

# Inteligência Artificial (L9027) 2021/2022

## Huffman Coding with patterns

Rogério Chaves, 11085218

December 11, 2021

### 1 Introduction

Huffman Coding is a method for coding a message  $\mathbf{m}$  made of characters  $\mathbf{c}$  such that every  $\mathbf{c}$  has a binary string  $\mathbf{S}$  to represent it based on how many time it repeats on  $\mathbf{m}$ . A highest number of repetitions of  $\mathbf{c}$  in  $\mathbf{m}$  translates in  $\mathbf{c}$  having a smaller length of  $\mathbf{S}$  associated with it.

### 2 Assumptions

We will be assuming that a text character is stored in 1 byte (8 bits), giving us a total of 256 possible different values

$$2^8 = 256 \tag{1}$$

let  $\mathbf{c}$  be our character in decimal value:

$$\mathbf{c} \in \mathbb{R} : \{0 \leq C \leq 255\} \tag{2}$$

my rule number 1, unlike ukkonens we have to keep track of the substring range, to do this I have created an extra rule on the ukkonens algorythm, basically, if our curent range is bigger in lenght than the range that is originally on this node we will have to split the node immediatly to account for repetition