

# 信息检索第五次作业报告

罗峻骁 计75

2017011364

## 摘要

本次作业基于第三次作业的检索系统，在词语及词性检索及位置约束的基础上，对文档进行了依存关系分析 (dependency parsing)，使得可以对被检索的词语指定其与其它词语的依存关系（主谓关系、动宾关系等）。

## 依存关系分析

使用了 [LTP](#) 工具对文档进行依存关系分析。为了使新工具达到尽量好的分析效果，没有使用第三次作业中的分词结果，而是使用此工具对文档重新进行分词，然后再在此基础上进行词性标注及依存关系分析。

有关依存分析的一些介绍[见此](#)。

## 功能实现

创建的索引仍然只有类型为 `text` 的 `content` 这一属性。此前为了支持词语及词性检索，通过恰当地设计 token filter，对于同一个词语，有如下三种元组用于建倒排索引：

- (词语, 词性)
- (词语,)
- (词性,)

为了保持这些功能，并增加新的依存关系检索功能，在保留上述这些元组的情况下，借助了 [pattern replace](#) 与 [multiplexer](#) 又增加了下面几种元组：

- (词语, 词性, 依赖头, 依赖关系)
- (词语, 依赖头, 依赖关系)
- (词性, 依赖头, 依赖关系)

此外，第三次作业中元组的存储通过下划线连接各部分，但其实分词后的文本中依然存在下划线。本次作业中使用了缩进字符 `\t` 来连接元组中的各部分，这样做的好处是缩进字符为空白字符，一定不存在于分词后的文本中，也就没有了歧义。

具体配置见 `word/index-config.json`。

搜索时，只需将用户输入拼接即可，基本与此前的实现类似。

## 操作界面

在保持原有功能输入的基础上，网页界面中增加了可以输入依存关系中依赖头和选择依赖关系的部分：

添加词语添加约束进行搜索

1	输入词语	选择词性	输入依赖头	选择依赖关系	删除	编号1	编号2	相邻	删除
排名	内容								评分
<div>选择依赖关系 主谓关系 动宾关系 间宾关系 前置宾语 兼语 定中关系 状中结构 动补结构 并列关系 介宾关系 左附加关系 右附加关系 独立结构 核心关系</div>									

## 结果展示

下面的例子中搜索 我 这个代词，并希望其与 看 字形成 主谓关系：

添加词语添加约束进行搜索

1	我	pronoun	看	主谓关系	删除	编号1	编号2	相邻	删除
---	---	---------	---	------	----	-----	-----	----	----

排名	内容	评分
1	我看着波罗的海碧蓝色的水面，我看着平静无云的海空，我看着海边开放的朵朵美丽的花丛，于是，德国人民心灵的欢呼又浮现在我的脑海来了：波罗的海，和平的海！	13.994086
2	然而吴先生不为所动，针锋相对地说：“你们这个东西我不看，我越看会越糊涂。	13.747726
3	我看过以连唱一百多句取胜的《对花枪》，我也看过以连场武打取胜的《对花枪》。	13.511891
4	“国安酒我看世界杯，我看中国队”有奖征文活动日前在京揭晓，刘德伟的《要真正做好“从娃娃抓起”》获特等奖，另有十位参与者的征文获得优秀奖。	13.094715
5	放眼看那座座高楼如同那稻麦，看眼前是人的海洋和交通的堵塞，我左看右看前看后可还是看不过来，这个那个我越看越奇怪。	13.094715
6	最近我看了湖南青年美展，感到其中不少作品是动情的，作者们征求我的意见，我凭直觉留言：“有脉搏、有血色、有泥土气息，因为我看过太多苍白的作品。	12.971487
7	我看三峡工程	12.697226
8	我看不尽然。	12.697226
9	我看是有。	12.639013
10	我看呆了。	12.639013

« 1 2 3 4 5 ... 71 »

Go

可以看出，结果中的句子都存在 我 与 看 构成主谓关系的成分。在第二个与第五个结果中，有 我不看 和 我越看 这样的情况，虽然这两个词没有相邻，但也构成诸位关系，也被检索了出来。

而当关系设为 动宾关系 时：

添加词语

添加约束

进行搜索

1

我

proun

看

动宾关系

删除

编号1

编号2

相邻

删除

排名	内容	评分
1	“导演陈颢嘶哑着嗓子喊着，一边用眼睛捕捉着一个个演员的神情：“所有的演员，都看着我，看着我！”	17.377151
2	惊看我	16.485403
3	玉柳不解地看着我。	16.185831
4	“看我的，跟我来。	16.04009
5	看我，就没为这个不高兴。	15.968201
6	写到这里，一个织布女工来看我。	15.896952
7	听说我病了，经理和门市部主任都来看我。	15.68697
8	学校老师们知道我生病了，都跑来看我。	15.68697
9	‘她倚着钢琴站着，眼睛直看着我。	15.68697
10	高泉楞住了，睁大眼睛莫名其妙地看着我。	15.68697

«

1

2

3

4

5

...

8

»

Go

可见检索结果中都存在 我 与 看 构成动宾关系的成分，如 看我， 看着我 等。