

一种面向室内智能机器人导航的路径自然语言处理方法

李新德¹ 张秀龙¹

摘 要 为了实现通过自然语言指导机器人导航的目的, 提出了一种利用描述路径的自然语言直接绘出机器人运行路径的方法. 为了提高机器人对路径自然语言的理解, 采用了基于组块的分析方法. 首先深入地研究了收集到的路径描述的自然语言语料库, 总结出了语义和句法的对应关系. 在此基础上, 构造了层叠条件随机场, 通过采用组块分析方法, 先进行名词短语组块的提取, 然后提取语义组块. 为了解决名词短语的解析问题, 引入了名词实体关系推理方法. 并且在提取出的语义组块的基础上逐个建立路径单元, 组成完整的路径信息. 最后, 给出了自然语言处理各个环节的实验结果, 并结合具体语句展示了从文字到机器人导航地图的映射过程, 对开展人机交互式和谐导航具有重要意义.

关键词 自然语言处理, 机器人导航, 层叠条件随机场, 贝叶斯推理

引用格式 李新德, 张秀龙. 一种面向室内智能机器人导航的路径自然语言处理方法. 自动化学报, 2014, 40(2): 289–305

DOI 10.3724/SP.J.1004.2014.00289

A Route Instruction Method Using Natural Language Processing for Indoor Intelligent Robot Navigation

LI Xin-De¹ ZHANG Xiu-Long¹

Abstract In this paper, a new route instruction method using natural language processing was investigated to guide mobile robot navigation. To gain a clear understanding of the route expressed by natural language, a chunk parsing method was adopted. Firstly, the collected corpus of the route expressed by natural language was deeply studied and the corresponding relationship between the sentence structure and semantics was presented. Based on this, cascaded conditional random fields were constructed and the noun phrase chunk is picked up and, its semantics chunk was also picked up by applying the chunk parsing method. To resolve the meaning of the noun phrases, a noun entities relationship reasoning method was presented. The units of the route was built with the picked-up semantic chunks and an integral route was finally completed. At last, the experimental results of every loop were presented. Some example sentences of the route expressed by natural language were given and processed to show the procedure from the sentences to map used in the robot navigation. It is very significant for robot navigation with harmony based on human-robot interaction to carry out this research.

Key words Natural language processing, robot navigation, cascaded conditional random fields, Bayesian inference

Citation Li Xin-De, Zhang Xiu-Long. A route instructions method using natural language processing for indoor intelligent robot navigation. *Acta Automatica Sinica*, 2014, 40(2): 289–305

在苹果公司推出的手机 iPhone 4s/5, 以及 iPad 3 中, 包含 Siri 系统, 其可以高度智能地完成与通话

相关的操作, 比如, 人们通过日常用语就能够控制手机自动完成打电话、查号码等操作, 甚至能够完成诸如查找音乐、网上搜索等功能. Siri 系统强大的自然语言理解系统给人印象深刻, 如果能用自然语言控制移动机器人, 进行和谐导航, 对实现智能机器人成为家庭普通成员具有重要意义. 于是本文提出一种用自然语言直接生成机器人前进路径的方法来实现机器人导航.

在基于汉语描述路径的自然语言 (简称路径自然语言) 处理方面, 相关研究工作还比较滞后, 目前还没有研究者提出基于完全自然语言的处理方法. 聂仙丽等^[1] 虽然提出过将自然语言处理用于机器人导航, 但是更多的是采用编程语言处理特定对话, 没有真正涉及到自然语言处理. 尽管如此, 其他领域中

收稿日期 2012-12-19 录用日期 2013-3-27
Manuscript received December 19, 2012; accepted March 27, 2013

国家自然科学基金 (61175091), 江苏省自然科学基金 (BK2010403), 东南大学优秀青年教师教学、科研资助计划 (3208001203), 江苏省“青蓝工程”资助计划资助

Supported by National Natural Science Foundation of China (61175091), Natural Science Foundation of Jiangsu Province (BK2010403), Excellent Young Teacher Foundation of Southeast University in Teaching and Science Research (3208001203), and Qing Lan Project

本文责任编辑 宗成庆
Recommended by Associate Editor ZONG Cheng-Qing

1. 东南大学自动化学院复杂工程测量与控制教育部重点实验室 南京 210096

1. Key Laboratory of Measurement and Control of CSE (School of Automation, Southeast University), Ministry of Education, Nanjing 210096

和路径相关的研究有一定借鉴作用, 因为路径自然语言中确定方位是基本的要求, 这和地理信息系统 (Geographic information system, GIS) 领域的部分工作比较相似. 乐小虬等^[2]采用基于分层有限状态自动机的方法来分析自然语言中的方位关系; 张雪英等^[3]从城市道路导航角度出发, 研究了汉语中路径语言的基本语法, 在此基础上提出了处理方法. 但实际上这样的处理更多的是受限自然语言处理, 而不是完全自然语言处理. 在基于英语的路径自然语言处理方面, Klippel 等^[4]在路径自然语言语义表达规律的基础上, 很详细地分析了路径的组成、路径中的方向和距离、路标的作用、路标和路径的关系等. Shimizu 等^[5]提出了 End-to-end 系统的概念, 认为用路径自然语言进行机器人导航应该分为两步, 即: 1) 路径自然语言分成若干系列运动; 2) 将这一系列的运动、地图、机器人初始方向结合. 这种系统划分被后来的研究者普遍接受. 尽管文中提出了使用组块进行分析的方法, 但是所采用的语料非常单一, 其算法只具有启发性. Yuan 等^[6]认为路径自然语言中经常有很多错误或者模糊信息, 因此只要通过环境感知方法推导出大环境的信息, 则可以通过与现有地图结合的方法进行机器人的路径推导. Kollar 等^[7]在文献 [6] 的基础上提出了 SDC (Spatial description clause) 用来分析路径自然语言中的方向, 并结合文献 [6] 的结果给出了相关的实验结果, 此后文献 [8–12] 在文献 [7] 的基础上逐步完善了该方法, 但是总体来说所用的方法依赖于环境的先验经验, 还没有达到理想的处理水平. 鲁汶大学的 Kordjamshidi 等^[13]则从更加宽广的角度研究了自然语言中物体的空间关系, 类比一般的语义角色标注方法, 他们采用了以方位词为中心的空间角色语义标注方法, 为路径自然语言提供了很好的借鉴. 在这个领域, 英文基于路径自然语言理解的机器人导航已经有了初步的发展. 总体来说, 英文的路径自然语言处理侧重于提取路标名词和方位词, 然后通过已有的地图进行路径规划, 自然语言的处理只涉及到句法层面. 自然语言处理仍然是一个辅助性的工具, 更多地依靠路径规划算法来提取路径.

本文的创新之处在于依靠自然语言处理方法获得路径自然语言的语义信息, 直接从路径自然语言中提取路线, 在此基础上机器人可以利用文献 [14] 中的方法进行导航. 因为这样的方法需要更高水平的路径自然语言处理方法, 更加智能, 既不需要依靠高精度的地图, 也不需要提前输入手绘地图.

本文第 1 节在文献 [4–5, 15] 的基础上分析了和路径自然语言中关键信息相关的句法, 提出了路径自然语言句法和句义的对应关系, 是本文提出方法的理论基础; 第 2 节和第 3 节在文献 [6, 16] 的基

础上构造了层叠条件随机场用来进行语义组块的分析, 然后分析了路径自然语言中名词短语内的名词实体的相互关系, 是本文具体的路径自然语言处理方法; 第 4 节在前面处理的基础上逐个提取路径单元, 并将其组合成运行路线; 第 5 节展示了本文提出方法的实验结果, 并结合具体语句给出了处理的过程.

1 句法和句义对应关系

1.1 语料的句法结构分析

通过志愿者调查问卷的形式收集语料, 调查分两次完成, 调查的对象是东南大学自动化学院大一到大三的学生. 在进行语料收集时, 调查人员给出具体的地图和地图中标出的路径^[16], 假设有人需要从地图中的起点走到终点, 但是不知怎么走. 要求被调查人员用自己日常的说说话习惯为其指路, 并将所说的内容记录在调查问卷上. 调查过程中调查人员不能给出具有倾向性的指导, 力求被调查者完全根据个人用语习惯描述. 在获得语料之后, 为了保证语料能反映被调查者真实的用语习惯, 只去除了重复和明显不符合日常用语习惯的语句, 还剩下 216 句.

总体来说, 收集到的句子的句法都比较简单, 大多数的语料只包含关键的信息, 例如方向、明显的标志物等. 但是对标志物的属性描述较少, 例如物体的大小、颜色等描述就很少提及. 比较典型的句子有:

1) “向前, 到椅子那右拐, 雨伞处右拐, 向前在吉他处左拐, 向前到达盆景柜”;

2) “直行到椅子前, 右转 90 度, 走过 L 型柜子, 向左斜前方向走到吉他与一字柜之间, 右转, 并越过一字柜再左转, 从中间空地直接走向盆景柜”;

3) “向前右拐绕过 L 型柜子, 向前再绕过一字柜, 向左前进到盆景树”.

对这些语料的有效信息进行了分析, 分析的内容主要包括: 运动方向、运动形式、路标及其相对方位等. 为了发现路径自然语言中的相关规律, 将与有效信息相关的短语单独列出进行了语法分析.

在语料库中和路径中的参考物相关的短语如例 1 所示.

例 1. 1) “到达柜子边缘”; 2) “到达初始点与椅子的中点”; 3) “在盆景前”.

用上下文无关语法 (Content-free grammar, CFG) 可表示为 (其中 S 表示句子, PP 为介词短语, NP 为名词短语, VP 为动词短语, MP 为数量词短语, n 为名词, v 为动词, f 为方位词, pf 为介词, m 表示数词, q 表示量词, udel 表示 “的”, 下文采用同样的表示):

1)

S → VP
 VP → v NP n → 柜子 — 边缘
 NP → n n v → 到达

2)

S → VP
 VP → v NP
 NP → NP udel NP
 NP → NP n
 NP → n
 n → 初始点 — 椅子 — 中点
 v → 到达

3)

S → PP f → 前
 PP → pf NP pf → 在
 NP → n f n → 盆景

以上三个例子是路径自然语言中比较常见的语句, 非常有代表性, 不难发现虽然各个短语不属于同一类(可能是介词短语, 或者是动词短语), 但是每个例子的第二条规则的重写部分都存在一个名词短语, 而名词短语之前存在介词或者动词, 名词短语包含着参考物体的信息. 在语料库中和参考物相关的短语基本都存在类似的句法结构.

在汉语路径自然语言中, 运动方向往往和动作相关, 而动作可以独立存在. 语料库中比较典型的直接表示运动方向和动作的短语结构如例 2 所示.

