Bài 25: Hãy phân tích các bước thực hiện việc nén dữ liệu theo thuật toán Huffman. Thực hiện nén bằng thuật toán này đối với chuỗi S="BHKJHBBMAABABKMBH". (1,5 điểm)

Giải

**Nguyên lý nén dữ liệu của thuật toán Huffman:**

* Bước 1: Thống kê tần suất xuất hiện của từng ký tự
* Bước 2: Sắp xếp các ký tự theo tần suất giảm/tăng dần
* Bước 3: Dựng cây
* Bước 4: Đánh mã (Trái – 0; phải – 1)
* Bước 5: Ghi mã
* Bước 6: Thay thế các ký tự

**Nén chuỗi: S=" BHKJHBBMAABABKMBH "**

Bước 1: Thống kê tần suất

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| B | H | K | J | M | A |
| 6 | 3 | 2 | 1 | 2 | 3 |

Bước 2: sắp xếp theo tần suất giảm dần

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| B | H | A | K | M | J |  |
| 6 | 3 | 3 | 2 | 2 | 1 | 17 |

Bước 3: Dựng cây

A diagram of a tree

Description automatically generated with medium confidence

Bước 4: Đánh mã: Trái – 0

Phải – 1

Bước 5: Ghi mã:

B: 00, H: 01, A: 11, K: 101, M: 1000, J: 1001

Bước 6: Thay thế các ký tự

S="BHKJHBBMAABABKMBH".

S’ = 00 - 01 - 101 - 1001 - 01 - 00 - 00 - 1000 - 11 - 11 - 00 - 11 - 00 - 101 - 1000 - 00 - 01