

南京农业大学
2007 年攻读硕士学位研究生入学考试试题

试题编号: 433 试题名称: 信息检索

注意: 答题一律答在答题纸上, 答在草稿纸或试卷上一律无效

一、名词解释 (每小题 3 分, 总 30 分)

- | | |
|-------------------|------------|
| 1. 排除词词典 | 2. 索引膨胀系数 |
| 3. 元搜索引擎 | 4. 萨尔顿 |
| 5. 平均查准率 (十一点插值法) | 6. 截词检索 |
| 7. 多媒体数据压缩比 | 8. 类重心 |
| 9. TREC | 10. 非相关检出率 |

二、名词辨析 (每小题 5 分, 总 25 分)

- 1、一次文献、二次文献、三次文献
- 2、目录、题录、目次
- 3、顺查法、倒查法、抽查法
- 4、文献检索、数据检索、事实检索
- 5、标题词、叙词、元词

三、试构建下列课题的布尔逻辑式 (15 分)

- (1) 抗坏血酸对兔子红细胞的影响
- (2) 冠状病毒引发病毒性肝炎的情况
- (3) 寄生蝇类的对比性研究
- (4) 水稻稻瘟病的基因差异
- (5) 西兰花高原栽培技术
- (6) 欧洲玉米螟幼虫死亡率的影响因素
- (7) 小麦蛋白质的人工选择
- (8) SARS 的致病因素
- (9) 禽流感的病毒变异
- (10) 去草胺除向日葵杂草

四、简答题 (每小题 10 分, 总 50 分)

- 1、什么是类? 给出类的几种形式化定义。
- 2、试分析比较绝对频率加权法、文内频率加权法、文外频率加权法、逆文献频率加权法与 tfx 加权公式。
- 3、为什么说查全率与查准率是独立于用户的?
- 4、机械分词法的优缺点。
- 5、等级聚类法的基本思想与算法效率分析。

三、论述题 (每小题 15 分, 总 30 分)

- 1、谈谈你所熟悉的中外文数据库 (十种以上)
- 2、谈谈向量空间模型中降维的意义, 请根据你的所能设计几种可行的降维方法。