时,命名实体识别的结果可以包括“ABC”、“AB”,但不包括“BC”。

[0070] 具体地,确定语句中需要进行实体识别干预的待干预语句后,服务器查询与业务流程对应预设的干预模板,并通过查询得到的干预模板对待干预语句进行实体识别干预,如确定识别出语句中与干预模板对应匹配的字段,或不识别语句中与干预模板对应匹配的字段,从而根据实体识别干预结果获得语句的命名实体识别结果。

[0071] 在一个具体应用中,业务流程对应的召回模板包括以正则表达式表达的“*123*”,干预模板包括“*12*”和“*23*”。其中,召回模板用户确定是否需要对话句进行实体识别干预,如语句中命中了召回模板,即与“*123*”匹配成功,则需要对话句中的“123”字段进行实体识别干预,以确定在该业务流程中对应的命名实体识别结果。“*12*”为正向干预模板,即命中该正向干预模板时,从语句中识别出“12”的字段;“*23*”为负向干预模板,命中该负向干预模板时,不识别语句中的“23”。对于不同的业务场景下,涉及到不同的业务流程时,可以通过配置不同的召回模板和干预模板,以实现对语句的针对性命名实体识别,得到准确的命名实体识别结果。例如,在A场景下对于语句“123”需要识别出“12”时,则可以配置召回模板为“*123*”,正向干预模板为“*12*”;若在B场景下对于语句“123”需要识别出“23”时,则可以配置召回模板为“*123*”,正向干预模板为“*23*”;若在C场景下对于语句“123”不识别“12”时,则可以配置召回模板为“*123*”,负向干预模板为“*12*”。

[0072] 在另一个应用中,对于语句“南京都”,如场景A对应的业务流程下需要识别出“南京”,则可以配置召回模板为“南京都”,正向干预模板为“南京”,从而可以根据语句“南京都”准确识别出“南京”;若在场景B对应的业务流程下需要识别出“京都”,则可以配置召回模板为“南京都”,正向干预模板为“京都”,而可以根据语句“南京都”准确识别出“京都”;若在场景C对应的业务流程下不识别出“南京”时,则可以配置召回模板为“南京都”,负向干预模板为“南京”,而可以根据语句“南京都”不识别出“南京”,而识别为其他字段。对于相同的语句,在不同的业务流程中,根据业务流程对应配置的召回模板和干预模板可以提高命名实体识别的针对性,从而提高针对不同业务流程的待识别语句进行命名实体识别处理的准确性。

[0073] 上述命名实体识别方法中,通过与待识别的语句对应的业务流程预设的至少一个召回模板,与待识别的语句进行匹配,确定在语句中与至少一个召回模板匹配成功的待干预语句,通过与业务流程对应预设的干预模板对待干预语句进行实体识别干预,并根据实体识别干预结果获得语句的命名实体识别结果。在命名实体识别的处理过程中,通过与业务流程对应预设的召回模板确定待识别的语句中的待干预语句,并通过与业务流程对应预设的干预模板对待干预语句进行实体识别干预,根据不同业务流程配置对应的召回模板和干预模板对语句进行命名实体识别,提高了在不同业务流程中命名实体识别的针对性,从而提高了针对不同业务流程的待识别语句进行命名实体识别处理的准确性。

[0074] 在一个实施例中,如图3所示,确定待干预语句的步骤,即将语句与至少一个召回模板进行匹配,根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句,包括:

[0075] 步骤302,确定至少一个召回模板中的召回特征字段。

[0076] 其中,召回特征字段为确定语句是否命中召回模板的特征字段,如对于召回模板“*123*”,则召回特征字段可以为“123”;又如召回模板为“*去|飞|到*”,召回特征字段可以为“去”“飞”“到”,其中,“|”表示“或”。每个召回模板中均设有对应的召回特征字段,通过与召回特征字段进行匹配,可以确定语句是否命中召回模板。

[0077] 具体地,服务器获得与业务流程对应的至少一个召回模板后,服务器确定召回模板中的召回特征字段。具体应用中,可以根据预设的召回模板格式对各召回模板进行解析,确定召回模板中的召回特征字段。

[0078] 步骤304,将召回特征字段在语句中进行字符匹配,当得到的字符匹配结果为匹配成功时,确定匹配成功的目标召回模板。

[0079] 确定召回模板中的召回特征字段后,将召回特征字段在语句中进行字符匹配,具体由服务器将召回特征字段与语句中的各字段依次进行字符匹配,当得到的字符匹配结果为匹配成功时,确定与语句匹配成功的目标召回模板,目标召回模板为语句匹配成功的召回模板。

[0080] 在具体实现时,召回模板中的召回特征字段可以为一个,如可以为“*A*”,也可以为多个,如可以为“*AB*”;召回特征字段可以为连续的,如“*ABC*”,也可以为非连续的,如“*A*B*CED*”。召回模板中召回特征字段的格式可以根据业务流程的实际需要预先配置。将召回特征字段在语句中进行字符匹配时,可以将召回特征字段与语句的字段一一进行匹配,在召回特征字段的数目为多个时,只有所有字段均按照对应顺序匹配成功,才可以确定对应的召回模板与语句匹配成功。例如,对于召回模板“*A*B*CED*”,其召回特征字段包括按序排列的“A”“B”和“CDE”,其中“A”和“B”之间可以有其它字符,而“C”“D”“E”为连续的,中间不包括其他字符;对于语句1为“*A*B*CED*”,其字段与召回模板的召回特征字段一一对应,则匹配成功;而对于语句2“*A*CED*B*”,虽然也包括“A”“B”和“CDE”字符,但字符顺序与召回模板中召回特征字段的顺序不同,则语句2与该召回模板并不匹配,匹配结果为匹配失败。

