

6.4

哈夫曼树

基本信息			
教学主题	哈夫曼树	课时安排	1 课时
所在章节	第 7 章第 8 小节		

【内容分析】

本次课的内容选自李春葆编著的《数据结构教程》(第 5 版)教材中第 7 章第 8 小节的内容。树结构是一种应用非常广泛的结构,在一些特定的应用中,树具有一些特殊特点,利用这些特点可以解决很多工程问题。随着大数据时代的到来,如何采用有效的数据压缩技术来节省数据文件的存储空间和网络传输时间越来越引起人们的重视。本节课所讲的哈夫曼树便可以应用于数据压缩技术中。哈夫曼树的构造算法较易理解,但其由来及原理会让学生有所困惑,导致学生出现知其然而不知其所以然的情况,因此在教学过程中应该注重介绍哈夫曼树的由来及构造原理,然后为学生介绍哈夫曼树的应用,让学生能够了解知识的产生、发展到应用的全过程。通过本讲学习,学生应该掌握哈夫曼树的概念、构造算法和哈夫曼编码算法的实现。

【教学思路】

教学思路如图 6.34 所示。

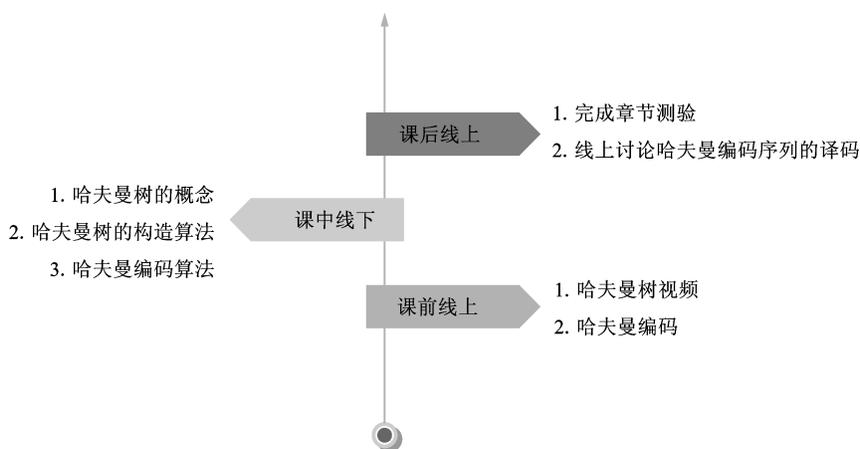


图 6.34 教学思路

一堂课是否能够吸引学生的注意力,且让学生从中获益,关键取决于导入以及案例讲解。本节课先提

出学生熟悉的成绩转换问题,让学生展示不同的判断流程,通过对比大量数据转换时不同流程所需比较次数,引出哈夫曼树的定义,并讲解哈夫曼树的特点及构造算法;然后通过讲述大数据时代,如何采用有效的数据压缩再次引入,给学生提出研讨主题,要求学生分组讨论并为一段字符给出编码方案;接着对比分析学生讨论的结果,为学生讲授哈夫曼编码;最后提出课后思考——“如何译码”,引导学生进一步思考。

▶ 【教学目标】

根据教学大纲的规定,按照质量工程的教育要求,并结合学生的实际情况,分别从知识获取、能力培养和价值塑造三个方面,确定本次课的教学目标如下:

知识目标:

- (1)理解哈夫曼树的概念及特点。
- (2)掌握哈夫曼树的构造算法。
- (3)掌握哈夫曼编码算法的实现。

能力目标:

- (1)通过学生熟知的成绩转换问题培养学生横向联系知识的能力,提升学生灵活运用能力;
- (2)通过分组讨论字符段的编码问题,将抽象知识点的被动输入学习转变为主动的探索学习,提升学生的主动思考能力和自主学习能力。

情感目标:

- (1)通过前沿热门话题的引入,让学生联系起当前学的知识,增强学生学习的动力,让学生看到希望。
- (2)通过具体的案例让学生讨论并实现,让学生在解决问题的同时获得成就感,增加学生的学习动力。