例 2. 1) “右转 90 度”; 2) “向左转”; 3) “向东 4 米”; 4) “直走”; 5) “向前走 5 米”.

可以用 CFG 表示为:

1)

S → VP f → 右
 VP → VP MP v → 转
 VP → f v m → 90
 MP → m q q → 度

2)

S → PP f → 左
 PP → pf VP v → 转
 VP → f v pf → 向

3)

S → PP f → 东
 PP → PP MP m → 4
 PP → pf f q → 米
 MP → m q pf → 向

4)

S → VP d → 直
 VP → d v v → 走

5)

S → PP f → 前

PP → PP VP v → 走
 VP → v MP pf → 向
 PP → pf f m → 5
 MP → m q q → 米

例 2 中 1)~4) 的句法结构差异较大, 但在句义层面上, 各个短语都表达了运动的方向和运动的形式; 5) 是日常用语中比较完整的表达, 前面四个短语都有 5) 中的部分句法结构, 可以认为是 5) 的一种缺省表达形式. 从上述的重写规则中不难发现 1)~4) 与 5) 有很多相似的句法结构, 都有介词短语、数量词短语、动词短语等其中的一个或多个, 且各个短语位置关系相对固定.

在语料库中还有一类用间接的方式表达运动方向和形式的短语, 如例 3 所示.

例 3. 1) 沿走廊走到主卧; 2) 朝盆景往前走; 3) 从客厅向着盆景走.

例 3 用 CFG 可以表示为:

1)

S → PP
 PP → PP VP pf → 沿
 n → 走廊|主卧
 PP → pf n v → 走到
 VP → v n

2)

S → PP
 PP → PP PP pf → 朝|往
 n → 盆景
 PP → pf n v → 走
 PP → pf VP f → 前
 VP → f v

3)

S → PP
 PP → PP PP pf → 从|向着
 n → 客厅|盆景
 PP → pf n v → 走
 PP → PP v

在例 3 中, 运动的方向不再用直接的方式表示出来, 而是用参考物体间接地表示运动方向. 在各个短语的重写规则中, 虽然有一部分句法成分不相同, 但是总体来说仍具有一定的相似性, 即短语本身都是介词短语(重写规则的第一条); CFG 的表达式中有很多相同的非终结符表明它们内部结构比较相似; 同时, 各个成分的相对位置关系也具有相似性, 比如动词是名词的后缀, 动词也是介词的后缀等.

以上所列出的三个例子是语料库中有效信息的主要表达方式, 当然这也符合我们的基本常识, 日常

交流中也存在大量这样的现象. 除去以上三个例子中所列举的短语信息, 语料库中还有其他的表示运动方向、物体方位的短语, 例如: “前方为雨伞” 这样的描述语句本身不改变运动的方向, 只是为路径提供了更多的参考, 而且这样的语句在语料库中所占比例比较低 (占 2.3% 左右), 在本文中暂时不予讨论. 通过以上分析可以发现, 语料库中和运动方向、运动形式、参考物体相关的短语往往具有相似的语法结构, 主要表现在构成短语的词性组合类似. 从另一方面来说, 这样的特点也是我们分辨出一个句子是否是路径自然语言的原因.

1.2 语义和句法结构的关系

由于自然语言表达的随意性, 一般来讲自然语言中句法结构和语义信息没有必然的关系. 例如 “开汽水” 和 “开汽车” 的句法结构都可以用重写规则 $VP \rightarrow v \ n$ 来表示, 且终结符也相似, 但是它们的意义完全不同. 在路径自然语言中, 表达方式和词汇的高度集中与表达方式的相对较少两者之间存在一定的关系.

从例 1 可以发现, 重写规则中的第二句都含有名词短语, 分别表示 “柜子边缘”、“初始点与椅子的中点” 和 “盆景前”, 这些名词短语语义上都是表示参考物. 这些名词的前缀即 “到达”, “在” 等分别表示动作和状态的词, 因此可以发现一种语义和句法相对应的关系, 即在路径自然语言中, 若有某个短语有类似于上面所示的句法结构, 则可以认为这样的短语包含了参考物体、运动形式或运动状态的语义信息. 我们定义例 1 所示的短语的句法结构为 SP1, 则 SP1 可以用重写规则表示为

$$SP1 \rightarrow PP_{SP1}/VP_{SP1} \ NP_{SP1}$$

其中, PP_{SP1} 、 VP_{SP1} 不含有名词成分. 定义 ACTION 表示语义层次上的动作, STATE 表示运动的状态, LANDMARK 表示和路标相关的信息, 则可以得到如下表达式:

$$NP_{SP1} \rightarrow LANDMARK \quad (1)$$

$$PP_{SP1} \rightarrow STATE \quad (2)$$

$$VP_{SP1} \rightarrow ACTION \quad (3)$$

注. 这里的表达式不是重写规则, 而是表示映射关系, 下文亦同.

式 (1) 表示如果能够在路径自然语言中获得 SP1 所示的句法结构, 则由名词短语可以推导得到参考物; 式 (2) 表示由介词短语可以获得运动的状态; 式 (3) 表示由动词短语可以得到运动的形式.

从句法角度分析可知例 2 中 1)~4) 都是 5) 的缺省表示, 由于人具有的语言常识和推导能力, 正常情况下, 仍然能够自动补全如 5) 所示的句法结构, 并正确理解. 比如 4) 中的运动方向一般都默认为 “前”, 运动方式都是 “走”, 自动补全之后的语句是: “直向前走”, 而行走的距离则通过常识进行控制. 因此我们可以看到 1)~4) 虽然句法结构各异, 但是在语义表达上和 5) 具有一致性, 而 5) 中的语义比较明确. 设类似于 5) 的短语的句法结构为 SP2, 则 SP2 可以表示为

$$SP2 \rightarrow PP_{SP2} \ VP_{SP2} \ MP_{SP2}$$

其中, PP_{SP2} 后缀中不包含动词成分. 定义 DIRECTION 表示运动方向, DISTANCE 表示运动距离, 则可以得到:

$$PP_{SP2} \rightarrow DIRECTION \quad (4)$$

$$VP_{SP2} \rightarrow ACTION \quad (5)$$

$$MP_{SP2} \rightarrow DISTANCE/DIRECTION \quad (6)$$

式 (4) 表示如果能够在路径自然语言中获得 SP2 所示的句法结构, 则从介词结构中可以得到方向信息; 式 (5) 表示从动词短语中可以得到运动的状态, 而且这样的动作是路径自然语言中的有效信息, 运动的状态可以由介词短语和数量词短语表示; 式 (6) 表示可以从数量词短语中获得距离信息或者方向信息.

参考对例 1 和例 2 的分析, 例 3 中句法结构所对应的语义角色也比较明晰. 设类似例 3 中所示的句法结构为 SP3. 用重写规则可以表示为

$$SP3 \rightarrow PP_{SP3} \ PP_{SP3}/VP_{SP3}$$

则可以得到:

$$PP_{SP3} \rightarrow LANDMARK/STATE \quad (7)$$

$$VP_{SP3} \rightarrow ACTION/LANDMARK \quad (8)$$

式 (7) 表示如果在路径自然语言中获得 SP3 所示的句法结构, 则认为可以由介词短语获得运动的状态或者参考物的信息; 式 (8) 表示可以从动词短语中得到参考物信息或者运动形式信息.

2 语义组块的提出

通过以上分析, 认为可以从某些特定语法结构中直接提取句义. 目前通用的浅层语义分析方法一般都是采用角色标注的方法进行处理. 通过获得句法树或者依赖关系提取特征, 然后通过谓词分析获得句子中的各种语义角色. 比较 “典型” 的就是基于

格语法的句义分析,但是目前的句义分析准确率普遍不高^[6].如果沿用通用的语义角色提取方法,鉴于我们的语料库规模非常小,将带来很大的数据稀疏问题,效果将非常有限.因此在以上分析的基础上,借鉴文献 [5, 17] 中的方法,使用组块分析的方法进行处理.在文献 [16] 中,定义了 13 个语义组块,通过实验发现过多的语义组块不但不能达到覆盖所有成分的目的,反而增加了语义组块之间的误判,因此本文中缩减为 7 种语义组块.

由句法结构 SP1 可得:

1) NL (Noun landmark, 名词性路标), 例如“L 型柜子”;

2) VL (Verb landmark, 动词性路标), 例如“走到墙角”;

3) PL (Preposition landmark, 介词性路标), 例如“在柜子旁边”;

由句法结构 SP2 可得:

4) DTM (Direction transformation module, 方位转换模块).

由句法结构 SP3 可得:

5) IDTM (基于参照物的方位转换模块或者间接方位转化模块).

剩余的语法结构为:

6) VP (Verb phrase, 与方位无关的动词短语), 例如“继续走”;

7) ROBO (这里指路径执行主体, Robot), 例如“你”.

由于语义组块和句法结构有直接的对应关系,因此可以采用组块分析^[17]的方法来处理路径自然语言的句义分析问题.本文中的组块处理方法属于序列信号的处理问题,该类问题目前有较多的方法,如文献 [18] 中所用的方法.但是考虑到算法的稳定性和构建层叠系统的需要,本文采用层叠随机条件场 (Cascaded conditional random fields, CCRFs) 来进行组块的提取.由于部分语义组块的句法结构和文献 [15] 中的模块相似,这里仍然采用相似的名称.

2.1 条件随机场

条件随机场 (Conditional random fields, CRFs) 模型是一种无向图模型,已经广泛应用于自然语言处理中分词系统、名词短语识别、地名识别等领域.设 $\mathbf{y} = (y_1, y_2, \dots, y_n)$ 表示状态值,在本文中表示各种语义组块; $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 表示观察值序列,在本文表示路径自然语言的输入句子.则 CRF 可以表示为

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{Z(\mathbf{x})} \prod_{j=1}^n \Psi_j(\mathbf{x}, \mathbf{y}) \quad (9)$$

其中,

$$Z(\mathbf{x}) = \sum_{\mathbf{y}'} \prod_{j=1}^n \Psi_j(\mathbf{x}, \mathbf{y}') \quad (10)$$

在本文中采用线性结构的 CRF, 因此:

$$\Psi_j(\mathbf{x}, \mathbf{y}) = \exp \left(\sum_{i=1}^m \lambda_i f_i(y_{j-1}, y_j, \mathbf{x}, j) \right) \quad (11)$$

故可以得到:

$$p(\mathbf{y}|\mathbf{x}) = \frac{1}{Z_{\lambda}(\mathbf{x})} \cdot \exp \left(\sum_{j=1}^n \sum_{i=1}^m \lambda_i f_i(y_{j-1}, y_j, \mathbf{x}, j) \right) \quad (12)$$

$$Z_{\lambda}(\mathbf{x}) = \sum_{\mathbf{y} \in \gamma} \exp \left(\sum_{j=1}^n \sum_{i=1}^m \lambda_i f_i(y_{j-1}, y_j, \mathbf{x}, j) \right) \quad (13)$$

其中, λ_i 为特征 f_i 的参数, γ 表示状态量可能的序列.

