



(12) 发明专利申请

(10) 申请公布号 CN 115965017 A

(43) 申请公布日 2023.04.14

(21) 申请号 202310009908.8

(22) 申请日 2023.01.04

(71) 申请人 北京三维天地科技股份有限公司
地址 100000 北京市海淀区西四环北路119号A座3层309室

(72) 发明人 金震 张京日 耿金帅

(74) 专利代理机构 北京冠和权律师事务所
11399
专利代理师 郑延斌

(51) Int. Cl.

G06F 40/289 (2020.01)

G06F 40/211 (2020.01)

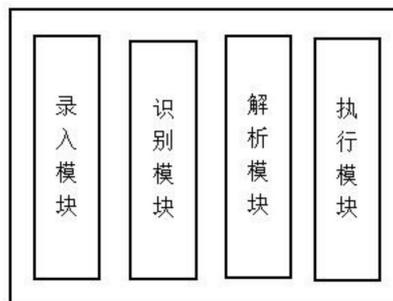
权利要求书3页 说明书9页 附图2页

(54) 发明名称

一种基于开发平台的多语言录入和解析系统及方法

(57) 摘要

本发明提供了一种基于开发平台的多语言录入和解析系统及方法包括:录入模块,用于录入文本信息,识别模块,用于获取所述文本信息的文本特征,将所述文本特征与每一语种对应的样本特征进行对比,得到所述文本信息对应的文本语种,根据所述文本语种建立译汉模型,解析模块,用于将所述文本信息输入到所述译汉模型中得到汉语信息,解析所述汉语信息,得到若干条文本数据,执行模块,用于将所述文本数据填充到开发平台中对应的文本框内进行显示,可以在开发平台中输入任何语种的文本信息,然后将其解析为统一的语言再输入到开发平台的显示页面上,打破了传统录入的弊端,实现了灵活录入。



1. 一种基于开发平台的多语言录入和解析系统,其特征在于,包括:
 - 录入模块,用于录入文本信息;
 - 识别模块,用于获取所述文本信息的文本特征,将所述文本特征与每一语种对应的样本特征进行对比,得到所述文本信息对应的文本语种,根据所述文本语种建立译汉模型;
 - 解析模块,用于将所述文本信息输入到所述译汉模型中得到汉语信息,解析所述汉语信息,得到若干条文本数据;
 - 执行模块,用于将所述文本数据填充到开发平台中对应的文本框内进行显示。
2. 如权利要求1所述的一种基于开发平台的多语言录入和解析系统,其特征在于,还包括:
 - 备份模块,用于根据所述文本数据在开发平台中文本框中的填充结果,得到汉语显示版图,根据所述文本语种将所述汉语显示版图翻译为目标显示版图;
 - 显示模块,用于显示所述汉语显示版图和目标显示版图。
3. 如权利要求1所述的一种基于开发平台的多语言录入和解析系统,其特征在于,所述录入模块,包括:
 - 输入单元,用于供用户输入原始文本;
 - 筛选单元,用于将所述原始文本进行文本筛选,得到目标文本;
 - 修正单元,用于根据时间顺序调整所述目标文本,得到文本信息。
4. 如权利要求1所述的一种基于开发平台的多语言录入和解析系统,其特征在于,所述识别模块,包括:
 - 训练单元,用于将所述文本信息进行语法训练,得到若干条文本语句;
 - 分析单元,用于分别提取每一文本语句对应的语句关键词,分析所述语句关键词对应的词语组成结构;
 - 识别单元,用于根据所述词语组成结构得到对应语句关键词的词特征,根据所述词特征的属性,确定所述文本信息对应的文本语种;
 - 建模单元,用于获取所述文本语种与汉语之间的翻译对照信息,建立译汉模型。
5. 如权利要求1所述的一种基于开发平台的多语言录入和解析系统,其特征在于,所述解析模块,包括:
 - 文本解析单元,用于获取所述文本语种的语句结构列表,根据所述语句结构列表分析所述文本信息中包含的若干中文本语句,建立所述文本信息的语句结构框架;
 - 模型解析单元,用于将所述文本信息输入到所述译汉模型中,在所述译汉模型中获取所述文本信息中包含的若干条文本语句,获取每一文本语句对应的若干个汉语词;
 - 文本修正单元,用于将所述汉语词输入到所述语句结构框架中进行语句重组,得到若干条第一汉语语句,当所述第一汉语语句中存在病句时,分析每一第一汉语语句对应的语句问题,并进行语句修正,得到第二汉语语句;
 - 文本转换单元,用于根据所述第二汉语语句建立汉语信息,得到若干条文本数据。
6. 如权利要求1所述的一种基于开发平台的多语言录入和解析系统,其特征在于,所述执行模块,包括:
 - 第一执行单元,用于解析所述文本数据,得到每一文本数据中包含的数据边界点,为每一文本数据的数据边界点上建立边界索引,得到每一文本数据对应的边界索引集;

第二执行单元,用于获取所述开发平台中每一文本框可填充的文本数据类型,为每一文本框建立数据索引;