▶ 【课程思政】

一、讲人物以励志——鼓励创新精神

1951年,哈夫曼在麻省理工学院(MIT)攻读博士学位,他修读的信息论课程需要完成学期报告。导师罗伯特·法诺(Robert Fano)出的学期报告题目是:查找最有效的二进制编码。由于无法证明哪个已有编码是最有效的,于是哈夫曼放弃对已有编码的研究,转向新的探索,最终发现了基于有序频率二叉树编码的想法,并很快证明了这个方法是最有效的。哈夫曼使用自底向上的方法构建二叉树,避免了次优算法香农-范诺编码(Shannon-Fano coding)的最大弊端——自顶向下构建树。

二、节约空间,减少数字冗余

通过哈夫曼编码在压缩软件中的应用,引导同学们在平时的编程过程中尽可能地压缩和节约空间,让数字世界少一点“垃圾”。

▶ 【教学内容】

❖ 教学重点

哈夫曼树的建立及哈夫曼编码的建立。

❖ 教学难点

哈夫曼树的实现。

❖ 重点及难点的处理

通过百分制成绩转换五分制成绩的判定树引出重点,之后通过具体实例讲解哈夫曼树的建立过程,再通过具体编码实例引出哈夫曼编码的优势,然后讲解哈夫曼编码的建立。对于教学难点——哈夫曼树的实现,先通过动画讲解让学生掌握哈夫曼树的实现过程,再分析代码。

▶ 【教学方法】

一、智慧课堂辅助——线上线下混合教学法

对教学内容进行分析,合理地划分线上线下部分,让学生课前通过微课视频预习,课上利用智慧课堂辅助记录学生参与讨论情况,以便老师能够实时了解所有学生的上课状态。

二、案例分析法

通过熟悉“成绩转换”案例分析,学生能够更深刻地理解哈夫曼树的概念及原理。通过“字符编码”案例,学生掌握哈夫曼编码算法。

三、研讨式教学法

给出编码的研讨主题,让学生分组讨论,并给出编码方案,再进行组间研讨,让学生能够通过讨论过程理解哈夫曼编码的优缺点,并在不断地改良编码方案的过程中提高解决问题的能力。

四、启发式教学法

在教学过程中不断地提出问题,深化问题,让学生在逐步解决问题的过程中不断深入思考,培养学生追根究底的探索精神。

五、探究式教学法

授课采用提出问题—分析问题—解决问题—问题延伸的教学设计层层深入,引导学生在理解新知识的同时进行探究,运用所学知识解决问题,从而提高学习的自信心。

六、多媒体辅助教学法

通过精心制作的动画展示哈夫曼树的构造过程以及哈夫曼编码算法过程,使教学更直观、更生动,帮助学生理解新课程知识。

▶ 【教学过程】

教学过程见表 6.6。

表 6.6 教学过程

教学环节	教师教学	学生活动	设计意图	时间/min
引入	案例引入： 请将百分制成绩转换成五分制成绩(给出成绩在各分数段的占比)	调整状态 沉浸在课堂中	引出本节课的主要内容“哈夫曼树”，通过流程判定树的平均比较次数，让学生理解哈夫曼树的意义	2
智慧课堂	线上选人展示课前准备结果 利用“学习通”的选人活动开展随机抽取学生展示课前准备的转换方案，然后分析平均比较次数并进行对比	参与智慧课堂，展示自己的方案，与老师共同计算和分析	通过不同的对比分析，引导学生总结出规律，从而引出本节课主讲内容哈夫曼树	3
教学过程	知识讲解：哈夫曼树 知识点： (1)带权路径长度 (2)最优二叉树 (3)哈夫曼树的构造 将引入中的判断流程转化成判定树，以之解释判定树中的平均比较次数便是带权路径长度。通过一个具体实例讲解哈夫曼树的构造过程 介绍哈夫曼勇于创新的故事，讲解哈夫曼树的由来	认真听讲解，学习哈夫曼树的相关概念，参与具体实例中哈夫曼树的构造过程	1. 讲授法加举例法讲解树的带权路径长度的概念，并通过实例比较法让学生加深对哈夫曼树的理解 2. 以名人事迹激励学生勇于创新	35
	研讨主题一：给一段特定的字符编码 学生分组讨论，每组给出一种编码方案，然后再组间相互质疑和解疑，最终对比各组结果进行总结	思考多种编码方案并进行对比，分组讨论，阐述小组结果，组间研讨	通过分组讨论与组间讨论培养学生的协同合作的能力，让学生在讨论中能够总结出最佳的编码方案——哈夫曼编码	
	知识讲解：哈夫曼编码 知识点：二叉树编码 根据上一环节的研讨结果，总结常见的几种编码，对比分析讲解哈夫曼编码的优势	仔细听讲解，参与计算几种常见编码的长度，学习哈夫曼编码思想	通过多种编码方案分析对比，培养学生严谨的治学态度和创新精神	
	研讨主题二：为研讨主题一中的字符建立哈夫曼编码 观察各组的哈夫曼编码过程及结果，及时点评并讲解	分组讨论，写出讨论结果，听点评和讲解	通过组内研讨及组间比较，让学生能够总结出哈夫曼编码的特点	
教学小结	教学小结： (1)哈夫曼树的特点及建立 (2)哈夫曼编码的特点	参与小结 听讲解	帮助学生更好地梳理与掌握本次课的重点内容	2
拓展	应用拓展： (1)最佳判定树 (2)数据压缩 文献拓展： 基于哈夫曼编码的稀疏矩阵的存储与计算	了解哈夫曼树的应用，感受节约空间，减少数字冗余的重复性；课后阅读文献	1. 拓展应用知识，培养学生阅读文献的能力 2. 通过哈夫曼编码在压缩软件中的应用，倡导同学们在平时编程时应尽可能压缩和节约空间，让数字世界少一些“垃圾”	2