[0081] 在召回模板的数量为多个时,可以将所有召回模板中的召回特征字段均分别在语句中进行字符匹配,根据字符匹配结果确定与语句匹配成功的目标召回模板,目标召回模板的数目为多个时,所有目标召回模板均有效,即可以根据所有命中的召回模板触发进行

实体识别干预处理。

[0082] 步骤306,基于目标召回模板从语句中确定待干预语句。

[0083] 得到目标召回模板后,服务器基于目标召回模板从语句中确定待干预语句。具体实现时,服务器可以根据目标召回模板的预设格式,从语句中确定需要进行实体识别干预处理的待干预语句。

[0084] 本实施例中,服务器将召回模板中得到召回特征字段在语句中进行字符匹配,根据字符匹配结果确定目标召回模板,并根据目标召回模板从语句中确定待干预语句,从而可以基于字符匹配确定语句所命中的目标召回模板,以便对语句中的待干预语句进行实体识别干预,提高命名实体识别的准确性。

[0085] 在一个实施例中,基于目标召回模板从语句中确定待干预语句,包括:确定目标召回模板中的待干预槽位;根据召回特征字段和待干预槽位将语句进行语句划分,得到语句划分结果;将语句划分结果中与待干预槽位对应的字段确定为待干预语句。

[0086] 其中,待干预槽位为召回模板中需要进行实体识别干预处理的字段,例如对于召回模板“*ABC {[w:1-4]}”,召回特征字段为“ABC”,而待干预槽位为“ABC”后1-4个字符,即语句匹配到“ABC”的召回特征字段,即语句命中该召回模板时,语句中位于“ABC”后的1-4个字符为待干预槽位,即需要对“ABC”后的1-4个字符进行实体识别干预处理。

[0087] 具体地,确定目标召回模板后,服务器进一步确定目标召回模板中的待干预槽位,待干预槽位在设置召回模板时预先配置,待干预槽位和召回特征字段可以包括不重叠的字段,如对于召回模板“*ABC {[w:1-4]}”;待干预槽位和召回特征字段也可以包括重叠的字段,即部分字段即作为召回特征字段,也作为待干预槽位,如对于召回模板“AB {[w:1-2] C}”,召回特征字段包括“AB”和“C”,而待干预槽位包括位于“AB”和“C”之间的1-2个字符,以及“C”。

[0088] 确定目标召回模板中的待干预槽位后,服务器根据召回特征字段和待干预槽位将语句进行语句划分,如由服务器按照召回特征字段和待干预槽位的顺序和格式,将语句的各字段进行划分,得到语句划分结果。服务器将语句划分结果中与待干预槽位对应的字段确定为待干预语句。待干预语句中的字段与待干预槽位对应,待干预语句中的字段可以包括召回特征字段,也可以不包括召回特征字段。

[0089] 在一个具体应用中,召回模板的配置中字段的配置包括“与”、“或”以及上下文。其中,“与”可以直接连接,不设置特定符号,如“AB”;“或”用小括号加以标识,且用“|”连接,如“(A|B|C)”,表示“A”或者“B”或者“C”;上下文可以根据字段的前后关系确定。例如,{[w:1-4]}表示长度为1~4的任意文本,要识别的待干预槽位存在上文或者下文,从而将符合长度的取最长的字符识别出来。对于只有上文的情况,如可以设置召回模板为“(A|B)C {[w:1-4]}”,具体可以为“(去|飞|到) {[w:2-4]}”,该召回模板中上文为“去”或“飞”或“到”时,将匹配到后面2~4个字符识别为待干预槽位,即将语句中位于“去”或“飞”或“到”后的2-4个字段确定为语句中的待干预语句。此外,召回模板还可以设置为只包括下文或同时包括上下文,具体根据实际需要进行灵活配置。

[0090] 在一个具体实现中,命名实体识别方法应用于多轮对话场景,可以根据针对语句的命名实体识别结果对多轮对话场景进行词槽填充。其中,多轮对话是一种在人机对话中,初步明确用户意图之后,获取必要信息以最终得到明确用户指令的方式,多轮对话与一件

事情的处理相对应。词槽是多轮对话过程中将初步用户意图转化为明确用户指令所需要补充的信息,一个词槽与一件事情的处理中所需要获取的一种信息相对应。此时,召回模板可以设置为“AB {[slot]C}”格式,从而将“[slot]+C”共同组成一个新的词槽,其中,{[slot]}表示系统词槽、自定义词槽、正则识别的结果,{ }外面可以没有上下文。例如,自定义词槽中罗列的姓名,召回模板为“我是 {[slot]先生}”,当语句中出现上文为“我是”,同时命中词槽“[slot]”,以及[slot]后为“先生”时,命中该召回模板,并将“[slot]先生”作为词槽召回,即将“[slot]先生”确定为待干预语句。又如,“[w:min-max]”,表示可以将从min到max数量的字符确定为待干预槽位,实现跳词,但召回模板中不允许两组[w:min-max]连续出现;又如,“{A[w:1-4]}”、“{A[slot]}”,可以将A和[slot]共同组成词槽,即作为待干预槽位。