第三执行单元,用于分别利用每一所述数据索引遍历每一文本数据对应的边界索引集,为每一数据索引匹配相应的边界索引,建立匹配列表;

第四执行单元,用于根据所述匹配列表将所述文本数据填充到开发平台中对应的文本框内进行显示。

7. 一种基于开发平台的多语言录入和解析方法,其特征在于,包括:

步骤1:录入文本信息;

步骤2:获取所述文本信息的文本特征,将所述文本特征与每一语种对应的样本特征进行对比,得到所述文本信息对应的文本语种,根据所述文本语种建立译汉模型;

步骤3:将所述文本信息输入到所述译汉模型中得到汉语信息,解析所述汉语信息,得到若干条文本数据;

步骤4:将所述文本数据填充到开发平台中对应的文本框内进行显示。

8. 如权利要求7所述的一种基于开发平台的多语言录入和解析方法,其特征在于,所述步骤2包括:

步骤21:将所述文本信息进行语法训练,得到若干条文本语句;

步骤22:分别提取每一文本语句对应的语句关键词,分析所述语句关键词对应的词语组成结构;

步骤23:根据所述词语组成结构得到对应语句关键词的词特征,根据所述词特征的属性,确定所述文本信息对应的文本语种;

步骤24:获取所述文本语种与汉语之间的翻译对照信息,建立译汉模型。

9. 如权利要求7所述的一种基于开发平台的多语言录入和解析方法,其特征在于,所述步骤3包括:

步骤31:获取所述文本语种的语句结构列表,根据所述语句结构列表分析所述文本信息中包含的若干中文本语句,建立所述文本信息的语句结构框架;

步骤32:将所述文本信息输入到所述译汉模型中,在所述译汉模型中获取所述文本信息包含的若干条文本语句,获取每一文本语句对应的若干个汉语词;

步骤33:将所述汉语词输入到所述语句结构框架中进行语句重组,得到若干条第一汉语语句,当所述第一汉语语句中存在病句时,分析每一第一汉语语句对应的语句问题,并进行语句修正,得到第二汉语语句;

步骤34:根据所述第二汉语语句建立汉语信息,得到若干条文本数据。

10. 如权利要求7所述的一种基于开发平台的多语言录入和解析方法,其特征在于,所述步骤4包括:

步骤41:解析所述文本数据,得到每一文本数据中包含的数据边界点,为每一文本数据的数据边界点上建立边界索引,得到每一文本数据对应的边界索引集;

步骤42:获取所述开发平台中每一文本框可填充的文本数据类型,为每一文本框建立数据索引;

步骤43:分别利用每一所述数据索引遍历每一文本数据对应的边界索引集,为每一数据索引匹配相应的边界索引,建立匹配列表;

步骤44:根据所述匹配列表将所述文本数据填充到开发平台中对应的文本框内进行显示。

一种基于开发平台的多语言录入和解析系统及方法

技术领域

[0001] 本发明涉及平台语言技术领域,特别涉及一种基于开发平台的多语言录入和解析方式及方法。

背景技术

[0002] 开发平台是指以某种语言或者某几种语言为基础,用户可以在这个产品上进行各种各样的软件产品的开发,并且在这个产品上进行开发的时候,只需要进行一些简单的配置,或者是写极少量的代码便可以完成一个业务系统的开发工作,但是目前不同的开发平台所使用的语言不相同,导致了在进行录入前需要先确定该开发平台的语言,加大了使用者的工作量,而且会使得开发平台的界面十分臃肿,增加了开发成本。

[0003] 因此,本发明提供了一种基于开发平台的多语言录入和解析系统及方法。

发明内容

[0004] 本发明一种基于开发平台的多语言录入和解析系统及方法,可以在开发平台中输入任何语种的文本信息,然后将其解析为统一的语言再输入到开发平台的显示页面上,打破了传统录入的弊端,实现了灵活录入。

[0005] 本发明提供了一种基于开发平台的多语言录入和解析系统及方法,包括:

[0006] 录入模块,用于录入文本信息;

[0007] 识别模块,用于获取所述文本信息的文本特征,将所述文本特征与每一语种对应的样本特征进行对比,得到所述文本信息对应的文本语种,根据所述文本语种建立译汉模型;

[0008] 解析模块,用于将所述文本信息输入到所述译汉模型中得到汉语信息,解析所述汉语信息,得到若干条文本数据;

[0009] 执行模块,用于将所述文本数据填充到开发平台中对应的文本框内进行显示。

[0010] 在一种可实施的方式中,

[0011] 还包括:

[0012] 备案模块,用于根据所述文本数据在开发平台中文本框中的填充结果,得到汉语显示版图,根据所述文本语种将所述汉语显示版图翻译为目标显示版图;

[0013] 显示模块,用于显示所述汉语显示版图和目标显示版图。

[0014] 在一种可实施的方式中,

[0015] 所述录入模块,包括:

[0016] 输入单元,用于供用户输入原始文本;

[0017] 筛选单元,用于将所述原始文本进行文本筛选,得到目标文本;