续表6.6

教学环节	教师教学	学生活动	设计意图	时间/min
思考	思考题: 当我们收到一串编码后的数据时如何译码呢? (编程实现编码及译码过程)	课后完成 在线讨论	1. 提升学生的编程能力 2. 督促学生思考	1

【预习与作业】

(1) 预习任务。

- ① 观看“哈夫曼树”和“哈夫曼编码”的预习视频，初步了解哈夫曼树的概念及构造。
- ② 完成案例“百分制转换成五分制”的转换方案，为课堂学习做准备。

(2) 课后作业。

- ① 完成“学习通”上的章节测验。
- ② 线上讨论哈夫曼编码序列的译码。
- ③ 阅读文献，撰写 300 字的阅读小报告。

板书设计

板书设计如图 6.35 所示。

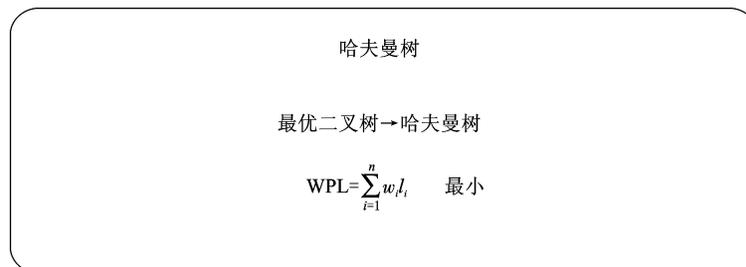


图 6.35 板书设计

教 学 过 程

◆ 课堂导入(2 分钟)

例：要求编写一个程序将百分制的考试成绩转换成五分制的成绩。各分数段占比如表 6.7 所示。

表 6.7 各分数段占比

分数段	0~59	60~69	70~79	80~89	90~100
比例/%	5	15	40	30	10

下面给出两种解决方案。

方案一(见图 6.36)：

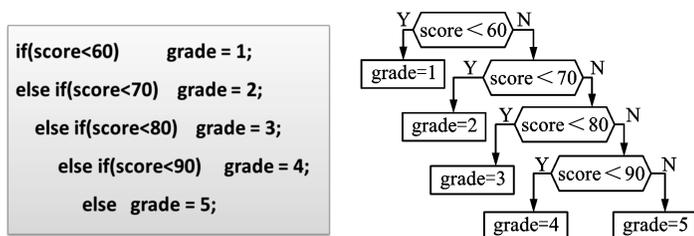


图 6.36 方案一解析图示

从判断树中可以看出，80%以上的数据需要进行三次或三次以上的比较，那么平均比较次数 = $0.05 \times 1 + 0.15 \times 2 + 0.4 \times 3 + 0.3 \times 4 + 0.1 \times 5 = 3.15$ (次)，如果有 10 0000 个输入数据需要比较 315000 次。