[0091] 本实施例中,服务器根据目标召回模板中的待干预槽位和召回特征字段将语句进行语句划分,并将语句划分结果中与待干预槽位对应的字段确定为待干预语句,从而按照与业务流程对应预设的召回模板从语句中提取得到待干预语句,以对待干预语句进行实体识别干预,获得准确的命名实体识别结果。

[0092] 在一个实施例中,如图4所示,实体识别干预的处理,即通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果,包括:

[0093] 步骤402,查询与业务流程对应预设的正向干预模板。

[0094] 其中,干预模板包括正向干预模板,正向干预模板在命中时,则识别出对应的字段,即语句与正向干预模板匹配时,确定语句中与正向干预模板匹配的字段为命名实体识别结果。例如,正向干预模板可以为“*AB*”,即语句中存在“AB”的字段时,与正向干预模板匹配成功,即语句命中正向干预模板,则可以识别得到语句中的命名实体识别结果包括“AB”。正向干预模板与业务流程对应预先根据实际需要配置。

[0095] 具体地,在进行实体识别干预的处理时,服务器查询与业务流程对应预设的正向干预模板。具体应用中,服务器可以查询业务流程对应预设的干预模板库,从该干预模板库中获得与业务流程对应预设的正向干预模板。

[0096] 步骤404,将正向干预模板与待干预语句的各字段进行正向匹配,得到正向匹配结果。

[0097] 得到正向干预模板后,服务器将正向干预模板与待干预语句的各字段进行正向匹配,具体可以由服务器将正向干预模板与待干预语句的各字段进行字符匹配,得到正向匹配结果。

[0098] 步骤406,当正向匹配结果为匹配成功时,将待干预语句中与正向干预模板匹配成功的字段,确定为语句的命名实体识别结果。

[0099] 当正向匹配结果为匹配成功时,即待干预语句的各字段中存在与正向干预模板一致的字段时,将待干预语句中与正向干预模板匹配成功的字段,确定为语句的命名实体识别结果,从而根据正向干预模板识别出语句中的命名实体。

[0100] 本实施例中,通过服务器将与业务流程对应预设的正向干预模板,与待干预语句的各字段进行正向匹配,将待干预语句中与正向干预模板匹配成功的字段,确定为语句的命名实体识别结果,从而利用正向干预模板识别出语句中的与业务流程关联的命名实体,可以避免语句中存在歧义时无法准确进行命名实体识别的问题,提高命名实体识别的准确

性。

[0101] 在一个实施例中,通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果,包括:查询与业务流程对应预设的负向干预模板;将负向干预模板与待干预语句的各字段进行负向匹配,得到负向匹配结果;当负向匹配结果为匹配成功时,确定语句的命名实体识别结果中不包括待干预语句中与负向干预模板匹配成功的字段。

[0102] 其中,干预模板包括负向干预模板,负向干预模板在命中时,则不识别出对应的字段,即语句与负向干预模板匹配时,确定语句中与负向干预模板匹配的字段不是命名实体识别结果,从而通过负向干预模板对识别结果进行排除。例如,负向干预模板可以为“*BC*”,语句中存在“*BC”的字段时,与负向干预模板匹配成功,即语句命中负向干预模板,则可以识别得到语句中的命名实体识别结果不包括“BC”,即不识别出语句中的“BC”字段,而识别为其他字段,如语句中为“*ABC*”,且命中“*BC*”的负向干预模板时,命名实体识别的结果可以包括“ABC”、“AB”,但不包括“BC”。

[0103] 具体地,服务器查询与业务流程对应预设的负向干预模板,具体应用中,服务器可以查询业务流程对应预设的干预模板库,从该干预模板库中获得与业务流程对应预设的负向干预模板。得到负向干预模板后,服务器将负向干预模板与待干预语句的各字段进行负向匹配,具体可以由服务器将负向干预模板与待干预语句的各字段进行字符匹配,得到负向匹配结果。当负向匹配结果为匹配成功时,即待干预语句的各字段中存在与负向干预模板一致的字段时,确定语句的命名实体识别结果中不包括待干预语句中与负向干预模板匹配成功的字段,从而将待干预语句中与负向干预模板匹配成功的字段的实体识别方式进行排序,不识别得到相应的命名实体识别结果。

[0104] 本实施例中,通过服务器将与业务流程对应预设的负向干预模板,与待干预语句的各字段进行负向匹配,将待干预语句中与负向干预模板匹配成功的字段,确定为语句的命名实体识别结果,从而利用负向干预模板排除语句中的与业务流程关联的命名实体,可以避免语句中存在歧义时无法准确进行命名实体识别的问题,提高命名实体识别的准确性。