[0018] 修正单元,用于根据时间顺序调整所述目标文本,得到文本信息。

[0019] 在一种可实施的方式中,

[0020] 所述识别模块,包括:

- [0021] 训练单元,用于将所述文本信息进行语法训练,得到若干条文本语句;
- [0022] 分析单元,用于分别提取每一文本语句对应的语句关键词,分析所述语句关键词对应的词语组成结构;
- [0023] 识别单元,用于根据所述词语组成结构得到对应语句关键词的词特征,根据所述词特征的属性,确定所述文本信息对应的文本语种;
- [0024] 建模单元,用于获取所述文本语种与汉语之间的翻译对照信息,建立译汉模型。
- [0025] 在一种可实施的方式中,
- [0026] 所述解析模块,包括:
- [0027] 文本解析单元,用于获取所述文本语种的语句结构列表,根据所述语句结构列表分析所述文本信息中包含的若干中文本语句,建立所述文本信息的语句结构框架;
- [0028] 模型解析单元,用于将所述文本信息输入到所述译汉模型中,在所述译汉模型中获取所述文本信息包含的若干条文本语句,获取每一文本语句对应的若干个汉语词;
- [0029] 文本修正单元,用于将所述汉语词输入到所述语句结构框架中进行语句重组,得到若干条第一汉语语句,当所述第一汉语语句中存在病句时,分析每一第一汉语语句对应的语句问题,并进行语句修正,得到第二汉语语句;
- [0030] 文本转换单元,用于根据所述第二汉语语句建立汉语信息,得到若干条文本数据。
- [0031] 在一种可实施的方式中,
- [0032] 所述执行模块,包括:
- [0033] 第一执行单元,用于解析所述文本数据,得到每一文本数据中包含的数据边界点,为每一文本数据的数据边界点上建立边界索引,得到每一文本数据对应的边界索引集;
- [0034] 第二执行单元,用于获取所述开发平台中每一文本框可填充的文本数据类型,为每一文本框建立数据索引;
- [0035] 第三执行单元,用于分别利用每一所述数据索引遍历每一文本数据对应的边界索引集,为每一数据索引匹配相应的边界索引,建立匹配列表;
- [0036] 第四执行单元,用于根据所述匹配列表将所述文本数据填充到开发平台中对应的文本框内进行显示。
- [0037] 本发明提供了一种基于开发平台的多语言录入和解析方法,包括:
- [0038] 步骤1:录入文本信息;
- [0039] 步骤2:获取所述文本信息的文本特征,将所述文本特征与每一语种对应的样本特征进行对比,得到所述文本信息对应的文本语种,根据所述文本语种建立译汉模型;
- [0040] 步骤3:将所述文本信息输入到所述译汉模型中得到汉语信息,解析所述汉语信息,得到若干条文本数据;
- [0041] 步骤4:将所述文本数据填充到开发平台中对应的文本框内进行显示。
- [0042] 在一种可实施的方式中,
- [0043] 所述步骤2包括:
- [0044] 步骤21:将所述文本信息进行语法训练,得到若干条文本语句;
- [0045] 步骤22:分别提取每一文本语句对应的语句关键词,分析所述语句关键词对应的词语组成结构;
- [0046] 步骤23:根据所述词语组成结构得到对应语句关键词的词特征,根据所述词特征

的属性,确定所述文本信息对应的文本语种;

[0047] 步骤24:获取所述文本语种与汉语之间的翻译对照信息,建立译汉模型。

[0048] 在一种可实施的方式中,

[0049] 所述步骤3包括:

[0050] 步骤31:获取所述文本语种的语句结构列表,根据所述语句结构列表分析所述文本信息中包含的若干中文本语句,建立所述文本信息的语句结构框架;

[0051] 步骤32:将所述文本信息输入到所述译汉模型中,在所述译汉模型中获取所述文本信息包含的若干条文本语句,获取每一文本语句对应的若干个汉语词;

[0052] 步骤33:将所述汉语词输入到所述语句结构框架中进行语句重组,得到若干条第一汉语语句,当所述第一汉语语句中存在病句时,分析每一第一汉语语句对应的语句问题,并进行语句修正,得到第二汉语语句;

[0053] 步骤34:根据所述第二汉语语句建立汉语信息,得到若干条文本数据。

[0054] 在一种可实施的方式中,

[0055] 所述步骤4包括:

[0056] 步骤41:解析所述文本数据,得到每一文本数据中包含的数据边界点,为每一文本数据的数据边界点上建立边界索引,得到每一文本数据对应的边界索引集;

[0057] 步骤42:获取所述开发平台中每一文本框可填充的文本数据类型,为每一文本框建立数据索引;

[0058] 步骤43:分别利用每一所述数据索引遍历每一文本数据对应的边界索引集,为每一数据索引匹配相应的边界索引,建立匹配列表;

[0059] 步骤44:根据所述匹配列表将所述文本数据填充到开发平台中对应的文本框内进行显示。

[0060] 本发明可以实现的有益效果为:为了使开发平台适应不同的语言,首先确定录入的文本信息的语种,为了方便操作将文本信息翻译成汉语,从而可以得到若干条文本数据,最后将文本数据填充到文本框内,实现了自动填充,使用者可以在开发平台的显示部读取信息,打破了传统操作中需要提前翻译语言的弊端,实现了灵活录入,节约了开发成本,提高了开发效率。