方案二(见图 6.37)：

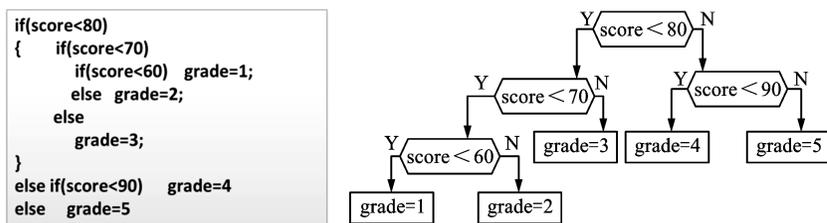


图 6.37 方案二解析图示

此方案的平均比较次数 = $0.05 \times 3 + 0.15 \times 3 + 0.4 \times 2 + 0.3 \times 2 + 0.1 \times 2 = 2.2$ (次)，如果有 10 0000 个输入数据需要比较 220000 次，相比方案一减少了 95000 次。

智慧课堂(3分钟)

线上选人展示课前准备结果。

利用“学习通”的选人活动开展随机抽取学生展示课前准备的转换方案,然后分析平均比较次数并进行对比。

思考:如何根据结点不同的查找频率构造更有效的搜索树呢?

(与学生共同总结发现:查找频率低的结点可以远离根结点,查找频率高的结点应该靠近根结点位置)

正文(35分钟)

• 知识讲解

(1)带权路径长度。

设二叉树具有 n 个带权值的叶结点,那么从根结点到各个叶结点的路径长度与相应结点权值的乘积的和,叫作二叉树的带权路径长度(WPL)。

$$WPL = \sum_{i=1}^n w_i l_i$$

(2)最优二叉树。

哈夫曼树也叫最优二叉树:WPL最小的二叉树

1951年,哈夫曼在麻省理工学院(MIT)攻读博士学位,他和修读信息论课程的同学得选择是完成学期报告还是期末考试。导师罗伯特·法诺(Robert Fano)出的学期报告题目是:查找最有效的二进制编码。由于无法证明哪个已有编码是最有效的,哈夫曼放弃对已有编码的研究,转向新的探索,最终发现了基于有序频率二叉树编码的想法,并很快证明了这个方法是最有效的。哈夫曼使用自底向上的方法构建二叉树,避免了次优算法香农-范诺编码(Shannon-Fano coding)的最大弊端——自顶向下构建树。

(3)哈夫曼树的构造。

①给定的 n 个权值 $\{W_1, W_2, \dots, W_n\}$ 构造 n 棵只有一个叶节点的二叉树,从而得到一个二叉树的集合 $F = \{T_1, T_2, \dots, T_n\}$ 。

②在 F 中选取根节点的权值最小和次小的两棵二叉树作为左、右子树构造一棵新的二叉树,这棵新的二叉树根节点的权值为其左、右子树根节点权值之和。

③在集合 F 中删除作为左、右子树的两棵二叉树,并将新建立的二叉树加入集合 F 中。

④重复②③两步,当 F 中只剩下一棵二叉树时,这棵二叉树便是所要建立的哈夫曼树。

构造哈夫曼树的原则:

①权值越大的叶节点越靠近根节点。

②权值越小的叶节点越远离根节点。

【PPT课件演示】

图 6.38 所示为哈夫曼树的构造。

★ 课程思政

讲人物以励志,鼓励创新精神。

03 PART THREE 哈夫曼树的构造 $W=\{0.05, 0.29, 0.07, 0.08, 0.14, 0.23, 0.03, 0.11\}$

建立哈夫曼树示例的演示

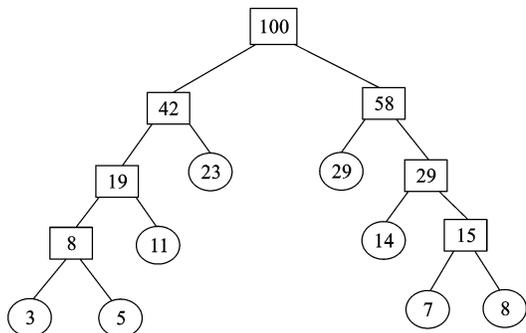


图 6.38 哈夫曼树的构造

❖ 研讨主题一

给一段特定的字符编码。学生分组讨论，每组给出一种编码方案，然后在组间相互质疑和解疑，最终对比各组结果进行总结。

(通过分组讨论与组间讨论培养学生的协同合作的能力，让学生在讨论中能够总结出最佳的编码方案——哈夫曼编码。)