2.2 层叠条件随机场

文献 [16] 中采用支持向量机 (Support vector machine, SVM) 对提出的各种语义角色进行了处理,有较多的语义角色都嵌套了名词短语.当名词短语为单个词时,正确率比较正常;但是当名词短语比较复杂时,正确率明显下降.因此引入一个多层次的 CRF 来处理这类嵌套的问题.如图 1 所示,在本文所用的 CCRFs 模型共分三层:底层为观测值,中层条件随机场模型处理名词短语识别的问题,高层条件随机场模型处理语义组块的识别问题.高层模型

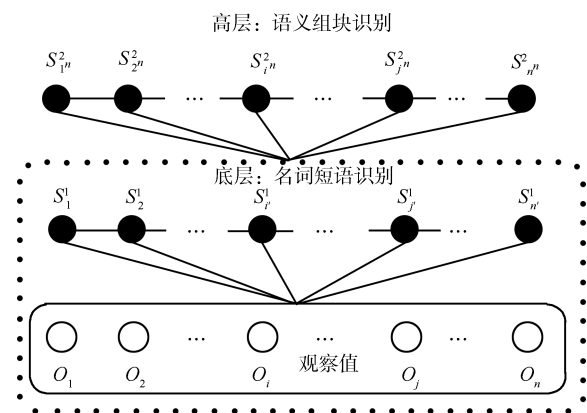


图 1 CCRFs 模型

Fig. 1 CCRFs model

的输入变量不仅包含观察值,而且包含了来自底层模型的识别结果,从而为高层条件随机场模型对比较复杂的语义组块识别提供了决策支持.各个层次

条件随机场模型的训练都是以我们建立的语料库作为训练语料。

2.3 基于 CCRFs 的语义组块识别

2.3.1 语料的标注和组块特征表示

为了满足层叠条件随机场的训练要求, 分别标注了三个语料库: 名词实体语料库、名词短语语料库、语义组块语料库。每个语料库中的句子完全一致, 只是标注的对象不同。例如, 例 4 的名词实体标注、名词短语标注、语义组块标注分别如 1) ~ 3) 所示。

例 4. “向厨房方向走, 经过客卫, 然后向右拐, 一直向前走, 走到主卧进入主卧, 然后走向主卧的斜对角, 进入书房”。

1) “向/p [厨房/n] NL 方向/n 走/v, /wd 经过/p [客/ng 卫/ng] NL 然后/c 向/p 右/f 拐/v, /wd 一直/d 向/p 前/f 走/v, /wd 走/v 到/v [主/ag 卧/n] NL 进入/v [主/ag 卧/n] NL, /wd 然后/c 走向/v [主/ag 卧/n] NL 的/ude1 斜/a [对角/n] NL, /wd 进入/v [书房/n] NL”。

2) “向/p [厨房/n] NL 方向/n 走/v, /wd 经过/p [客卫/n] NL 然后/c 向/p 右/f 拐/v, /wd 一直/d 向/p 前/f 走/v, /wd 走/v 到/v [主/ag 卧/n] NL 进入/v [主/ag 卧/n] NL, /wd 然后/c 走向/v [主/ag 卧/n 的/ude1 斜/a 对角/n] NL, /wd 进入/v [书房/n] NL”。

3) “[向/p 厨房/n 方向/n 走/v] IDTM, /wd [经过/p 名词/n] PL 然后/c [向/p 右/f 拐/v] DTM, /wd 一直/d [向/p 前/f 走/v] DTM, /wd [走/v 到/v 名词/n] VL [进入/v 名词/n] VL, /wd 然后/c [走向/v 名词/n] VL, /wd [进入/v 书房/n] VL”。

在 3) 中, 在 PL、VL 等组块中出现了“名词/n”,

这是将 2) 中的名词短语进行了替换, 但是一部分比较简单的短语没有进行替换。例如, 例 4 中的“主卧的斜对角”在 2) 中是作为一个名词短语进行标注的, 在 3) 中就被“名词/n”替换。

训练的语料将按照 IOB2 (I for inside, O for outside, B for begin) 标记方法进行标记^[8], 处理完的样本可以直接应用于训练和测试。

2.3.2 特征模板的构造

图 2 为整个路径自然语言处理的流程, 其中区域 1 为 CCRFs 处理部分, 生语料处理的具体流程是:

- 1) 对图 2 中的名词短语语料和语义组块语料进行训练, 得到对应的特征参数;
- 2) 对生语料进行词性标注;
- 3) 进行名词短语识别;
- 4) 用语言学规则对名词短语进行辅助处理;
- 5) 将所有识别出来的名词短语都进行替换;
- 6) 将所得句子用 CRF 进行语义组块分析, 得到语义组块的分析结果。

在词性标注部分, 本文直接使用中国科学院自然语言研究所的 ICTCLAS (Institute of Computing Technology; Chinese Lexical Analysis System) 系统^[19]。通过测试, 总体来讲符合使用的要求。由于我们的语料库比较小, CRF 常见的训练效率低下等问题并不突出, 因此直接使用开源的 CRF++ 程序包^[20]进行训练。

由于语料库规模较小, 训练复杂度有限, 训练时间有限, 因此使用了尽可能多的特征, 特征模板如表 1 所示。

关于 CRF 具体的训练和推导的详细过程见文献^[21]。

表 1 CRF 特征模板
Table 1 The template of CRF

特征	意义	特征	意义
W_{-2}	前第二个词汇的词形	W_{-1}	前第一个词汇的词形
W_0	当前词汇的词形	W_1	后第一个词汇的词形
W_2	后第二个词汇的词形	P_{-2}	前第二个词的词性
P_{-1}	前第一个词的词性	P_0	当前词的词性
P_1	后第一个词的词性	P_2	后第二个词的词性
W_{-1}/W_0	前一个词和当前词的词形	W_0/W_1	后一个词和当前词的词形
P_{-2}/P_{-1}	前两个词和前一个词的词性	P_{-1}/P_0	前一个词和当前词的词性
P_0/P_1	前一个词和当前词的词形	P_1/P_2	后一个词和后两个词的词形
$P_{-2}/P_{-1}/P_0$	前两个词、前一个词与当前词的词性	$P_{-1}/P_0/P_2$	前一个词、当前词与后两个词的词性
$P_0/P_1/P_2$	当前词、后一个词与后两个词的词性		

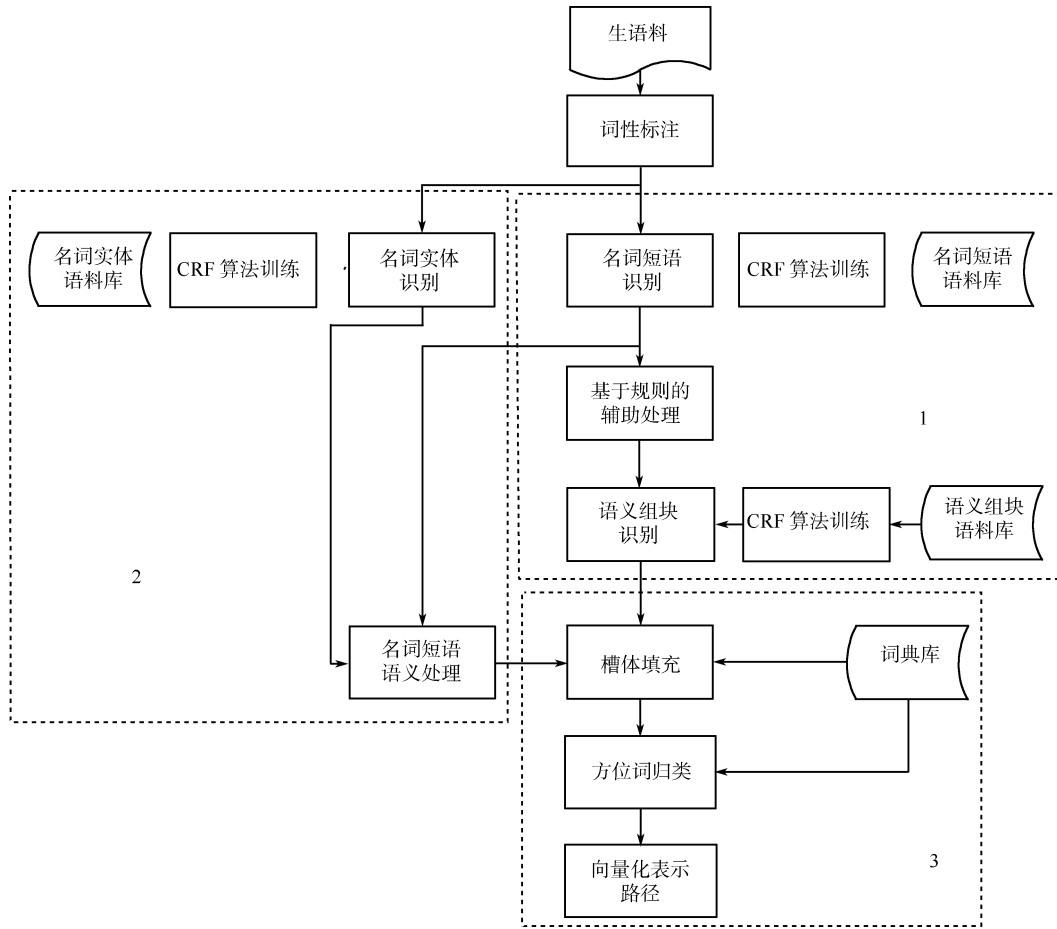


图2 整体处理流程

Fig. 2 The flowchart of global processing

在 CCRFs 中第一层提取出名词短语之后, 为了提高名词短语的正确率, 根据语言学知识对名词短语进行辅助处理. 由语言学知识可知, 形容词、数词、量词、名词都可以作为修饰成分修饰名词, 所以如果一个名词短语组块边界外存在如上所述的成分且这些成分不属于其他名词组块, 则认为这些词也属于该名词短语组块.

3 名词短语处理

名词实体识别是自然语言中一个重要的研究方向. 为了能够获得路径自然语言中的名词实体, 如图 2 中的区域 2 所示, 我们采用 CRF 模型来处理, 所使用的模板如前所示. 由式 (2) 和式 (11) 可知, 从 SP1 和 SP3 结构的名词短语中可以获得路标的相关信息, 在本文所用的语料库中大多数的名词短语都是单个名词实体, 这样的名词实体即路标.