[0061] 本发明的其它特征和优点将在随后的说明书中阐述,并且,部分地从说明书中变得显而易见,或者通过实施本发明而了解。本发明的目的和其他优点可通过在所写的说明书、权利要求书、以及附图中所特别指出的结构来实现和获得。

[0062] 下面通过附图和实施例,对本发明的技术方案做进一步的详细描述。

附图说明

[0063] 附图用来提供对本发明的进一步理解,并且构成说明书的一部分,与本发明的实施例一起用于解释本发明,并不构成对本发明的限制。在附图中:

[0064] 图1为本发明实施例中一种基于开发平台的多语言录入和解析系统的组成示意图;

[0065] 图2、图3为本发明实施例中一种基于开发平台的多语言录入和解析系统及方法的开发平台的界面显示示意图;

[0066] 图4为本发明实施例中一种基于开发平台的多语言录入和解析方法的工作流程示意图。

具体实施方式

[0067] 以下结合附图对本发明的优选实施例进行说明,应当理解,此处所描述的优选实施例仅用于说明和解释本发明,并不用于限定本发明。

[0068] 实施例1

[0069] 本实施例提供了一种基于开发平台的多语言录入和解析系统及方法,如图1所示,包括:

[0070] 录入模块,用于录入文本信息;

[0071] 识别模块,用于获取所述文本信息的文本特征,将所述文本特征与每一语种对应的样本特征进行对比,得到所述文本信息对应的文本语种,根据所述文本语种建立译汉模型;

[0072] 解析模块,用于将所述文本信息输入到所述译汉模型中得到汉语信息,解析所述汉语信息,得到若干条文本数据;

[0073] 执行模块,用于将所述文本数据填充到开发平台中对应的文本框内进行显示。

[0074] 该实例中,文本特征表示可以区分不同文本信息的特征;

[0075] 该实例中,样本特征表示用来区分不同语种,可以表现该语种特点的特征;

[0076] 该实例中,文本语种表示文本信息所对应的语种;

[0077] 该实例中,译汉模型表示将不同语种翻译成汉语的模型;

[0078] 该实例中,开发平台的界面显示为普通文本输入框,如图2所示,聚焦输入框时显示下拉面板,在面板中录入多种语言文本,如图3所示,将两种语言保存到一个数据对象中,界面显示时根据当前系统语言环境,显示对应的语言下的文本内容。

[0079] 上述技术方案的工作原理以及有益效果:为了使开发平台适应不同的语言,首先确定录入的文本信息的语种,为了方便操作将文本信息翻译成汉语,从而可以得到若干条文本数据,最后将文本数据填充到文本框内,实现了自动填充,使用者可以在开发平台的显示部读取信息,打破了传统操作中需要提前翻译语言的弊端,实现了灵活录入,节约了开发成本,提高了开发效率。

[0080] 实施例2

[0081] 在实施例1的基础上,所述一种基于开发平台的多语言录入和解析系统,还包括:

[0082] 备案模块,用于根据所述文本数据在开发平台中文本框中的填充结果,得到汉语显示版图,根据所述文本语种将所述汉语显示版图翻译为目标显示版图;

[0083] 显示模块,用于显示所述汉语显示版图和目标显示版图。

[0084] 该实例中,汉语显示版图表示将文本信息转换为汉语信息后填充到文本框内的结果;

[0085] 该实例中,目标显示版图表示将汉语显示版图转换为与文本信息语种一致的版图。

[0086] 上述技术方案的工作原理以及有益效果:为了实现多元化显示,在开发平台中不仅显示汉语版图,还显示文本语种的版图,这样一来使用者可以根据自己的需求调取不同

的版图,满足了不同使用者的使用需求。

[0087] 实施例3

[0088] 在实施例1的基础上,所述一种基于开发平台的多语言录入和解析系统,所述录入模块,包括:

[0089] 输入单元,用于供用户输入原始文本;

[0090] 筛选单元,用于将所述原始文本进行文本筛选,得到目标文本;

[0091] 修正单元,用于根据时间顺序调整所述目标文本,得到文本信息。

[0092] 上述技术方案的工作原理以及有益效果:为了避免用户输入的文本中含有冗余数据,导致文本分析量变大,所以在用户输入原始文本后对文本进行筛选,然后再根据时间顺序对其进行调整,最后可以得到文本信息简化了文本,提高了文本的解析效率。

[0093] 实施例4

[0094] 在实施例1的基础上,所述一种基于开发平台的多语言录入和解析系统,所述识别模块,包括:

[0095] 训练单元,用于将所述文本信息进行语法训练,得到若干条文本语句;

[0096] 分析单元,用于分别提取每一文本语句对应的语句关键词,分析所述语句关键词对应的词语组成结构;

[0097] 识别单元,用于根据所述词语组成结构得到对应语句关键词的词特征,根据所述词特征的属性,确定所述文本信息对应的文本语种;