[0105] 在一个具体实现中,命名实体识别方法应用于多轮对话场景,可以根据针对语句的命名实体识别结果对多轮对话场景进行词槽填充。此时,在召回模板确定待干预语句后,当且仅当待干预语句命中正向干预模板时,将与正向干预模板匹配的字段进行词槽填充;在待干预语句命中负向干预模板时,将与负向干预模板匹配的字段排除,不作为词槽填充的字段。

[0106] 具体地,干预模板的配置中字段的配置包括“与”、“或”、上下文、以及其他词槽key。其中,“与”可以直接连接,不设置特定符号,如“AB”;“或”用小括号加以标识,且用“|”连接,如“(A|B|C)”,表示“A”或者“B”或者“C”;上下文可以根据字段的前后关系确定;其他词槽key可以为[D]。例如,对于正向干预模板“(A|B)C{{slot}}”,{{slot}}表示系统词槽、自定义词槽、正则识别,模板抽取的结果。具体如“(去|飞|到){{start_city}}”,该正向干预模板为上文为“去”或“飞”或“到”时,同时匹配到后文一个已经识别出的城市词槽时,将命中正向干预模板,并对{}里面的[start_city]进行干预,即识别出[start_city]进行词槽填充。又如干预模板“A[D]{{slot}}”,允许上文中出现其他词槽名来表示一类词槽,具体

如“从[start_city]到{[end_city]}”,当“{”前面为“从”+start_city类型的词槽+“到”,同时后面命中一个识别end_city类型的词槽,命中干预模板,并对end_city的词槽进行干预,即识别出end_city进行词槽填充或不进行词槽填充。此外,干预模板还可以设置为只包括下文或同时包括上下文,具体根据实际需要进行灵活配置。干预模板也可以通过“[w:min-max]”实现跳词,表示可以将从min到max数量的字符进行实体识别干预处理,但干预模板中不允许两组[w:min-max]连续出现;此外,“{A[w:1-4]}”、“{A[slot]}”,可以将A和[slot]共同组成词槽,即进行词槽填充或不进行词槽填充。

[0107] 在具体实现时,正向干预模板和负向干预模板具体的配置格式和规则可以相同,但对应的配置结果不同,命中正向干预模板时为识别出对应字段,而命中负向干预模板时不为识别出对应的字段。

[0108] 在一个实施例中,通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果,包括:查询与业务流程对应预设的正向干预模板和负向干预模板;将负向干预模板和正向干预模板分别与待干预语句的各字段进行匹配,得到负向匹配结果和正向匹配结果;当负向匹配结果和正向匹配结果均为匹配成功、且待干预语句中与负向干预模板匹配成功的字段和待干预语句中与正向干预模板匹配成功的字段为相同字段时,确定语句的命名实体识别结果中不包括相同字段。

[0109] 其中,干预模板同时包括正向干预模板和负向干预模板,分别通过正向干预模板和负向干预模板在待干预语句中进行匹配,并根据匹配结果确定语句对应的命名实体识别结果。

[0110] 具体地,服务器查询与业务流程对应预设的正向干预模板和负向干预模板,将负向干预模板和正向干预模板分别与待干预语句的各字段进行匹配,具体可以为进行字符匹配,得到负向匹配结果和正向匹配结果。当负向匹配结果和正向匹配结果均为匹配成功,表明语句同时命中了负向干预模板和正向干预模板,服务器确定待干预语句中与负向干预模板匹配成功的字段和待干预语句中与正向干预模板匹配成功的字段是否为相同字段,即语句同时命中了负向干预模板和正向干预模板的字段是否为同一字段,若是,则确定语句的命名实体识别结果中不包括相同字段。即在语句同一字段同时命中正向干预模板和负向干预模板时,负向干预模板的优先级高于正向干预模板,确定语句的命名实体识别结果中不包括相同字段,从而将相同字段进行排除。

[0111] 本实施例中,在语句同时命中正向干预模板和负向干预模板时,确定语句的命名实体识别结果中不包括相同字段,将相同字段进行排除,使将相同字段进行排除,可以避免正向干预模板和负向干预模板之间的冲突,确保命名实体识别结果的准确性。

[0112] 在一个实施例中,命名实体识别方法还包括:确定业务流程对应的待填充词槽;将命名实体识别结果填充至待填充词槽中,得到业务流程对应的业务词槽;当业务词槽满足流程执行条件时,基于业务词槽执行业务流程。

[0113] 其中,待填充词槽将用户意图转化为明确的业务流程指令所需要补充的信息,如预订机票的业务流程中,待填充词槽可以包括出发地点、出发时间和到达地点。业务词槽为根据命名实体识别结果对待填充词槽进行填充后得到的信息。流程执行条件用于触发业务流程的执行,如预订机票的业务流程中,出发地点、出发时间和到达地点等三个词槽均填充

完成,且不冲突时,如出发地点与到达地点不同,此时可以执行预订机票的业务流程,则满足流程执行条件,触发业务流程的执行,进行机票预订处理。

[0114] 具体地,得到语句的命名实体识别结果后,服务器确定业务流程对应的待填充词槽,不同的业务流程对应于不同的待填充词槽,对业务流程对应的待填充词槽进行填充后,满足流程执行条件时,可以触发业务流程的执行。具体由服务器将命名实体识别结果填充至待填充词槽中,得到业务流程对应的业务词槽,服务器根据业务词槽确定是否满足业务流程对应的流程执行条件,若是,则触发业务流程的执行,服务器基于业务词槽执行业务流程。在具体应用时,不同的业务流程对应于不同的待填充词槽,具有不同的流程执行条件,根据实际需要进行预先配置。