例：假设有一段文本，包含 58 个字符，并由以下 7 个字符构成：a, e, i, s, t, 空格(sp), 换行(nl) (见图 6.39)；这 7 个字符出现的次数不同。如何对这 7 个字符进行编码，使得总编码空间最少？

C_i	a	e	i	s	t	sp	nl
f_i	10	15	12	3	4	13	1

图 6.39 研讨主题一题目

• 知识讲解

分析上一研讨环节的问题，讲解哈夫曼编码。

(1) 用等长 ASCII 编码：58×8=464 位；

(2) 用等长 3 位编码：58×3=174 位；

(3) 不等长编码：出现频率高的字符用的编码短些，出现频率低的字符则可以编码长些。

怎么进行不等长编码呢？看看下面这种编码可以吗？

a: 1e: 0s: 10t: 11……

1011 是什么字符串的编码？

aeaa: 1 0 1 1

aet: 1 0 11

st: 10 11

前缀码 prefix code: 任何字符的编码都不是另一个字符编码的前缀。

提出二叉树编码：左右分支：0、1；字符只在叶节点上。

例：四个字符的频率：a:4, u:1, x:2, z:1 (见图 6.40)

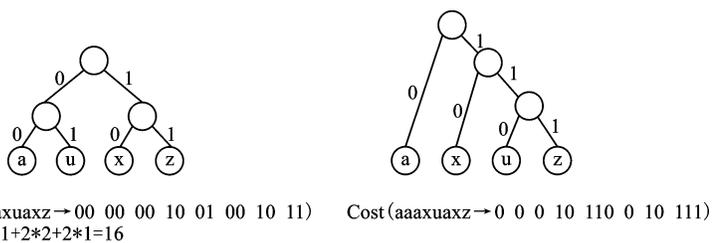


图 6.40 四个字符的两种编码方案图

分析编码代价的计算方式与带权路径长度计算相同，从而推理出编码代价最小的二叉树是哈夫曼树。

❖ 研讨主题二

为研讨主题一中的字符建立哈夫曼编码

通过组内研讨及组间比较，学生能够总结出哈夫曼编码的特点。

【PPT 课件演示】

图 6.41 所示为哈夫曼编码。

C	a	e	i	s	t	sp	nl
f _i	10	15	12	3	4	13	1

a: 111
 e: 10
 i: 00
 s: 11011
 t: 1100
 sp: 01
 nl: 11010
 Cost=3*10+2*15+2*12+5*3+4*4+2*13+5*1=146

哈夫曼编码不唯一
 相同字符对应的编码长度相同
 编码代价相同

图 6.41 哈夫曼编码

◆ 教学小结(2 分钟)

- (1) 哈夫曼树的特点及建立。
- (2) 哈夫曼编码的特点。

◆ 拓展(2 分钟)

- 应用拓展
 - (1) 最佳判定树。
 - (2) 数据压缩。
- 文献拓展
如图 6.42 所示。

★ 课程思政
节约空间，减少数字冗余。通过哈夫曼编码在压缩软件中的应用，倡导同学们在平时编程时应尽可能压缩和节约空间，让数字世界少一些“垃圾”。



基于哈夫曼编码的稀疏
矩阵的存储与计算.pdf

图 6.42 文献拓展

◆ 思考(1 分钟)

当我们收到一串编码后的数据时如何译码呢？（编程实现编码及译码过程）

教 学 反 思

本节课改革了以往直接介绍哈夫曼树概念的讲授法，通过学生熟悉的案例，让学生深入思考程序流程，再引出课堂主题——哈夫曼树，让学生从应用的方面理解哈夫曼树的概念，使得学生能够在理解的基础上记忆概念。课堂中采用了研讨式教学方法，让学生在研讨中学习，在研讨中总结，并在研讨中进步。

从育人方面讲述哈夫曼的故事，介绍哈夫曼树的由来，以名人事迹激励学生勇于探索和创新，追求真理的科学精神。课堂最后对哈夫曼树的应用进行拓展，并提供文献供学生课后阅读，有效激发学生自主学习和探究学习的积极性。