[0098] 建模单元,用于获取所述文本语种与汉语之间的翻译对照信息,建立译汉模型。

[0099] 该实例中,语法训练表示将文本信息中所包含的文本划分为若干个句子,以及分析每一个句子的句子结构的过程;

[0100] 该实例中,语句关键词表示在一条文本语句中,可以表示该语句含义的词语;

[0101] 该实例中,词语组成结构包括语句关键词的词性、词义以及该语句关键词在对应语句中的位置,以及该位置前后词语的词性、词义;

[0102] 该实例中,翻译对照信息表示将文本语种翻译为汉语时的对照表,其中包括字翻译、词翻译、固定句子翻译。

[0103] 上述技术方案的工作原理以及有益效果:在进行语言解析时,先将文本信息进行语法训练可以得到若干条文本语句,然后根据其语句关键词的词语组成结构来分析其词特征,从而可以确定该文本信息的文本语种,这样一来就可以建立一个由文本语种翻译为汉语的模型,为后续进行语言解析做基础。

[0104] 实施例5

[0105] 在实施例1的基础上,所述一种基于开发平台的多语言录入和解析系统,所述解析模块,包括:

[0106] 文本解析单元,用于获取所述文本语种的语句结构列表,根据所述语句结构列表分析所述文本信息中包含的若干中文本语句,建立所述文本信息的语句结构框架;

[0107] 模型解析单元,用于将所述文本信息输入到所述译汉模型中,在所述译汉模型中获取所述文本信息包含的若干条文本语句,获取每一文本语句对应的若干个汉语词;

[0108] 文本修正单元,用于将所述汉语词输入到所述语句结构框架中进行语句重组,得到若干条第一汉语语句,当所述第一汉语语句中存在病句时,分析每一第一汉语语句对应

的语句问题,并进行语句修正,得到第二汉语语句;

[0109] 文本转换单元,用于根据所述第二汉语语句建立汉语信息,得到若干条文本数据。

[0110] 该实例中,语句结构列表表示使用不同文本语种时的造句结构统计表,例如:就中文与英文而举例,英文的语句结构为倒装句;

[0111] 该实例中,语句结构框架表示文本信息中语句结构的组成,例如:一个文本信息中第一句陈述句,第二句为递进句,第三句为疑问句;

[0112] 该实例中,汉语词表示将文本语句中的所有词语翻译成汉语后的结果;

[0113] 该实例中,语句重组表示根据现有的语句结构框架来调整汉语词排列顺序的结果。

[0114] 上述技术方案的工作原理以及有益效果:为了避免不同语种的语言习惯导致翻译结果有误,在进行翻译时先根据文本语种的语句结构列表来建立文本信息的结构框架,然后将译汉模型中输出的汉语词输入到结构框架上进行语句重组,将重组后的语句进行病句修正,最后可以建立汉语信息,从而得到若干条文本数据,这样一来不仅实现了语种翻译,还提高了翻译的精确度,避免出现语义歧义导致显示结果错误,为使用者提供了精确的显示结果。

[0115] 实施例6

[0116] 在实施例1的基础上,所述一种基于开发平台的多语言录入和解析系统,所述执行模块,包括:

[0117] 第一执行单元,用于解析所述文本数据,得到每一文本数据中包含的数据边界点,为每一文本数据的数据边界点上建立边界索引,得到每一文本数据对应的边界索引集;

[0118] 第二执行单元,用于获取所述开发平台中每一文本框可填充的文本数据类型,为每一文本框建立数据索引;

[0119] 第三执行单元,用于分别利用每一所述数据索引遍历每一文本数据对应的边界索引集,为每一数据索引匹配相应的边界索引,建立匹配列表;

[0120] 第四执行单元,用于根据所述匹配列表将所述文本数据填充到开发平台中对应的文本框内进行显示。

[0121] 该实例中,数据边界点表示一个文本数据中的首位数据、末尾数据以及空位数据;

[0122] 该实例中,边界索引表示文本数据中的数据停顿标志;

[0123] 该实例中,数据索引表示文本框可填充的文本数据类型的标志;

[0124] 该实例中,为每一数据索引匹配相应的边界索引的过程表示分析文本数据的数据类型是否符合文本数据类型的过程。

[0125] 上述技术方案的工作原理以及有益效果:为了使文本数据准确无误的填充到文本框中,先分析文本数据的数据边界点,为其建立边界索引,以及根据每一文本框可填充的文本数据类型为其建立数据索引,然后利用边界索引和数据索引的配对情况进行数据填充,这样一来不仅提高了填充效率,还确保了填充的准确性,为用户提供良好的使用感受。

[0126] 实施例7

[0127] 本发明提供了一种基于开发平台的多语言录入和解析方法,如图4所示,包括:

[0128] 步骤1:录入文本信息;

[0129] 步骤2:获取所述文本信息的文本特征,将所述文本特征与每一语种对应的样本特

征进行对比,得到所述文本信息对应的文本语种,根据所述文本语种建立译汉模型;