[0115] 本实施例中,根据语句的命名实体识别结果对业务流程对应的待填充词槽进行填充,从而执行相应的业务流程,可以通过识别到的准确的命名实体识别结果进行词槽填充,能够确保业务流程执行的准确性。

[0116] 应该理解的是,虽然图2-4的流程图中的各个步骤按照箭头的指示依次显示,但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明,这些步骤的执行并没有严格的顺序限制,这些步骤可以以其它的顺序执行。而且,图2-4中的至少一部分步骤可以包括多个步骤或者多个阶段,这些步骤或者阶段并不必然是在同一时刻执行完成,而是可以在不同的时刻执行,这些步骤或者阶段的执行顺序也不必然是依次进行,而是可以与其它步骤或者其它步骤中的步骤或者阶段的至少一部分轮流或者交替地执行。

[0117] 在一个实施例中,如图5所示,提供了一种命名实体识别装置500,包括:语句获取模块502、召回模板查询模块504、待干预语句确定模块506和识别干预模块508,其中:

[0118] 语句获取模块502,用于获取待识别的语句,并确定语句对应的业务流程;

[0119] 召回模板查询模块504,用于查询与业务流程对应预设的至少一个召回模板;

[0120] 待干预语句确定模块506,用于将语句与至少一个召回模板进行匹配,根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句;

[0121] 识别干预模块508,用于通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果。

[0122] 在一个实施例中,待干预语句确定模块506包括特征字段确定模块、字符匹配模块和目标召回模板模块;其中:特征字段确定模块,用于确定至少一个召回模板中的召回特征字段;字符匹配模块,用于将召回特征字段在语句中进行字符匹配,当得到的字符匹配结果为匹配成功时,确定匹配成功的目标召回模板;目标召回模板模块,用于基于目标召回模板从语句中确定待干预语句。

[0123] 在一个实施例中,目标召回模板模块包括待干预槽位确定模块、语句划分模块和划分结果处理模块;其中:待干预槽位确定模块,用于确定目标召回模板中的待干预槽位;语句划分模块,用于根据召回特征字段和待干预槽位将语句进行语句划分,得到语句划分结果;划分结果处理模块,用于将语句划分结果中与待干预槽位对应的字段确定为待干预语句。

[0124] 在一个实施例中,识别干预模块508包括正向模板查询模块、正向匹配模块和正向干预模块;其中:正向模板查询模块,用于查询与业务流程对应预设的正向干预模板;正向匹配模块,用于将正向干预模板与待干预语句的各字段进行正向匹配,得到正向匹配结果;

正向干预模块,用于当正向匹配结果为匹配成功时,将待干预语句中与正向干预模板匹配成功的字段,确定为语句的命名实体识别结果。

[0125] 在一个实施例中,识别干预模块508包括负向模板查询模块、负向匹配模块和负向干预模块;其中:负向模板查询模块,用于查询与业务流程对应预设的负向干预模板;负向匹配模块,用于将负向干预模板与待干预语句的各字段进行负向匹配,得到负向匹配结果;负向干预模块,用于当负向匹配结果为匹配成功时,确定语句的命名实体识别结果中不包括待干预语句中与负向干预模板匹配成功的字段。

[0126] 在一个实施例中,识别干预模块508包括干预模块查询模块、干预模板匹配模块和干预处理模块;其中:干预模块查询模块,用于查询与业务流程对应预设的正向干预模板和负向干预模板;干预模板匹配模块,用于将负向干预模板和正向干预模板分别与待干预语句的各字段进行匹配,得到负向匹配结果和正向匹配结果;干预处理模块,用于当负向匹配结果和正向匹配结果均为匹配成功、且待干预语句中与负向干预模板匹配成功的字段和待干预语句中与正向干预模板匹配成功的字段为相同字段时,确定语句的命名实体识别结果中不包括相同字段。

[0127] 在一个实施例中,还包括词槽确定模块、词槽填充模块和业务处理模块;其中:词槽确定模块,用于确定业务流程对应的待填充词槽;词槽填充模块,用于将命名实体识别结果填充至待填充词槽中,得到业务流程对应的业务词槽;业务处理模块,用于当业务词槽满足流程执行条件时,基于业务词槽执行业务流程。

[0128] 关于命名实体识别装置的具体限定可以参见上文中对于命名实体识别方法的限定,在此不再赘述。上述命名实体识别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。

[0129] 在一个实施例中,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6所示。该计算机设备包括通过系统总线连接的处理器、存储器和网络接口。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、计算机程序和数据库。该内存储器为非易失性存储介质中的操作系统和计算机程序的运行提供环境。该计算机设备的数据库用于存储数据。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。该计算机程序被处理器执行时以实现一种命名实体识别方法。

[0130] 本领域技术人员可以理解,图6中示出的结构,仅仅是与本申请方案相关的部分结构的框图,并不构成对本申请方案所应用于其上的计算机设备的限定,具体的计算机设备可以包括比图中所示更多或更少的部件,或者组合某些部件,或者具有不同的部件布置。

