

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，共 30 分）**

- |         |                |
|---------|----------------|
| 1. 追溯法  | 6. review      |
| 2. 三次文献 | 7. E. Garfield |
| 3. 概念组配 | 8. 事实检索        |
| 4. 字面成族 | 9. 目录服务        |
| 5. 分组标引 | 10. 原文传递       |

**二. 简答题（每小题 10 分，共 70 分）**

- 1、简述叙词语言词间关系控制的主要手段及方法。
- 2、分类法和主题法结构的兼容性主要表现在哪些方面？
- 3、简要介绍《中图法》（2—4 版）索引的类型和编制方法？
- 4、简介三个以上主要的生命科学文献数据库？
- 5、简述词区分值加权标引法的算法思想。
- 6、在全文检索中如何用排除词词典提高查准率？
- 7、基于文本的多媒体信息检索方法的原理、优点与局限性。

**三. 论述题（1、2 题每题 15 分，第 3 题 20 分，共 50 分）**

- 1、阐明文献数据库标引质量的定量测评指标和方法。
- 2、有一位老师正在从事棉花纤维（cotton fiber）品质方面的研究课题，需要查找近五年内的中外文资料，请为他设计一份详细的文献调研计划（包括使用的主要中外文检索工具，检索步骤，适合的检索策略，检索时应注意的问题等）。
- 3、词典分词法、部件词典法、切分标记法、语法分析法是目前可行的几种汉语语词切分方法，请逐一描述其算法原理。词典分词法的最大弊端是分词词典不能适应新词变化，你能不能提出某些改进措施，使得词典分词法也能切分出某些原分词词典中没有的、新出现的汉语词汇。