[0130] 步骤3:将所述文本信息输入到所述译汉模型中得到汉语信息,解析所述汉语信息,得到若干条文本数据;

[0131] 步骤4:将所述文本数据填充到开发平台中对应的文本框内进行显示。

[0132] 该实例中,文本特征表示可以区分不同文本信息的特征;

[0133] 该实例中,样本特征表示用来区分不同语种,可以表现该语种特点的特征;

[0134] 该实例中,文本语种表示文本信息所对应的语种;

[0135] 该实例中,译汉模型表示将不同语种翻译成汉语的模型;

[0136] 该实例中,开发平台的界面显示为普通文本输入框,如图2所示,聚焦输入框时显示下拉面板,在面板中录入多种语言文本,如图3所示,将两种语言保存到一个数据对象中,界面显示时根据当前系统语言环境,显示对应的语言下的文本内容。

[0137] 上述技术方案的工作原理以及有益效果:为了使开发平台适应不同的语言,首先确定录入的文本信息的语种,为了方便操作将文本信息翻译成汉语,从而可以得到若干条文本数据,最后将文本数据填充到文本框内,实现了自动填充,使用者可以在开发平台的显示部读取信息,打破了传统操作中需要提前翻译语言的弊端,实现了灵活录入,节约了开发成本,提高了开发效率。

[0138] 实施例8

[0139] 在实施例7的基础上,所述一种基于开发平台的多语言录入和解析方法,所述步骤2包括:

[0140] 步骤21:将所述文本信息进行语法训练,得到若干条文本语句;

[0141] 步骤22:分别提取每一文本语句对应的语句关键词,分析所述语句关键词对应的词语组成结构;

[0142] 步骤23:根据所述词语组成结构得到对应语句关键词的词特征,根据所述词特征的属性,确定所述文本信息对应的文本语种;

[0143] 步骤24:获取所述文本语种与汉语之间的翻译对照信息,建立译汉模型。

[0144] 该实例中,语法训练表示将文本信息中所包含的文本划分为若干个句子,以及分析每一个句子的句子结构的过程;

[0145] 该实例中,语句关键词表示在一条文本语句中,可以表示该语句含义的词语;

[0146] 该实例中,词语组成结构包括语句关键词的词性、词义以及该语句关键词在对应语句中的位置,以及该位置前后词语的词性、词义;

[0147] 该实例中,翻译对照信息表示将文本语种翻译为汉语时的对照表,其中包括字翻译、词翻译、固定句子翻译。

[0148] 上述技术方案的工作原理以及有益效果:在进行语言解析时,先将文本信息进行语法训练可以得到若干条文本语句,然后根据其语句关键词的词语组成结构来分析其词特征,从而可以确定该文本信息的文本语种,这样一来就可以建立一个由文本语种翻译为汉语的模型,为后续进行语言解析做基础。

[0149] 实施例9

[0150] 在实施例7的基础上,所述一种基于开发平台的多语言录入和解析方法,其特征在于,所述步骤3包括:

[0151] 步骤31:获取所述文本语种的语句结构列表,根据所述语句结构列表分析所述文本信息中包含的若干中文本语句,建立所述文本信息的语句结构框架;

[0152] 步骤32:将所述文本信息输入到所述译汉模型中,在所述译汉模型中获取所述文本信息包含的若干条文本语句,获取每一文本语句对应的若干个汉语词;

[0153] 步骤33:将所述汉语词输入到所述语句结构框架中进行语句重组,得到若干条第一汉语语句,当所述第一汉语语句中存在病句时,分析每一第一汉语语句对应的语句问题,并进行语句修正,得到第二汉语语句;

[0154] 步骤34:根据所述第二汉语语句建立汉语信息,得到若干条文本数据。

[0155] 该实例中,语句结构列表表示使用不同文本语种时的造句结构统计表,例如:就中文与英文而举例,英文的语句结构为倒装句;

[0156] 该实例中,语句结构框架表示文本信息中语句结构的组成,例如:一个文本信息中第一句陈述句,第二句为递进句,第三句为疑问句;

[0157] 该实例中,汉语词表示将文本语句中的所有词语翻译成汉语后的结果;

[0158] 该实例中,语句重组表示根据现有的语句结构框架来调整汉语词排列顺序的结果。

[0159] 上述技术方案的工作原理以及有益效果:为了避免不同语种的语言习惯导致翻译结果有误,在进行翻译时先根据文本语种的语句结构列表来建立文本信息的结构框架,然后将译汉模型中输出的汉语词输入到结构框架上进行语句重组,将重组后的语句进行病句修正,最后可以建立汉语信息,从而得到若干条文本数据,这样一来不仅实现了语种翻译,还提高了翻译的精确度,避免出现语义歧义导致显示结果错误,为使用者提供了精确的显示结果。

[0160] 实施例10

[0161] 在实施例7的基础上,所述一种基于开发平台的多语言录入和解析方法,所述步骤4包括:

[0162] 步骤41:解析所述文本数据,得到每一文本数据中包含的数据边界点,为每一文本数据的数据边界点上建立边界索引,得到每一文本数据对应的边界索引集;