但是剩余的名词短语中, 路标不能直接获得, 必须区分出哪些名词实体是路标, 哪些是修饰路标的名词实体. 例如, 例 4 中“主卧的斜对角”这个名词短语中有两个名词“主卧”、“斜对角”, 而路标是“斜

对角”, “主卧”是“斜对角”的修饰成分, 表明了斜对角所处的大环境; 例 1 中的第二个名词短语“初始点与椅子的中点”中的路标是“中点”, 但是同时存在三个名词“初始点”、“椅子”、“中点”. 如何确定名词实体之间的关系涉及到名词短语语义处理问题. 目前中文名词短语语义处理研究还不深入, 文献 [22] 中的分析方法借鉴了通常的语义处理方法, 用谓词识别方法进行处理. 但是本文中较多的名词短语无法找到参考的动词, 例如 PL 语义组块. 因此, 本文从名词实体关系的角度对名词短语语义的确定进行了相关的研究.

3.1 名词短语语义问题的表示

设 A 、 B 为同一个名词短语 P 中的名词实体, 则 A 、 B 之间存在三种关系: 1) A 依赖于 B , 表示为 $A \rightarrow B$, 例如, “主卧门口”中 A 为“门口”, B 为“主卧”; 2) B 依赖于 A , 表示为 $B \rightarrow A$; 3) A 和 B 并列, 表示为 $A = B$, 例如“电视和茶几之间的椅子”中的“电视”、“茶几”.

需要解决的问题: 在已知 A 、 B 之间特征

$\Omega(A, B)$ 的情况下, 推导 A 、 B 之间的关系 R , 对应上面三种关系, A 、 B 之间三种关系表示为: 1) $R_{A \rightarrow B}$; 2) $R_{B \rightarrow A}$; 3) $R_{A=B}$.

3.2 名词实体关系中的特征

通过考察语料库中的包含两个以上名词实体的短语, 总结出 4 种实体关系的特征.

1) 名词归类, 用 f_A 、 f_B 来表示 (f_A 同 f_B). 在路径自然语言中, 名词实体大都是具体实物或者空间方位概念, 因此将名词按照物理空间的大小分为: a) 较大空间名词, 例如卧室、走廊等, 定义为 1 类名词; b) 具体实物名词, 例如电视、椅子等, 定义为 2 类名词; c) 指代部位名词, 例如边缘、头部等, 定义为 3 类名词; d) 其他, 定义为 4 类名词. 如图 2 中区域 3 所示, 其中的词典库包含以上各类名词实体词典, 因为所涉及的名词实体只是关于空间、建筑、物品等, 这样的词典基本能够覆盖到对应的日常应用词汇.

2) 两个实体名词之间是否有连词, 用 f_C 来表示.

3) 两个实体名词之间是否有“的”, 用 f_D 来表示.

4) 两个实体名词的相对位置, 用 f_P 来表示.

3.3 名词实体关系的推导

由贝叶斯式可得:

$$P(R|\Omega(A, B)) = \frac{P(\Omega(A, B)|R)P(R)}{P(\Omega(A, B))} \quad (14)$$

为了便于推导, 默认各个特征之间的关系式相互独立, 则:

$$\Omega(A, B) = f_A f_B f_C f_D f_P \quad (15)$$

故可得:

$$\begin{aligned} P(R|\Omega(A, B)) &= \frac{P(f_A f_B f_C f_D f_P | R) P(R)}{P(\Omega(A, B))} = \\ &= \frac{P(f_A | R) P(f_B | R) P(f_C | R) P(f_D | R) P(f_P | R) P(R)}{P(\Omega(A, B))} \approx \\ &= \frac{P(f_A | R) P(f_B | R) P(f_C | R) P(f_D | R) P(f_P | R) P(R)}{P(\Omega(A, B))} \end{aligned} \quad (16)$$

当 $A = B$ 时, 由语法知识可知, 即使 A 、 B 的相对位置发生变化, 也应该不影响其之间的关系, 所以概率 $p(R_{A=B}|\Omega(A, B))$ 不会发生变化. 所以有:

当 $P_A > P_B$ 时, $P(R_{A=B}|\Omega(A, B)) = P_1$, $\Omega(A, B) = f_A f_B f_C f_D f_P$.

当 $P_A < P_B$ 时, $P(R_{A=B}|\Omega'(A, B)) = P_2$, $\Omega'(A, B) = f_A f_B f_C f_D f_{P'}$.

P_A , P_B 分别表示名词实体 A , B 在名词短语中的位置. 因为位置不影响条件概率大小, 所以 $P_1 = P_2$, 则可得:

$$\begin{aligned} 2P(R_{A=B}|\Omega(A, B)) &= \frac{P(\Omega(A, B)|R_{A=B})P(R_{A=B})}{P(\Omega(A, B))} + \\ &= \frac{P(\Omega'(A, B)|R_{A=B})P(R_{A=B})}{P(\Omega(A, B))} = \\ &= \frac{P(\Omega(A, B))}{(P(f_A f_B f_C f_D f_P | R_{A=B})) +} \\ &= \frac{P(\Omega(A, B))}{P(f_A f_B f_C f_D f_{P'} | R_{A=B})} P(R_{A=B}) = \\ &= \frac{[P(f_A | R) P(f_B | R) P(f_C | R) P(f_D | R) (P(f_P | R) + P(f_{P'} | R)) P(R)]}{[P(\Omega(A, B))]} \end{aligned} \quad (17)$$

因为 $P(f_P | R)$ 与 $P(f_{P'} | R)$ 分别表示 A 、 B 互换位置之后的概率, 因此有:

$$P(f_P | R) + P(f_{P'} | R) = 1$$

故:

$$\begin{aligned} 2P(R_{A=B}|\Omega(A, B)) &= \frac{P(f_A | R) P(f_B | R) P(f_C | R) P(f_D | R) P(R)}{P(\Omega(A, B))} \\ &= \frac{P(f_A | R) P(f_B | R) P(f_C | R) P(f_D | R) P(R)}{P(\Omega(A, B))} \end{aligned} \quad (18)$$

所以, 当 A 、 B 为并列实体时:

$$\begin{aligned} P(R_{A=B}|\Omega(A, B)) &\approx \\ &= \frac{1}{2} P(f_A | R_{A=B}) P(f_B | R_{A=B}) \times \\ &= \frac{1}{2} P(f_C | R_{A=B}) P(f_D | R_{A=B}) P(R_{A=B}) \end{aligned} \quad (19)$$

由式 (19) 可知, 当 A 、 B 为并列关系时, 它们的相互位置对其之间的关系没有影响. 当 A 、 B 为偏正关系时:

$$\begin{aligned} P(R|\Omega(A, B)) &\approx P(f_A | R) \times \\ &= P(f_B | R) P(f_C | R) P(f_D | R) P(f_P | R) P(R) \end{aligned} \quad (20)$$

3.4 训练和推导

为了便于训练和推导, 上文中的特征都进行量化表示.

1)

$$f_A = \begin{cases} 0, & \text{如果 } A \text{ 为 1 类名词} \\ 1, & \text{如果 } A \text{ 为 2 类名词} \\ 2, & \text{如果 } A \text{ 为 3 类名词} \\ 3, & \text{如果 } A \text{ 为 4 类名词} \end{cases} \quad (21)$$

在训练时通过查询词典库来确定特征的值.

2) f_B 取值同上.

3)

$$f_C = \begin{cases} 0, & \text{如果 } A, B \text{ 之间存在连词} \\ 1, & \text{如果 } A, B \text{ 之间不存在连词} \end{cases} \quad (22)$$

4)

$$f_D = \begin{cases} 0, & \text{如果 } A, B \text{ 之间存在“的”} \\ 1, & \text{如果 } A, B \text{ 之间不存在“的”} \end{cases} \quad (23)$$

5)

$$f_P = \begin{cases} 0, & \text{如果 } A \text{ 在 } B \text{ 之前} \\ 1, & \text{如果 } A \text{ 在 } B \text{ 之后} \end{cases} \quad (24)$$

由于复杂名词短语的数量有限, 本文采用如下方法进行相关概率的求解:

$$p(f_i = j|R) = \frac{N_{f_i=j} + \delta}{N_R + 2\delta} \quad (25)$$

其中, $p(f_i = j|R)$ 表示当 A, B 关系为 R 时, 特征 f_i 值为 j 的概率, 其中, j 为整数, $f_i \in \{f_A, f_B, f_C, f_D, f_P\}$; $N_{f_i=j}$ 表示在语料库中, 特征 f_i 值为 j 时出现的次数; N_R 为 A, B 关系 R 在语料库中出现的次数. δ 是一个较小的值, 为防止数据稀疏的参数, 这里赋值为 0.0001.

为了获得各个特征对应的概率, 首先将语料中的所有非单个名词实体的名词短语标注出主要路标和次要路标. 然后利用式 (18)~(20) 统计出现的特征种类及其数量, 最后利用式 (21) 计算其对应的概率, 结果如表 2 所示. 表 2 表示存在两个名词实体 A, B , 其中 A 为不同类型名词时, A, B 之间关系所对应的概率. 其中第 1 行表示 A 和 B 并列关系的

概率, 第 2 行表示 A 依赖于 B 的概率, 第 3 行表示 B 依赖于 A 的概率.

表 3 中分别表示两个名词实体 A, B 之间存在连词和存在“的”时, A, B 之间的关系所对应的概率, 这里主要关注 A, B 是否是并列关系. 表 3 第 1 行表示 A, B 是并列关系时所对应的概率, 第 2 行表示 A, B 为非并列关系时的概率.

表 4 表示位置对名词依赖性的影响, 第 1 列表示 A 的位置在 B 之前时 A, B 之间所对应的依赖关系的概率, 第 2 列表示 A 的位置在 B 之后时所对应的概率. 一般来讲, 位置靠前的名词都是修饰后面的.

利用以上各个概率进行测试, 测试中所用的 11 个样本均能准确地推导出来, 准确率为 100%. 总的来说, 我们收集的样本还比较简单, 一般一个名词短语中名词实体的数量不超过 4 个, 因此推理的难度比较小.

在推导时, 首先提取名词短语中的名词实体, 将其两两组合, 由式 (19) 和 (20) 分别计算 $P(R_{A \rightarrow B})$ 、 $P(R_{B \rightarrow A})$ 、 $P(R_{A=B})$ 的值, 三者中最大的值对应的那个名词实体关系即为所求.

4 导航意向图的形成

4.1 路径结构化

本文沿用文献 [15] 中的路径描述原理, 但是在自然语言中不能规定系统输入语句的句法结构, 因此重新定义路径结构.

