## 数据结构的学习与应用

数据结构是指计算机存储、组织和管理数据的方式，使得数据可以高效地被访问和修改。数据结构的研究内容涵盖了各种不同类型的数据组织方式以及它们之间的相互关系。使用合适的数据结构可以大大优化数据访问和操作的时间复杂度,改善程序运行效率。数据结构为解决问题提供了抽象和建模工具。一个合适的数据结构通常可以很直观地表示问题的本质。

在数据结构学习中我遇到哈希表，哈希表是一种用于实现映射关系的数据结构，它通过哈希函数将键映射到索引，从而实现对数据的快速插入、查找和删除操作。

哈希表的关键思想在于将键映射到数组索引，使得访问元素的时间复杂度接近常数级别。但当遇到字符串哈希时，”多次询问子串哈希，单次计算一个字符串的哈希值复杂度是 O(n)，其中 n 为串长，与暴力匹配没有区别，如果需要多次询问一个字符串的子串的哈希值，每次重新计算效率非常低下。”(OI-wiki 字符串哈希 https://oi-wiki.org/string/hash/)，所以当我们处理大量字符串的哈希时需要考虑如何提高效率，以下是我学习总结后的方法:假定哈希函数将字符串随机地映射到大小为 M 的值域中，总共有 n 个不同的字符串，对目标字符串进行两次大质数取模。因为单一的哈希函数容易出现哈希冲突的情况,不同的字符串可能哈希到同一个槽位中。使用双哈希函数可以大大减少冲突,降低错误匹配的概率。同时单一哈希函数可能会导致某些特殊分布而出现数据聚集在某些槽位的情况。而双哈希函数通常可以获得更加均匀的分布。

因此两个哈希函数的结果可以相互校验,如果不一致则说明数据经过了修改。所以,使用双哈希函数可以获得更低冲突率、更高安全性、更均匀分布、更好并行性和容错性等优势,适合处理大规模字符串数据。

--2203050320 闻家尉