[0163] 步骤42:获取所述开发平台中每一文本框可填充的文本数据类型,为每一文本框建立数据索引;

[0164] 步骤43:分别利用每一所述数据索引遍历每一文本数据对应的边界索引集,为每一数据索引匹配相应的边界索引,建立匹配列表;

[0165] 步骤44:根据所述匹配列表将所述文本数据填充到开发平台中对应的文本框内进行显示。

[0166] 该实例中,数据边界点表示一个文本数据中的首位数据、末尾数据以及空位数据;

[0167] 该实例中,边界索引表示文本数据中的数据停顿标志;

[0168] 该实例中,数据索引表示文本框可填充的文本数据类型的标志;

[0169] 该实例中,为每一数据索引匹配相应的边界索引的过程表示分析文本数据的数据类型是否符合文本数据类型的过程。

[0170] 上述技术方案的工作原理以及有益效果:为了使文本数据准确无误的填充到文本框中,先分析文本数据的数据边界点,为其建立边界索引,以及根据每一文本框可填充的文

本数据类型为其建立数据索引,然后利用边界索引和数据索引的配对情况进行数据填充,这样一来不仅提高了填充效率,还确保了填充的准确性,为用户提供良好的使用感受。

[0171] 实施例11

[0172] 在实施例9的基础上,所述一种基于开发平台的多语言录入和解析方法,包括:

[0173] 获取所述文本语句与第一汉语语句;

[0174] 将所述文本语句与第一汉语语句进行语句匹配,得到每一第一汉语语句对应的源语句;

[0175] 根据公式(1)将所述第一汉语语句与所述源语句的语句含义相似度;

$$[0176] \quad D = \frac{\sqrt{\left(\alpha \frac{\sum_{i=1}^n X_i}{n}\right)^2 - \left(\beta \frac{\sum_{k=1}^m Y_k}{m}\right)^2}}{|m - n|}$$

[0177] 其中,D表示所述第一汉语语句与所述源语句的语句含义相似度, α 表示所述第一汉语语句的句式参数, β 表示所述源语句的句式参数,且 $\alpha + \beta = 1$, X_i 表示所述第一汉语语句中第*i*个词语对应的词义, n 表示所述第一汉语语句中词语的总数, Y_k 表示所述源语句中第*k*个词语对应的词义, m 表示所述源语句中词语的总数;

[0178] 根据公式(1)的计算结果,提取语句含义相似度低于预设相似度的第一汉语语句,记作病句。

[0179] 上述技术方案的工作原理以及有益效果:为了避免翻译过程中出现错翻的现象,将翻译后的第一汉语语句与源语句进行比较,分析二者的语句含义是否相同,及时发现存在问题的语句并进行调整,保证后续填充结果的精确度。

[0180] 显然,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样,倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内,则本发明也意图包含这些改动和变型在内。