[0131] 在一个实施例中,提供了一种计算机设备,包括存储器和处理器,存储器中存储有计算机程序,该处理器执行计算机程序时实现以下步骤:

[0132] 获取待识别的语句,并确定语句对应的业务流程;

[0133] 查询与业务流程对应预设的至少一个召回模板;

[0134] 将语句与至少一个召回模板进行匹配,根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句;

[0135] 通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果。

[0136] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:确定至少一个召回模板中的召回特征字段;将召回特征字段在语句中进行字符匹配,当得到的字符匹配结果为匹配成功时,确定匹配成功的目标召回模板;基于目标召回模板从语句中确定待干预语句。

[0137] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:确定目标召回模板中的待干预槽位;根据召回特征字段和待干预槽位将语句进行语句划分,得到语句划分结果;将语句划分结果中与待干预槽位对应的字段确定为待干预语句。

[0138] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:查询与业务流程对应预设的正向干预模板;将正向干预模板与待干预语句的各字段进行正向匹配,得到正向匹配结果;当正向匹配结果为匹配成功时,将待干预语句中与正向干预模板匹配成功的字段,确定为语句的命名实体识别结果。

[0139] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:查询与业务流程对应预设的负向干预模板;将负向干预模板与待干预语句的各字段进行负向匹配,得到负向匹配结果;当负向匹配结果为匹配成功时,确定语句的命名实体识别结果中不包括待干预语句中与负向干预模板匹配成功的字段。

[0140] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:查询与业务流程对应预设的正向干预模板和负向干预模板;将负向干预模板和正向干预模板分别与待干预语句的各字段进行匹配,得到负向匹配结果和正向匹配结果;当负向匹配结果和正向匹配结果均为匹配成功、且待干预语句中与负向干预模板匹配成功的字段和待干预语句中与正向干预模板匹配成功的字段为相同字段时,确定语句的命名实体识别结果中不包括相同字段。

[0141] 在一个实施例中,处理器执行计算机程序时还实现以下步骤:确定业务流程对应的待填充词槽;将命名实体识别结果填充至待填充词槽中,得到业务流程对应的业务词槽;当业务词槽满足流程执行条件时,基于业务词槽执行业务流程。

[0142] 在一个实施例中,提供了一种计算机可读存储介质,其上存储有计算机程序,计算机程序被处理器执行时实现以下步骤:

[0143] 获取待识别的语句,并确定语句对应的业务流程;

[0144] 查询与业务流程对应预设的至少一个召回模板;

[0145] 将语句与至少一个召回模板进行匹配,根据匹配结果确定在语句中与至少一个召回模板匹配成功的待干预语句;

[0146] 通过与业务流程对应预设的干预模板,对待干预语句进行实体识别干预,根据实体识别干预结果获得语句的命名实体识别结果。

[0147] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:确定至少一个召回模板中的召回特征字段;将召回特征字段在语句中进行字符匹配,当得到的字符匹配结果为匹配成功时,确定匹配成功的目标召回模板;基于目标召回模板从语句中确定待干预语句。

[0148] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:确定目标召回模

板中的待干预槽位;根据召回特征字段和待干预槽位将语句进行语句划分,得到语句划分结果;将语句划分结果中与待干预槽位对应的字段确定为待干预语句。

[0149] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:查询与业务流程对应预设的正向干预模板;将正向干预模板与待干预语句的各字段进行正向匹配,得到正向匹配结果;当正向匹配结果为匹配成功时,将待干预语句中与正向干预模板匹配成功的字段,确定为语句的命名实体识别结果。

[0150] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:查询与业务流程对应预设的负向干预模板;将负向干预模板与待干预语句的各字段进行负向匹配,得到负向匹配结果;当负向匹配结果为匹配成功时,确定语句的命名实体识别结果中不包括待干预语句中与负向干预模板匹配成功的字段。

[0151] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:查询与业务流程对应预设的正向干预模板和负向干预模板;将负向干预模板和正向干预模板分别与待干预语句的各字段进行匹配,得到负向匹配结果和正向匹配结果;当负向匹配结果和正向匹配结果均为匹配成功、且待干预语句中与负向干预模板匹配成功的字段和待干预语句中与正向干预模板匹配成功的字段为相同字段时,确定语句的命名实体识别结果中不包括相同字段。

[0152] 在一个实施例中,计算机程序被处理器执行时还实现以下步骤:确定业务流程对应的待填充词槽;将命名实体识别结果填充至待填充词槽中,得到业务流程对应的业务词槽;当业务词槽满足流程执行条件时,基于业务词槽执行业务流程。

[0153] 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,是可以通过计算机程序来指令相关的硬件来完成,所述的计算机程序可存储于一非易失性计算机可读存储介质中,该计算机程序在执行时,可包括如上述各方法的实施例的流程。其中,本申请所提供的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易失性存储器中的至少一种。非易失性存储器可包括只读存储器(Read-Only Memory,ROM)、磁带、软盘、闪存或光存储器等。易失性存储器可包括随机存取存储器(Random Access Memory,RAM)或外部高速缓冲存储器。作为说明而非局限,RAM可以是多种形式,比如静态随机存取存储器(Static Random Access Memory,SRAM)或动态随机存取存储器(Dynamic Random Access Memory,DRAM)等。

