

840 《数学-数据方向基础综合》考试大纲

840 《数学-数据方向基础综合》适用于以下专业及研究方向：

招生专业代码及方向	硕士研究生方向
	MS Research Area
0812J3 数据科学与信息技术 01(全日制) 数据科学交叉学科	工业工程与运筹学方向 Industrial Engineering and Operational Research
	智能交通与物流系统方向 Intelligent Transportation and Logistics Systems
	低碳经济与金融风险分析研究方向 Low Carbon Economy and Financial Risk Analysis Research
	智能电网与可再生能源方向 Smart Grid and Renewable Energy
	大数据与人工智能方向 Big data and artificial intelligence
	计算机系统研究 Computer Systems
	传感与测控方向 Sensor, measurement and control
	光电子方向 Photonics

注意：2025 级硕士招生目录及招生人数请以当年清华大学研究生招生网公布的为准。

一. 学科概述

“数据科学和信息技术”是清华大学自主设置的交叉学科，面向未来的社会发展需求并已按国家有关文件要求完成备案的新型学科。

欢迎但不限于以下专业背景的同学报考：电子科学与技术、信息与通信工程、计算机科学与技术、电气工程、动力工程及工程热物理、光学、应用经济学、数学、物理、化学、仪器科学与技术、机械工程、控制科学与工程、土木工程、管理科学与工程、航空宇航科学与技术、社会学等。

二. 参考书目：

《数据结构》（C 语言版）（严蔚敏、吴伟民 清华大学出版社）

三. 考试内容：

- 1.1 什么是数据结构
- 1.2 基本概念和术语
- 1.3 抽象数据类型的表示与实现

1.4 算法和算法分析

1.4.1 算法

1.4.2 算法设计的要求

1.4.3 算法效率的度量

1.4.4 算法的存储空间需求

2 线性表

2.1 线性表的类型定义

2.2 线性表的顺序表示和实现

2.3 线性表的链式表示和实现

2.3.1 线性链表

2.3.2 循环链表

2.3.3 双向链表

2.4 一元多项式的表示及相加

3 栈和队列

3.1 栈

3.1.1 抽象数据类型栈的定义

3.1.2 栈的表示和实现

3.2 栈的应用举例

3.2.1 数制转换

3.2.2 括号匹配的检验

3.2.3 行编辑程序

3.2.4 迷宫求解

3.2.5 表达式求值

3.3 栈与递归的实现

3.4 队列

3.4.1 抽象数据类型队列的定义

3.4.2 链队列——队列的链式表示和实现

3.4.3 循环队列——队列的顺序表示和实现

3.5 离散事件模拟

4 串

4.1 串类型的定义

4.2 串的顺序表示和实现

4.2.1 定长顺序存储表示

4.2.2 堆分配存储表示

4.2.3 串的块链存储表示

4.3 串的模式匹配算法

4.3.1 求子串位置的定位函数 Index (S, T, pos)

4.3.2 模式匹配的一种改进算法

4.4 串操作应用举例

4.4.1 文本编辑

4.4.2 建立词索引表

5 数组和广义表

5.1 数组的定义

5.2 数组的顺序表示和实现

5.3 矩阵的压缩存储

5.3.1 特殊矩阵

5.3.2 稀疏矩阵

5.4 广义表的定义

5.5 广义表的存储结构

5.6 元多项式的表示

5.7 广义表的递归算法

5.7.1 求广义表的深度

5.7.2 复制广义表

5.7.3 建立广义表的存储结构

6 树和二叉树

6.1 树的定义和基本术语

6.2 二叉树

6.2.1 二叉树的定义

6.2.2 二叉树的性质

6.2.3 二叉树的存储结构

6.3 遍历二叉树和线索二叉树

6.3.1 遍历二叉树

6.3.2 线索二叉树

6.4 树和森林

6.4.1 树的存储结构

6.4.2 森林与二叉树的转换

6.4.3 树和森林的遍历

6.5 树与等价问题

6.6 赫夫曼树及其应用

6.6.1 最优二叉树（赫夫曼树）

6.6.2 赫夫曼编码

6.7 回溯法与树的遍历

6.8 树的计数

7 图

7.1 图的定义和术语

7.2 图的存储结构

7.2.1 数组表示法

7.2.2 邻接表

7.2.3 十字链表

7.2.4 邻接多重表

7.3 图的遍历

7.3.1 深度优先搜索

- 7.3.2 广度优先搜索
- 7.4 图的连通性问题
 - 7.4.1 无向图的连通分量和生成树
 - 7.4.2 有向图的强连通分量
 - 7.4.3 最小生成树
 - 7.4.4 关节点和重连通分量
- 7.5 有向无环图及其应用
 - 7.5.1 拓扑排序
 - 7.5.2 关键路径
- 7.6 最短路径
 - 7.6.1 从某个源点到其余各顶点的最短路径
 - 7.6.2 每一对顶点之间的最短路径

- 8 动态存储管理
 - 8.1 概述
 - 8.2 可利用空间表及分配方法
 - 8.3 边界标识法
 - 8.3.1 可利用空间表的结构
 - 8.3.2 分配算法
 - 8.3.3 回收算法
 - 8.4 伙伴系统
 - 8.4.1 可利用空间表的结构
 - 8.4.2 分配算法
 - 8.4.3 回收算法
 - 8.5 无用单元收集
 - 8.6 存储紧缩

- 9 查找
 - 9.1 静态查找表
 - 9.1.1 顺序表的查找
 - 9.1.2 有序表的查找
 - 9.1.3 静态树表的查找
 - 9.1.4 索引顺序表的查找
 - 9.2 动态查找表
 - 9.2.1 二叉排序树和平衡二叉树
 - 9.2.2 B 树和 B+树
 - 9.2.3 键树
 - 9.3 哈希表
 - 9.3.1 什么是哈希表
 - 9.3.2 哈希函数的构造方法
 - 9.3.3 处理冲突的方法
 - 9.3.4 哈希表的查找及其分析

- 10 内部排序

- 10.1 概述
- 10.2 插入排序
 - 10.2.1 直接插入排序
 - 10.2.2 其他插入排序
 - 10.2.3 希尔排序
- 10.3 快速排序
- 10.4 选择排序
 - 10.4.1 简单选择排序
 - 10.4.2 树形选择排序
 - 10.4.3 堆排序
- 10.5 归并排序
- 10.6 基数排序
 - 10.6.1 多关键字的排序
 - 10.6.2 链式基数排序
- 10.7 各种内部排序方法的比较讨论

- 11 外部排序
 - 11.1 外存信息的存取
 - 11.2 外部排序的方法
 - 11.3 多路平衡归并的实现
 - 11.4 置换-选择排序
 - 11.5 最佳归并树

- 12 文件
 - 12.1 有关文件的基本概念
 - 12.2 顺序文件
 - 12.3 索引文件
 - 12.4 ISAM 文件和 VSAM 文件
 - 12.4.1 ISAM 文件
 - 12.4.2 VSAM 文件
 - 12.5 直接存取文件（散列文件）
 - 12.6 多关键字文件
 - 12.6.1 多重表文件
 - 12.6.2 倒排文件