设路径自然语言 S 所对应的路径为 P , 则 $P = f(S)$. 由于歧义和语言本身模糊性的存在, P 与 S 是多对一的关系, 但本文不涉及所有可能路径的推理, 只解决字面意义上的路径自然语言理解问题, 所以在本文中 P 与 S 是一对一的关系, P 的结构可以定义为:

表 2 名词归类特征所对应的概率

Table 2 The probability corresponding to classified noun feature

	$f_A = 0$	$f_A = 1$	$f_A = 2$	$f_A = 3$
$P(f_A R_{A=B})$	0.048 388 6	0.935 482	1.6129×10^{-6}	0.016 130 6
$P(f_A R_{A \rightarrow B})$	0.047 620 5	0.730 158	0.174 604	0.047 620 5
$P(f_A R_{B \rightarrow A})$	0.523 809	0.460 318	1.5873×10^{-6}	0.015 874 6

表 3 连接词特征和是否含“的”的特征分布

Table 3 Feature distribution of conjunction and “of”

	$f_C = 0$	$f_C = 1$	$f_D = 0$	$f_D = 1$
$P(f_A R_{A=B})$	0.193 55	0.806 45	0.967 739	0.032 261 1
$P(f_A \overline{R_{A=B}})$	0.984 125	0.015 874 6	0.523 809	0.476 191

表 4 名词位置的特征分布
Table 4 Feature distribution of noun position

	$f_P = 0$	$f_P = 1$
$P(f_A R_{A \rightarrow B})$	0.999 998	1.5873×10^{-6}
$P(f_A R_{B \rightarrow A})$	1.5873×10^{-6}	0.999 998

1) P 可分为若干路径单元 $p_i, i = 1, 2, 3, \dots, n$, 即 $P = \bigcup p_i$;

2) $p_i = \{r_i, f_i\}$, 其中 r_i 表示路径的方向和距离 (如果该单元包含距离信息), 称为路径单元向量; $f_i = \{lf_i, di, lb_i\}$ 表示路径的其他导航辅助信息, 其中位于路径单元开始位置的路标为前向路标, 即 lf_i , 位于终点位置处的路标为后向路标, 即 lb_i ; di 表示该路径单元是否包含距离信息, 如果包含, 机器人可以以此为参照进行导航, 反之则完全依靠传感器进行导航; 每一个路径单元只能表示一个运动方向.

4.2 语义信息的提取

语义信息提取的目的是将已得到的各个组块中的有效信息组成一系列的路径单元.

4.2.1 组块的模板匹配

在提取出语料中的各个语义组块之后, 通过模板匹配的方法可以提取出各个组块中的关键信息. 现阶段, 我们主要关注路标的名称、距离、角度等信息.

语义组块 NL、VL、PL 主要是用于提取路标相关的信息, 例如, 名称、大小、颜色等. 这些组块统称为路标相关组块. 但是现阶段我们只关心路标的名称. 因此路标相关组块的模板的定义如表 5 所示.

表 5 路标相关组块模板槽

Table 5 The template slot of chunk related to landmark

槽名	注释
序号	表示该组块在句子中的序号
核心名词实体	起主要作用的名词实体, 数量可以大于 1
辅助名词实体	起辅助作用的名词实体, 数量可以大于 1

在填充路标相关组块时, 先获得该组块内部的名词实体 (图 2 中, 区域 2), 然后采用第 3 节中的名词短语处理方法确定组块内部的核心名词实体和辅助名词实体, 然后逐个填入槽中.

DTM 模块中, 主要提取距离、方向、转弯角度等信息. 模板定义如表 6 所示.

在填充 DTM 的模板时, 主要是依靠词性来确定哪个词填到哪个槽中:

1) 词性为 f (方位词)、 s (处所词) 的词就填充到“方向”槽中;

2) 词性为 v (动词)、 vi (内动词)、 vn (动名

词)、 vf (趋向动词) 则填充到“动作”槽中;

3) 词性为 m (数词) 则填充到“数量”槽中;

4) 词性为 q (量词), mq (数量词), qv (动量词) 则填充到“单位标示”槽中;

5) 词性为 d (副词) 则填充到“副词”槽中. 其中“数量”和“单位标示”的次序是一对的. 例如短语“走 10 米右转 45 度”, 填槽时就需要先填入“10”、“米”, 然后再分别填进“45”、“度”.

表 6 DTM 相关组块模板槽

Table 6 The template slot of chunk related to DTM

槽名	注释
序号	表示该组块在句子中的序号
方向	包含的关于方向的词, 例如: 左、南等
动作	有些动词本身包含动作, 例如: 转弯、掉头等
数量	距离, 或者转弯的角度
单位标示	表示量词属性的单位, 例如: 公里、度 (角度)
副词	表示动作的幅度或者属性

IDTM 模块中, 主要提取路标名称、距离、方向、转弯角度等信息. IDTM 的填槽和 DTM 类似, 主要是依靠词性来确定词填到哪个槽中:

1) 对 IDTM 中出现在靠前位置的名词短语采用表 5 中的槽, 进行填充, 其结果即为表 7 中的“前向路标”, “后向路标”的处理方法相同;

2) 表 7 中和表 6 中相同名称的槽的填充方法是相同的;

3) 词性为 p (介词) 的词填充到“介词”槽中.

表 7 IDTM 相关组块模板槽

Table 7 The template slot of chunk related to IDTM

槽名	注释
序号	表示该组块在句子中的序号
前向路标	IDTM 中位置靠前的名词短语经过表 5 填充完成的槽
后向路标	IDTM 中位置靠后的名词短语经过表 5 填充完成的槽
方向	包含的关于方向的词, 例如: 左、南等
动作	有些动词本身包含动作, 例如: 转弯、掉头等
数量	距离, 或者转弯的角度
单位标示	表示量词属性的单位, 例如: 公里、度 (角度)
副词	表示动作的幅度或者属性
介词	IDTM 中可能会出现多个介词

4.2.2 路径单元的提取

这部分的工作主要就是从经过填槽操作获得的各个依次排列槽体中获得路径单元.

通常来说, 如果表示路径的语言用没有表示具体的方向, 一般就是指方向为“前”, 例如“过了红绿灯左拐, 走 30 米就到了”. 这里“走 30 米”没有具体说是“左拐”之后向哪边走, 但是正常的理解是“左拐, 向前走 30 米”. 因此, 在本文处理过程中, 没有标明方向的运动, 都默认为是“前”. 对于没有标

注具体距离的运动, 本文都默认是运动了的单位距离, 即为 1. 在处理过程中定义默认的路径向量为 $\mathbf{r}_0 = (0, 1)$.

根据路径单元的定义, 当前需提取的路径单元为 $p_i, i \geq 1$, 其对应的路径向量为 $\mathbf{r}_i$, 前向路标为 lf_i , 后向路标为 lb_i , 距离标记为 d_i . 为了便于处理, 定义一个用来保存前一个路径单元的路径向量的路径向量 $\mathbf{rf}_i$. 设输入的组块为 $c_j, j = 1, 2, \dots, n, n$ 表示组块的数量.

在提取路径单元时, 顺序读入各个组块对应的槽体. 提取流程如图 3 所示, 解释如下:

- 1) 输入一个组块;
- 2) 判断当前组块是不是路标相关的组块, 如果不是, 则转到步骤 7), 如果是, 则转到步骤 3);
- 3) 将当前路标组块中的路标作为 lb_i ;
- 4) 如果当前组块是第一个组块, 则 $\mathbf{r}_i = \mathbf{r}_0$, 保存 $p_i, \mathbf{rf}_i = \mathbf{r}_i$, 否则, 转 5);
- 5) 判断前一个组块是不是 DTM 或者 IDTM, 如果是, 则保存 p_i , 如果不是, 且 $i > 1$, 则 $\mathbf{r}_i = \mathbf{rf}_i$, 保存 p_i ;
- 6) 判断当前组块是不是最后一个组块, 如果不是, 则转向 1), 如果是, 则结束;
- 7) 如果是 DTM 组块, 则转到 8), 如果不是, 则转到 19);
- 8) 如果该组块不包含方向, 则转向 9), 反之则转向 14);
- 9) 如果该组块是第一个组块, 则 $\mathbf{r}_i = \mathbf{r}_0$, 否则计算 $\mathbf{r}_i$, 具体方法见第 4.3 节;
- 10) 如果当前组块包含距离, 其值为 l , 则 $d_i = 1, \mathbf{r}_i = \mathbf{r}_i \times l$, 否则, $d_i = 0$;
- 11) 如果前一个组块是 DTM, 转到 12), 否则, 转到 13);
- 12) 如果 $i > 1$ 且 d_{i-1} 值为 1, 则 $\mathbf{r}_i = \mathbf{rf}_i + \mathbf{r}_i$, 保存 p_i , 转到 6);
- 13) 如果后一个组块不是路标相关组块或者当前组块为最后一个组块, 则保存 p_i , 转向 6), 否则直接转向 6);
- 14) 判断组块内部是不是带明确的旋转角度, 如果没有则转到 15), 如果有则转到 18);
- 15) 判断当前组块是否是第一个组块, 如果是, 则通过判断方位词确定方向, 如果不是, 则通过 $\mathbf{rf}_i$ 和组块中的方位词来确定当前路径单元的方向向量, 具体处理过程见第 4.3 节;
- 16) 如果当前组块包含距离, 其值为 l , 则 $d_i = 1, \mathbf{r}_i = \mathbf{r}_i \times l$, 否则, $d_i = 0$, 然后, 转到 13);
- 17) 如果是第一个组块, 则利用旋转角度计算 $\mathbf{r}_i$, 反之, 则利用 $\mathbf{rf}_i$ 和旋转角度计算当前的方向向量 $\mathbf{r}_i$, 然后, 转到 16);

18) 如果是 IDTM 组块, 则转向 19), 否则结束;

19) 将 IDTM 中的前向、后向路标赋值给当前路径单元, 转到 8).

4.3 方位问题的处理

路径自然语言中的方位一般分为三种: 1) 相对方位, 例如前、后、左、右等, 这类方位涉及到方位的推导; 2) 绝对方位, 例如东、西、南、北等, 这类方向文中不做处理, 机器人导航时将靠硬件识别, 这里不予讨论; 3) 间接方位, 一般用物体指代, 例如, “从书房走到卧室”, 这里书房和卧室的位置关系没有明确指出来, 但是人一般都能理解其方位是由书房指向卧室, 这类方位的识别需要在导航时确定方向, 这里默认方向都是向前. 对于相对方位, 参考笛卡尔坐标系, 当机器人没有运动时, 定义与 Y 轴正方向相同的方向为“前”, 与 X 轴正方向相同的方向为“右”. 则可以用单位向量量化表示各个方向, 如表 8 所示.