[0154] 以上实施例的各技术特征可以进行任意的组合,为使描述简洁,未对上述实施例中的各个技术特征所有可能的组合都进行描述,然而,只要这些技术特征的组合不存在矛盾,都应当认为是本说明书记载的范围。

[0155] 以上所述实施例仅表达了本申请的几种实施方式,其描述较为具体和详细,但并不能因此而理解为对发明专利范围的限制。应当指出的是,对于本领域的普通技术人员来说,在不脱离本申请构思的前提下,还可以做出若干变形和改进,这些都属于本申请的保护范围。因此,本申请专利的保护范围应以所附权利要求为准。

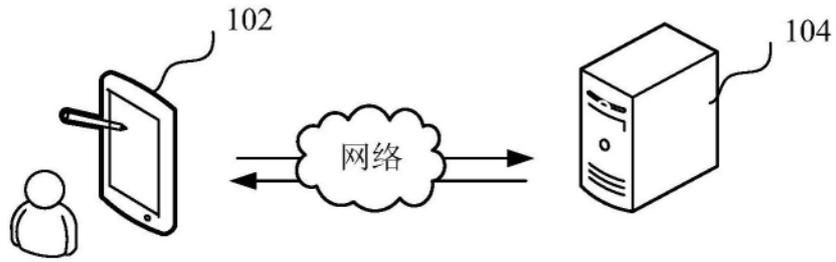


图1

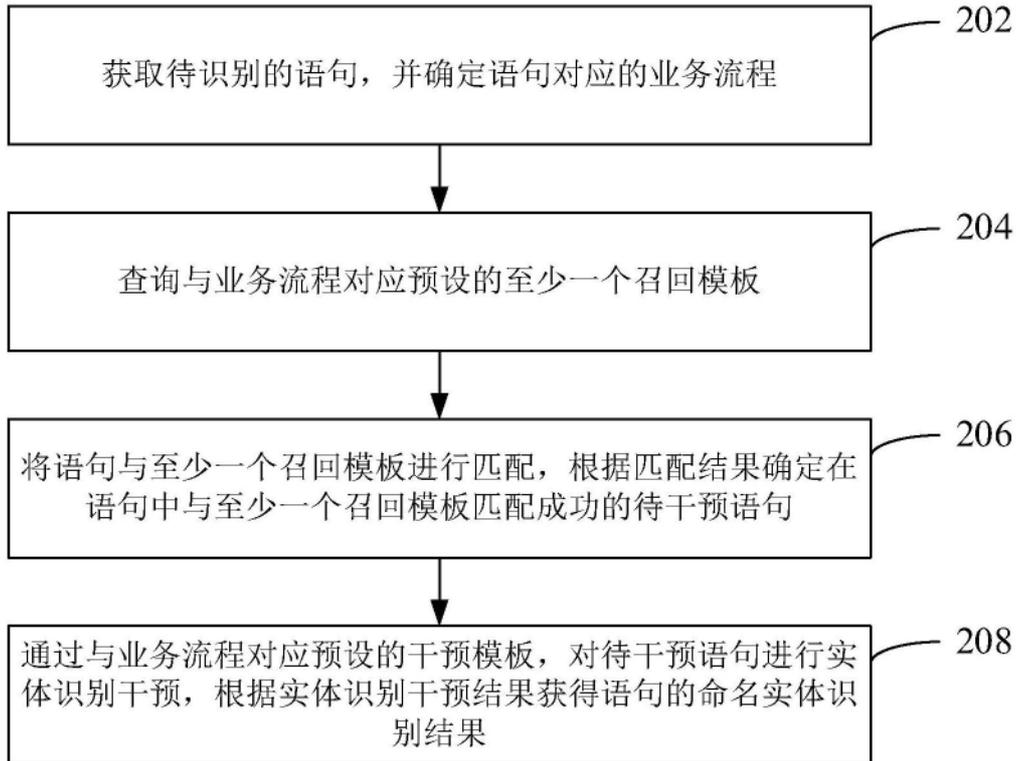


图2

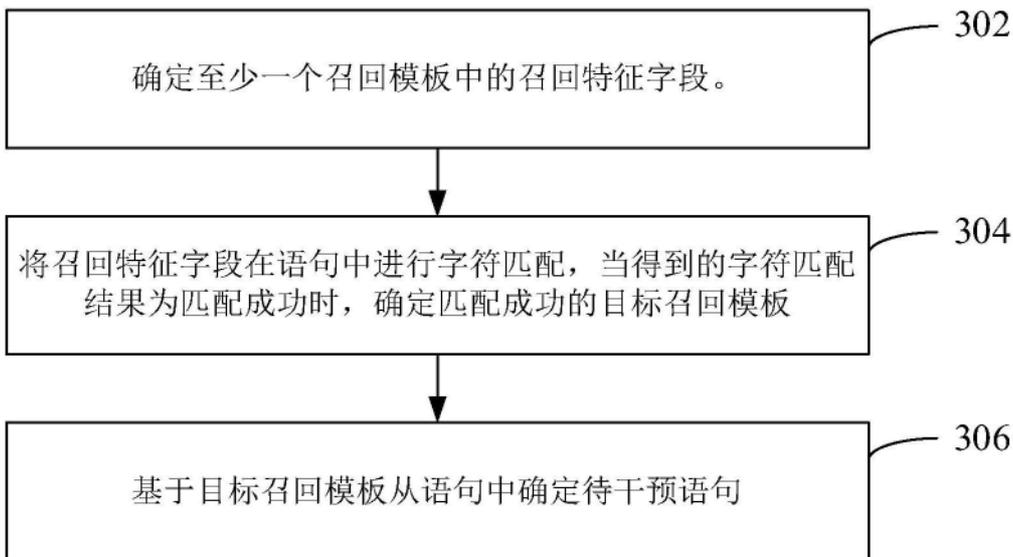


图3

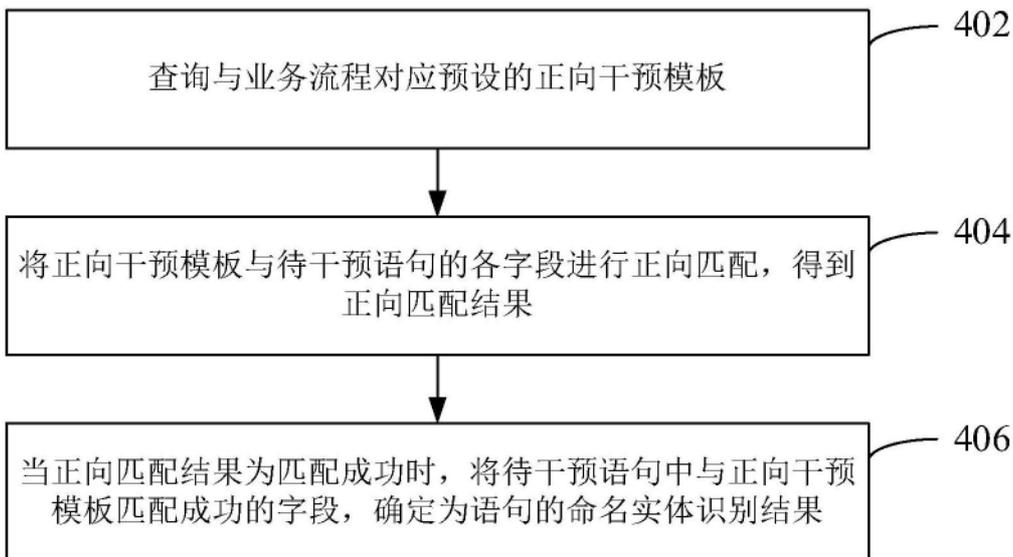


图4

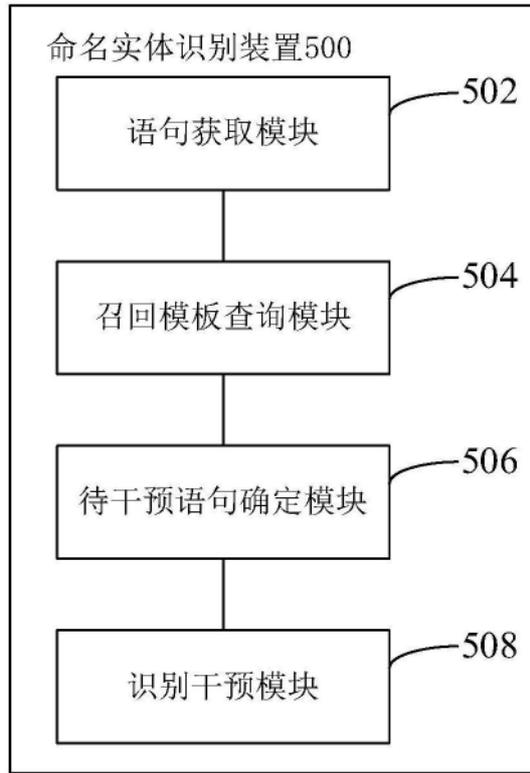


图5

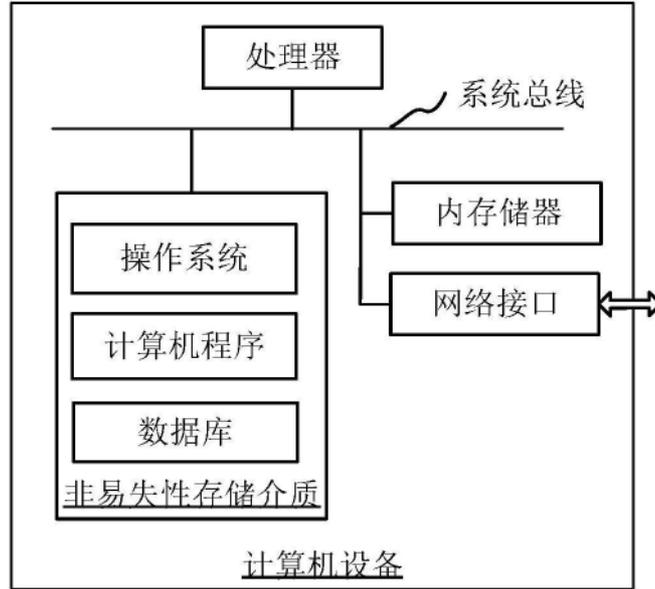


图6