|  |
| --- |
|  |
| HW2 |
|  |
| PB15111604金泽文 |

[**Web Information Processing and Applications**](http://staff.ustc.edu.cn/~jpq/courses/webinfo.html)

HW2

PB15111604金泽文

1. 假设词典中包括词｛王公，公子，研究，研究生，生命，起源｝以及所有单字集合，请分别给出句子“王公子在研究生命的起源”的FMM和BMM分词结果。

*Answer：*

*FMM - 王公 / 子 / 在 / 研究生 / 命 / 的 / 起源*

*BMM - 王 / 公子 / 在 / 研究 / 生命 / 的 / 起源*

1. 下面的句子存在哪种类型的分词歧义？为什么？
   1. 吉林省长春药店

*1. 吉林 / 省长 / 春药 / 店*

*2. 吉林省长 / 春药 / 店*

*3. 吉林省 / 长春 / 药店*

*1和2属于组合歧义: 吉林,省长,吉林省长都是词典的词*

*1和3,2和3属于交叉歧义: 吉林,省长,春药,店,吉林省,长春,药店,吉林省长,都是词典的词.*

* 1. 东北大学生联合会

*1. 东北 / 大学生 / 联合会*

*2. 东北大学 / 生 / 联合会*

*3. 东北 / 大学 / 生 / 联合会*

*1和2属于交叉歧义: 东北,大学生,东北大学,生都是词典的词*

*2和3属于组合歧义:东北,大学,东北大学都是词典的词.*

* 1. 人大代表群体性事件

*1. 人大 / 代表 / 群体 / 性 / 事件*

*2. 人大代表 / 群体性事件*

*3. 人大代表 / 群体 / 性事 / 件*

*4. 人大代表 / 群体性 / 事件*

*1和2属于组合歧义: 人大,代表,群体,性,事件人大代表,群体性事件都是词典的词.*

*3和4属于交叉歧义: 群体性,事件,群体,性事,件都是词典的词.*