表 8 相对方位向量化的初始定义

Table 8 The initial definition of relative direction vectors

方向	单位向量
右	$\mathbf{e}_1 = (1, 0)$
左	$\mathbf{e}_2 = (-1, 0)$
后	$\mathbf{e}_3 = (0, -1)$
前	$\mathbf{e}_4 = (0, 1)$
右后	$\mathbf{e}_5 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$
右前	$\mathbf{e}_6 = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$
左后	$\mathbf{e}_7 = \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$
左前	$\mathbf{e}_8 = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$

当机器人开始运动时, 以顺时针方向为正、逆时针方向为负定义转过的角度, 可以得到表 9.

设单位向量 $\mathbf{d}_i = (x_i, y_i)$ 为机器人运行过程中在第 i ($i = 1, 2, \dots, n$) 个路径单元内相对上一路径单元的方向, 则机器人朝向推导规则如下:

1) 当机器人还未启动时, 机器人朝向为 $\mathbf{d}_0$, 根据路径自然语言的描述, $\mathbf{d}_0$ 值可以从表 8 中得到;

2) 当机器人启动之后, 若第 i 个路径单元的相对方位变化为 $\alpha_j, j = 1, 2, \dots, 8$, 则 $\mathbf{d}_{i+1}$ 的值可以由向量旋转式推导得出:

$$\begin{cases} \mathbf{d}_{i+1} = (x_{i+1}, y_{i+1}) \\ x_{i+1} = x_i \cdot \cos \alpha_j - y_i \cdot \sin \alpha_j \\ y_{i+1} = y_i \cdot \cos \alpha_j + x_i \cdot \sin \alpha_j \end{cases} \quad (26)$$

在实际的处理过程中, 方位词的同义词较多. 为了简化处理, 本文定义了同义词词典 (共 151 个), 在处理过程中将所有同义的进行归类, 然后转化为对应的标准同义词, 如图 2 中区域 3 所示.

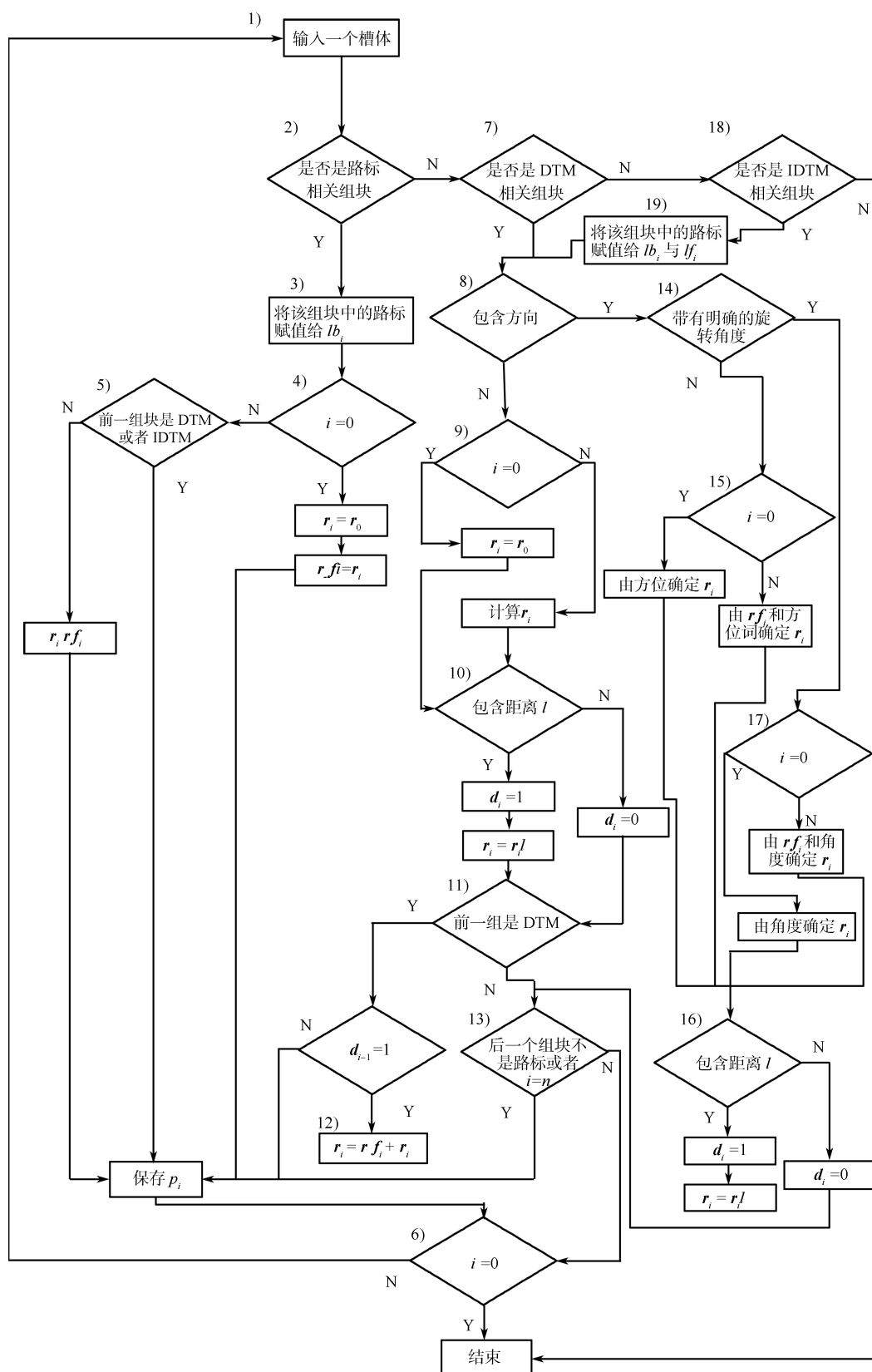


图 3 路径单元提取流程图

Fig. 3 The flowchart of picking-up route unit

表 9 相对方位的角度变化
Table 9 The change of angulars about relative orientation

方向	单位向量
右	$\alpha_1 = \frac{\pi}{2}$
左	$\alpha_2 = -\frac{\pi}{2}$
后	$\alpha_3 = \pi$
前	$\alpha_4 = 0$
右后	$\alpha_5 = \frac{3\pi}{4}$
右前	$\alpha_6 = \frac{\pi}{4}$
左后	$\alpha_7 = -\frac{3\pi}{4}$
左前	$\alpha_8 = -\frac{\pi}{4}$

4.4 路径的生成

每一个路径单元中的路径 $\mathbf{r}_i = \mathbf{d}_i \cdot l_i$, 其中 $\mathbf{d}_i$ 表示第 i 个路径单元方向的单位向量, l_i 表示该路径单元的距离, 如果该路径单元中没有距离信息, 则用单位距离表示, 如果有则按照实际距离表示. 设 L 为 P 所对应的运行路线. 则 $L = \cup \mathbf{r}_i$, 表示由 $\mathbf{r}_i$ 首尾相接形成的路线. 例如: 一条路径由三个路径单元组成, $\mathbf{r}_1 = (1, 0)$, $\mathbf{r}_2 = (0, 1)$, $\mathbf{r}_3 = (1, 0)$, 则对应的 L 如图 4 所示.

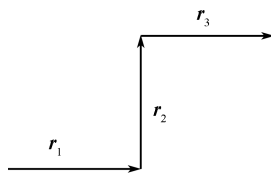


图 4 L 表示的路线

Fig. 4 The route represented by L

5 实验结果

下文中所涉及的实验指标定义如下:

$$P = \frac{N_b}{N_t} \quad (27)$$

$$R = \frac{N_b}{N_z} \quad (28)$$

$$F = \frac{(\beta^2 + 1) \times R \times P}{\beta^2 \times R + P} \times 100\% \quad (29)$$

这里, N_b 表示正确标注的组块个数, N_t 表示标注组块的总个数, N_z 表示正确组块的总个数, 这里设 $\beta = 1$.

5.1 各个组块分析的处理结果

因为语料库总体规模较小, 没有专门区分出训

练语料和测试语料. 为了使训练语料达到最大化, 测试时, 采用 Cross validation (交叉验证) 方法进行了相关的实验. 将 216 句样本随机分成 54 份, 每份 4 句样本. 测试时其中一份作为测试样本, 其他则作为训练样本, 因此每组实验都需要进行 54 次. 本文统计了各组实验正确率 (P)、召回率 (R) 和 F 值的区间分布以及最后的平均值.

实验 1. 名词实体提取实验

图 5 是名词实体提取实验结果的分布图, 由图可知大部分实验的准确率都在 95% ~ 100% 的区间, 但是相比准确率, 召回率和 F 值在更小的区间有较多的分布, 说明有较多的组块没有被识别出来. 名词实体实验的平均结果见表 10.

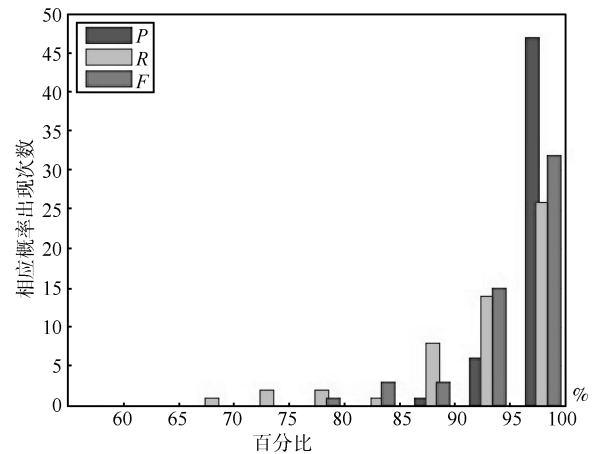


图 5 名词实体提取实验结果总体分布

Fig. 5 The global distribution of experimental result for picking up noun entity

表 10 名词实体识别的平均结果

Table 10 The average recognition result of noun entity

实验指标	P	R	F
平均值 (%)	97.29	92.64	94.81

实验 2. 名词短语提取实验

图 6 是名词短语提取实验结果的分布图, 由图可知大部分实验的准确率也都在 95% ~ 100% 的区间, 分布相对比较集中. 相比图 5, 召回率和 F 值在 95% ~ 100% 中的分布更多. 名词短语提取实验平均结果见表 11.

表 11 名词短语组块处理结果

Table 11 The processed result of noun phrase chunk

实验指标	P	R	F
平均值 (%)	96.28	91.92	93.98

实验 3. 语义组块提取实验

图 7 是语义组块提取实验结果总体分布图, 由图可知大部分实验的准确率在 80% ~ 85% 区间,

总体呈现中间分布多、两边分布少的情况, 分布相对比较分散. 语义组块提取实验平均结果见表 12.