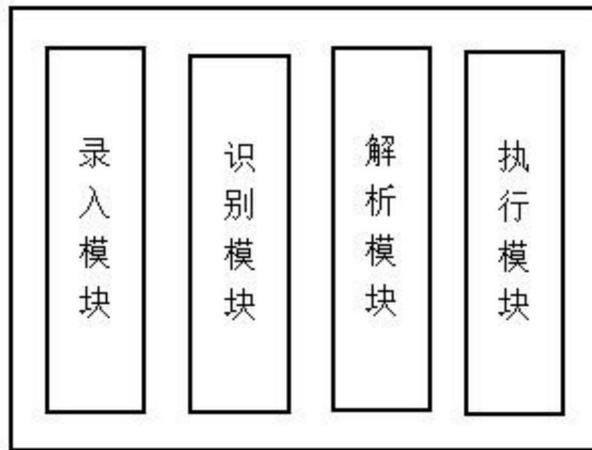


图1

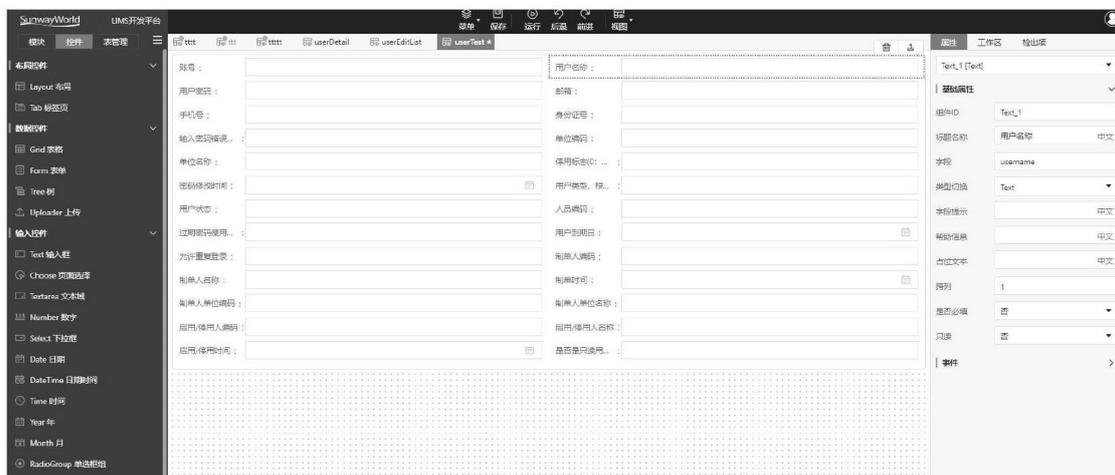


图2

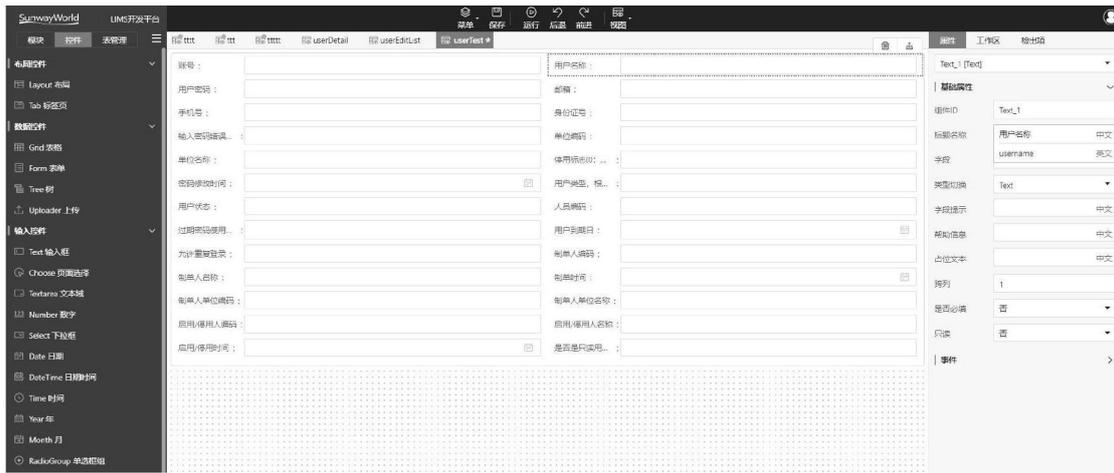


图3

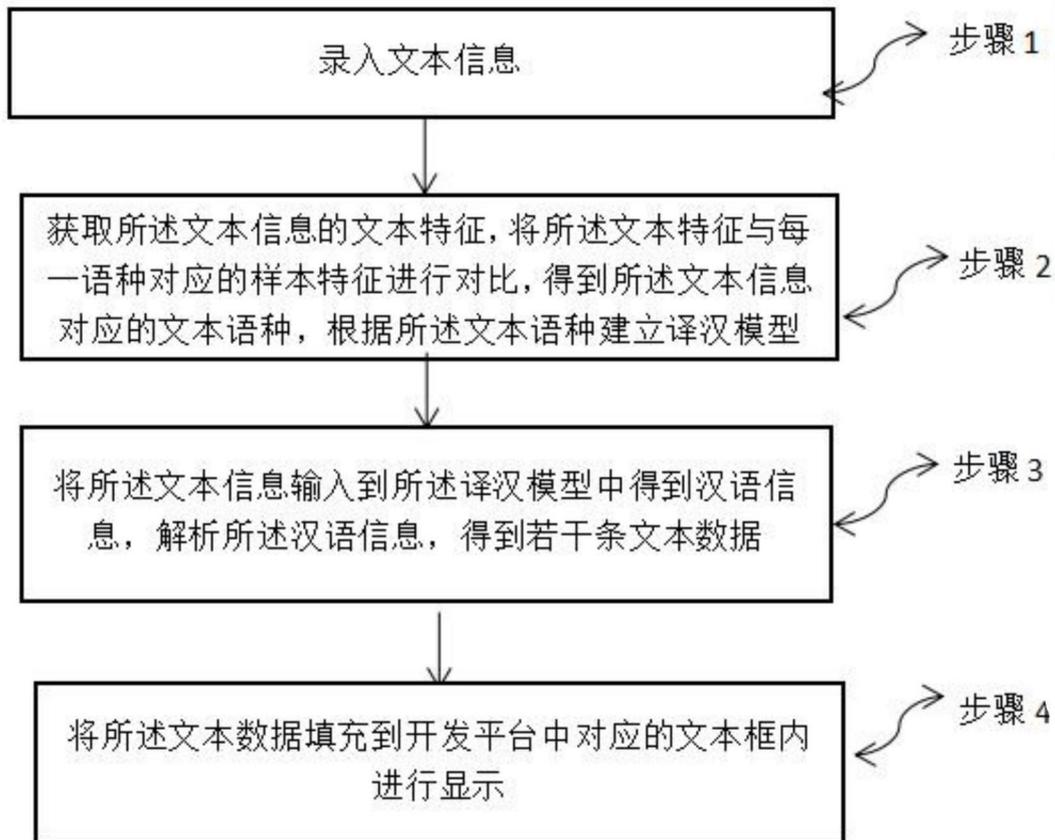


图4