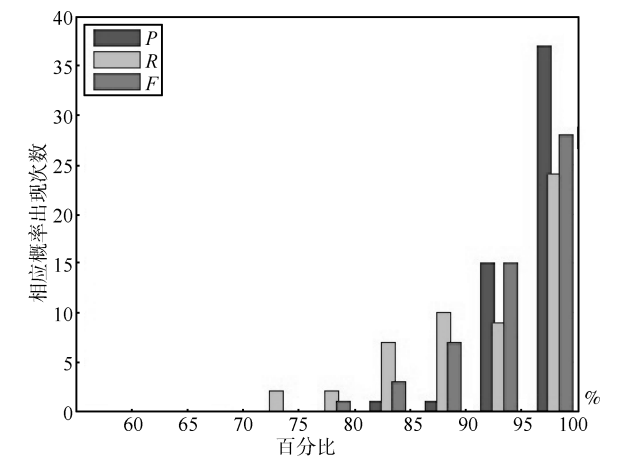


图 6 名词短语提取实验结果总体分布

Fig. 6 The global distribution of experimental result for picking up noun phrase

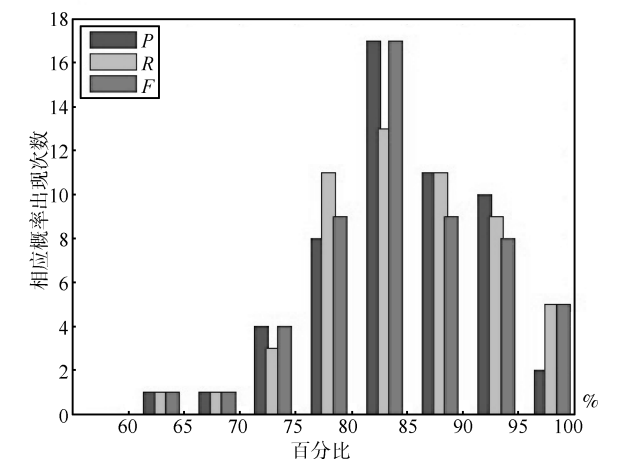


图 7 语义组块提取实验结果总体分布

Fig. 7 The global distribution of experimental result for picking up semantic chunk

表 12 语义组块处理结果

Table 12 The processed result of semantic chunk

	P	R	F
平均值 (%)	83.65	84.33	83.98

语义组块处理的准确率总体而言比采用一般的浅层语义标注方法^[23]要高(因为采用的不是相同的语料库, 且语料库规模也相差比较大, 应用背景不同, 没有绝对可比性, 只能说有一定的启发意义). 从图 7 可知语义组块提取的准确度稳定性还不高, 主要问题在于有部分的语义组块的句法组成有一定的相似性, 主要集中在 PL 和 IDTM. 例如, 如果一个 IDTM 的结构比较简单, 主要以名词短语组成, 不包含方向和距离时, 就有可能出现误判的情况, 从而可

能导致准确率下降.

实验 4. CRFs 和 CCRFs 的对比实验

采用随机抽取的方法, 我们抽取了 20 组样本, 分别采用 CRFs 和 CCRFs 进行了实验, 如图 8 和图 9 所示.

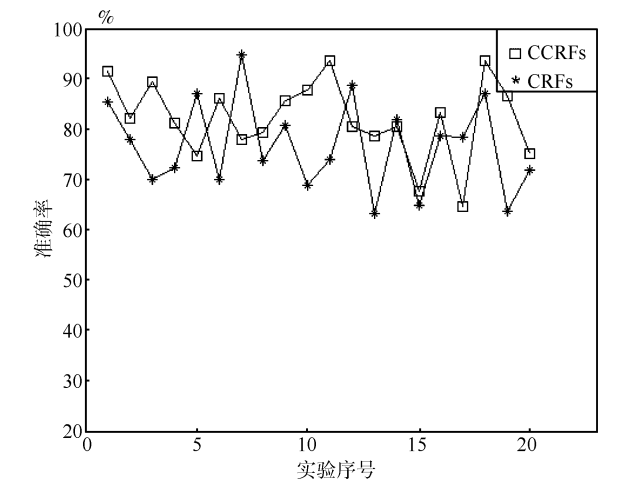


图 8 采用 CCRFs 前后的准确率对比

Fig. 8 The precision rate comparison between before and after adopting CCRFs

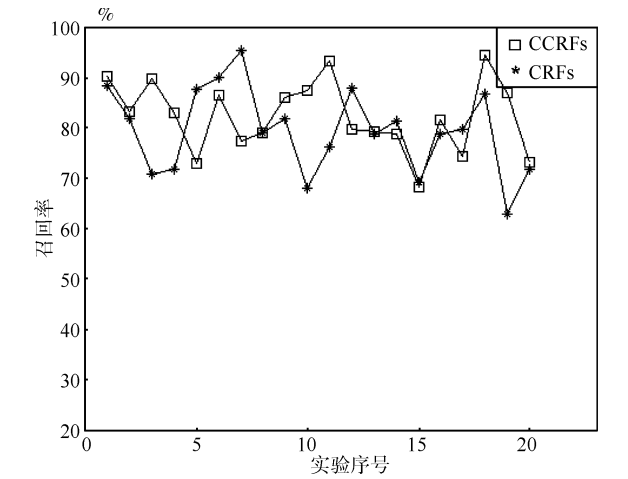


图 9 采用 CCRFs 前后的召回率对比

Fig. 9 The recall rate comparison between before and after adopting CCRFs

从图中可以看出, 采用 CCRFs 方法平均准确率和召回率分别为 81.9868% 和 82.3226%, 而采用 CRFs 方法平均准确率和召回率分别为 76.6268% 和 79.4495%, 总体来说采用 CCRFs 方法后准确率和召回率都有所提高, 约分别提高 7% 和 3.6%. 但是, 还有部分实验组的准确率和召回率反而有所下降, 这主要是因为采用 CCRFs 之后算法复杂度提高, 上一层名词短语识别的错误会带入下一层, 使错误出现叠加现象, 会影响到准确率和召回率. 这方面的问题将是我们今后研究的重点.

5.2 路径的生成实验

本文选取语料库中的一句路径自然语言作为示例, 整个处理流程如图 2 所示. 所用的生活语料如例句 1 所示.

例句 1. “先直走, 碰到椅子后右拐, 沿着 L 型柜向前走到吉他处左拐, 再在饮水机前左拐并直走至盆景处”.

名词实体提取的结果为: 先/d 直/d 走/v 碰到/v [椅子/n] NL 后/f 右/f 拐/v 沿着/p [L/x 型/k 柜/ng] NL 向前/vi 走/v 到/v [吉他/n] NL 处/n 左/f 拐/v 再/d 在/p [饮水/n 机/ng] NL 前/f 左/f 拐/v 并/cc 直/d 走/v 至/p [盆景/n 处/n] NL.

名词短语处理结果为: 先/d 直/d 走/v 碰到/v [椅子/n] NL 后/f 右/f 拐/v 沿着/p [L/x 型/k 柜/ng] NL 向前/vi 走/v 到/v [吉他/n 处/n] NL 左/f 拐/v 再/d 在/p [饮水/n 机/ng] NL 前/f 左/f 拐/v 并/cc 直/d 走/v 至/p [盆景/n 处/n] NL.

语义提取的处理结果为: [先/d]VL [直/d 走/v] DTM [碰到/v 名词/n] VL 后/f [右/f 拐/v] DTM [沿着/p 名词/n 向前/vi] IDTM [走/v 到/v 名词/n] VL [左/f 拐/v] DTM 再/d [在/p 名词/n] PL 前/f [左/f 拐/v 并/cc 直/d 走/v] DTM [至/p 名词/n] VL.

从以上语义提取的结果可知, 总共提取出 10 个语义组块, 其中, 第一个组块 “[先/d] VL” 为误识别, 其余都正确. 按照语义提取的方法, 可以得到语句中的关键信息如下: “直走”, “椅子”, “右”, “拐”, “L 型柜子”, “前”, “吉他”, “左”, “饮水机”, “左拐”, “直走”, “盆景”.

按照每个路径单元只能有一个方向的原则, 将上述关键信息组织成相应的路径单元:

- 1) “前”, “椅子”;
- 2) “右”, “L 型柜子”;
- 3) “前”, “吉他”;
- 4) “左”, “饮水机”;
- 5) “左”, “盆景处”.

默认机器人面向前方, 则按照第 4.1 节和第 4.2 节中的方法可以计算出路径. 机器人导航的地图如图 10 所示, 图中方框表示路标在机器人前进路径中的位置.

例句 2. “先往前走 2 米, 向右转前进 2 米, 再次向右转, 前进 3 米, 再向左转前进 3 米, 向左转向前进 5 米停下”. 形成的地图如图 11 所示.

例句 3. “向前走到椅子处右拐前进, 经过 L 型柜子后朝斜前方椅子前进, 然后左转一定角度, 从机器人与电视柜之间的路走向终点”. 形成的地图如图 12 所示.

例句 2 和例句 3 的处理过程与例句 1 相似. 其中例句 2 中出现了“虚拟路标”, 这里是表示路径方向发生改变的情形, 便于机器人导航^[16, 24-25]. 例句 3 则是反映了名词短语处理的结果, 其中“机器人”和“电视柜”都是用来修饰路的名词, 在图中反映了相关的位置关系.

6 总结与展望

本文所有采用的方法都是在研究语料的基础上逐步提出来的, 上面若干实验数据和路径结果也都是建立在标注语料库的基础上. 由于语料库的规模较小, 可能导致某些语言现象没有反映出来, 进而没

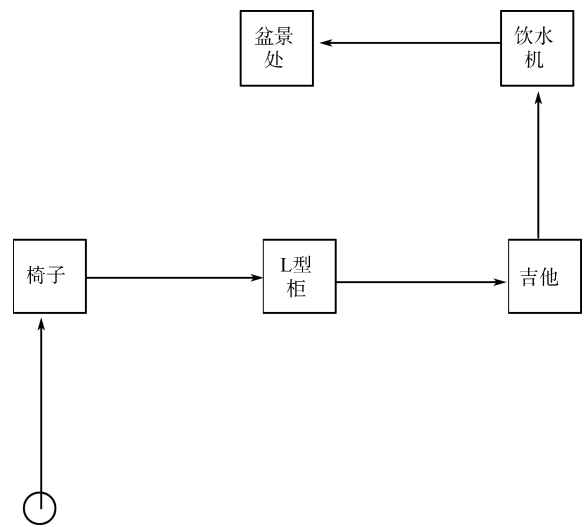


图 10 例句 1 导航结果

Fig. 10 The navigation result corresponding to Example sentence 1

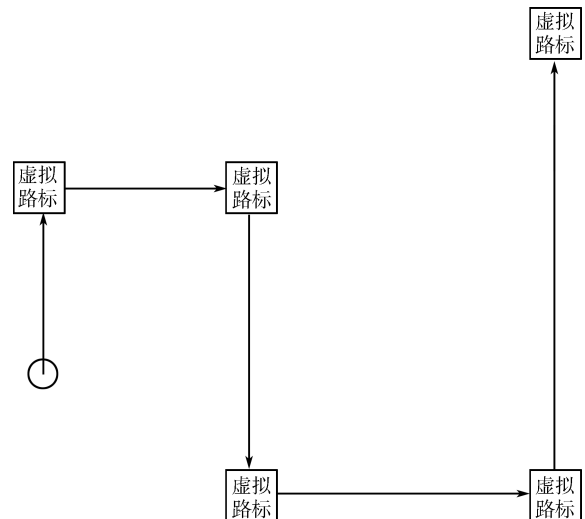


图 11 例句 2 导航结果

Fig. 11 The navigation result corresponding to Example sentence 2

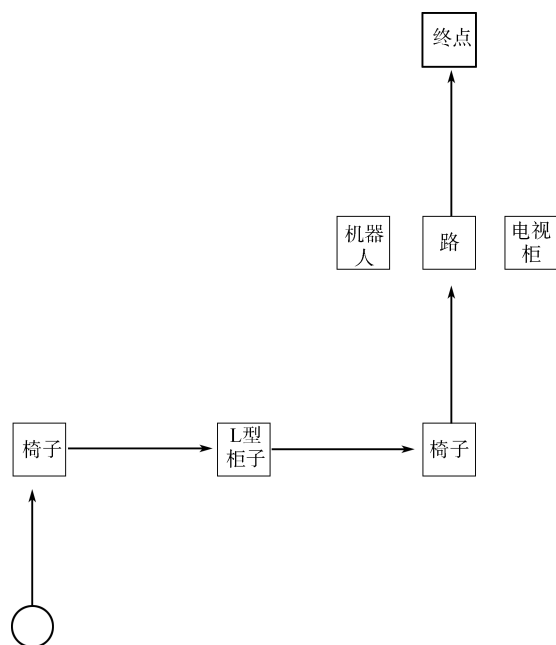


图 12 例句 3 导航结果

Fig. 12 The navigation result corresponding to

Example sentence 3

有被考虑进我们的方法中。但是建立大型的具有广泛代表性的语料库需要巨大的人力和物力,更重要的是需要长时间地研究其中的规律,进而提出高效的标注方法。但是,由于中文路径自然语言还很少有人关注,当前确实存在一些难度。在没有大型语料库、没有通用测试环境的条件下,本文给出的实验结果更多的是定性的结论,在一定程度上反映文中提出方法的可行性。客观地讲,本文方法还不是很完善,距离成熟的商业应用还有一定的差距,但还是期望我们的工作能给后来的研究以启发。

References

- Nie Xian-Li, Jiang Ping, Chen Hui-Tang. Natural language programming for tasks for a mobile robot. *Robot*, 2003, **25**(4): 308–317
(聂仙丽, 蒋平, 陈辉堂. 采用自然语言的移动机器人任务编程. *机器人*, 2003, **25**(4): 308–317)
- Le Xiao-Qiu, Yang Chong-Jun, Yu Wen-Yang. Spatial concept extraction based on spatial semantic role in natural language. *Geomatics and Information Science of Wuhan University*, 2005, **30**(12): 1100–1103
(乐小虬, 杨崇俊, 于文洋. 基于空间语义角色的自然语言空间概念提取. *武汉大学学报 (信息科学版)*, 2005, **30**(12): 1100–1103)
- Zhang Xue-Ying, Lv Guo-Nian, Huan Jian. Approach to route description in Chinese natural language for electronic maps. *Geo-Information Science*, 2008, **10**(6): 757–762
(张雪英, 阚国年, 宦建. 面向汉语的自然语言路径描述方法. *地球信息科学*, 2008, **10**(6): 757–762)
- Klippel A, Hansen S, Richter K F, Winter S. Urban granularities — a data structure for cognitively ergonomic route directions. *Geoinformatica*, 2009, **13**(2): 223–247
- Shimizu N, Haas A. Learning to follow navigational route instructions. In: *Proceedings of the 21st International Joint Conferences on Artificial Intelligence*. San Francisco, USA: Morgan Kaufmann Publishers Inc., 2009. 1488–1493
- Yuan W, Brunskill E, Kollar T, Roy N. Where to go: interpreting natural directions using global inference. In: *Proceedings of the 2009 IEEE International Conference on Robotics and Automation*. Kobe, Japan: IEEE, 2009. 3761–3767
- Kollar T, Tellex S, Roy D, Roy N. Toward understanding natural language directions. In: *Proceedings of the 5th ACM/IEEE International Conference on Human Robot Interaction*. Piscataway, NJ, USA: IEEE, 2010. 259–267
- Kollar T, Dickerson S, Tellex S, Banerjee A G, Walter M R, Tellex S, Roy N. Towards Understanding Hierarchical Natural Language Commands for Robotic Navigation and Manipulation. Technical Report MIT-CSAIL-TR-2011-007, Massachusetts Institute of Technology, USA, 2011
- Tellex S, Kollar T, Dickerson S, Walter M R, Banerjee A G, Teller S J, Roy N. Approaching the symbol grounding problem with probabilistic graphical models. *AI Magazine*, 2011, **32**(4): 243–252
- Nick R, Tellex S, Thaker P, Deits R, Simeonov D, Kollar T, Roy N. Toward a Probabilistic Approach to Acquiring Information from Human Partners Using Language. Technical Report MIT-CSAIL-TR-2012-002, Massachusetts Institute of Technology, USA, 2012
- Simeonov D, Tellex S, Kollar T, Roy N. Toward interpreting spatial language discourse with grounding graphs. In: *Proceedings of the 2011 RSS Workshop on Grounding Human-Robot Dialog for Spatial Tasks*. California, USA: Robotics: Science and Systems, 2011
- Tellex S, Thaker P, Deits R, Kollar T, Roy N. Toward information theoretic human-robot dialog. In: *Proceedings of the 2012 Robotics: Science and Systems*. Sydney, Australia: Robotics: Science and Systems, 2012. 1432–1440
- Kordjamshidi P, van Otterlo M, Moens M F. Spatial role labeling: towards extraction of spatial relations from natural language. *ACM Transactions on Speech and Language Processing*, 2011, **8**(3): 1–36
- Wu Xue-Jian. Visual Navigation Method Research for Mobile Robot Based on Hand-drawn Map [Master dissertation], Southeast University, China, 2011
(吴雪建. 基于手绘地图的移动机器人视觉导航方法研究 [硕士学位论文], 东南大学, 中国, 2011)
- Li Xin-De, Zhang Xiu-Long, Dai Xian-Zhong. A visual navigation method of mobile robot based on constrained natural language processing. *Robot*, 2011, **33**(6): 742–749
(李新德, 张秀龙, 戴先中. 一种基于受限自然语言处理的移动机器人视觉导航方法. *机器人*, 2011, **33**(6): 742–749)
- Zhang Xiu-Long, Li Xin-De, Dai Xian-Zhong. Semantic role labeling method for route natural language based on chunk parsing. *Journal of Southeast University (Natural Science Edition)*, 2012, **42**(S1): 127–131
(张秀龙, 李新德, 戴先中. 基于组块分析的路径自然语言语义角色标注方法. *东南大学学报 (自然科学版)*, 2012, **42**(S1): 127–131)

- 17 Ding Wei-Wei, Chang Bao-Bao. Chinese semantic role labeling based on semantic chunking. *Journal of Chinese Information Processing*, 2009, **23**(5): 53–61
(丁伟伟, 常宝宝. 基于语义组块分析的汉语语义角色标注. 中文信息学报, 2009, **23**(5): 53–61)
- 18 Tang Bu-Zhou, Wang Xiao-Long, Wang Xuan. Confidence-weighted online sequence labeling algorithm. *Acta Automatica Sinica*, 2011, **37**(2): 188–195
(汤步洲, 王晓龙, 王轩. 置信度加权在线序列标注算法. 自动化学报, 2011, **37**(2): 188–195)
- 19 ICTCLAS Chinese word segmentation system [Online], available: <http://ictclas.org>, July 3, 2012
- 20 CRF++: Yet Another CRF toolkit [Online], available: <http://crfpp.googlecode.com>, February 13, 2013
- 21 Lafferty J D, McCallum A, Pereira F C N. Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: Proceedings of the 18th International Conference on Machine Learning. San Francisco, USA: Morgan Kaufmann Publishers Inc., 2001. 282–289
- 22 Zhang Nan. The Semantic Analysis on Verb Phrase and Noun Phrase in Restricted Domain [Master dissertation], Tianjin Normal University, China, 2008
(张楠. 限定领域内动词短语和名词短语的语义分析 [硕士学位论文], 天津师范大学, 中国, 2008)
- 23 Liu Ting, Che Wan-Xiang, Li Sheng. Semantic role labeling with maximum entropy classifier. *Journal of Software*, 2007, **18**(3): 565–573
(刘挺, 车万翔, 李生. 基于最大熵分类器的语义角色标注. 软件学报, 2007, **18**(3): 565–573)
- 24 Sun Zuo-Lei, Mao Xu-Chu. Quantification of relative movement uncertainty based on laser scan matching and graphical model for mobile robot. *Acta Automatica Sinica*, 2011,

37(2): 205–213

(孙作雷, 茅旭初. 基于激光束匹配和图模型的移动机器人相对运动估计误差分析. 自动化学报, 2011, **37**(2): 205–213)

- 25 Li Bo, Yang Dan, Deng Lin. Visual vocabulary tree with pyramid TF-IDF scoring match scheme for loop closure detection. *Acta Automatica Sinica*, 2011, **37**(6): 665–673
(李博, 杨丹, 邓林. 移动机器人闭环检测的视觉字典树金字塔 TF-IDF 得分匹配方法. 自动化学报, 2011, **37**(6): 665–673)



李新德 东南大学自动化学学院副教授, 博士生导师. 主要研究方向为智能机器人, 人机交互, 机器感知, 信息融合, 不确定推理和机器视觉. 本文通信作者.

E-mail: xindeli@seu.edu.cn

(**LI Xin-De** Associate professor at the School of Automation, Southeast University. His research interest covers

intelligent robot, human-robot interaction, robot perception, information fusion, and reasoning under uncertainty, machine vision. Corresponding author of this paper.)



张秀龙 2013 年获得东南大学自动化学学院硕士学位. 主要研究方向为机器人导航和路径自然语言理解.

E-mail: zhang_seu@163.com

(**ZHANG Xiu-Long** Received his master degree from the School of Automation, Southeast University in 2013. His research interest covers robot navigation and route instructions using natural language processing.)