

Árvore de Huffman

Prof. Marcos Carrard
carrard@univali.br
carrard@gmail.com

introdução

- Árvore do código de Huffman é utilizada como algoritmo de compactação de arquivos ASCII.
- O algoritmo de compactação de Huffman se baseia na frequência de caracteres e na codificação dos mesmos.



introdução

- A tabela ASCII possui 256 caracteres possíveis e a eles são atribuídos um código de 0 à 255, necessitando de oito bits.
- Dado que um arquivo de texto codificado em ASCII tenha 900 letras, quantos bytes necessitaríamos?



introdução

- A tabela ASCII possui 256 caracteres possíveis e a eles são atribuídos um código de 0 à 255, necessitando de oito bits.
- Dado que um arquivo de texto codificado em ASCII tenha 900 letras, quantos bytes necessitaríamos?
 - a resposta é 900 bytes.



introdução

- O algoritmo de Huffman, utiliza a frequência dos caracteres e estabelece uma codificação diferenciada para cada um deles.
 - Símbolos que aparecem numa **maior frequência**, ganham uma codificação em bits menor;
 - Símbolos que aparecem em **menor frequência** recebem codificação maior.



introdução

- Exemplo de árvore de Huffman



Algoritmo huffman

passarinho que anda com morcego acaba dormindo de ponta cabeça

Quantos bytes e bits possui este texto?



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

Resposta:

62 B

496 bits



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

Dado o texto ao lado, vamos contar a frequência de cada uma das letras presentes no texto.



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

Temos 2 símbolos
"p"

2
p



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

Temos 10 símbolos
"a"

(note que a
listagem abaixo
está sendo
montada por
ordem de
frequência)

2	10
p	a



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

2	2	10
p	s	a

Temos 2 símbolos
"s"

(note que a listagem abaixo está sendo montada por ordem de frequência)



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

2	2	3	10
p	s	r	a

Temos 3 símbolos
"r"

(note que a listagem abaixo está sendo montada por ordem de frequência)



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

2	2	2	3	10
p	s	i	r	a

Temos 2 símbolos
"i"

(note que a listagem abaixo está sendo montada por **ordem de frequência**)



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

2	2	2	3	4	10
p	s	i	r	n	a

Temos 4 símbolos
"n"

(note que a listagem abaixo está sendo montada por **ordem de frequência**)



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	2	2	2	3	4	10
h	p	s	i	r	n	a

Temos 1 símbolos
"h"

(note que a
listagem abaixo
está sendo
montada por
**ordem de
frequência**)



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	2	2	2	3	4	7	10
h	p	s	i	r	n	o	a

Temos 7 símbolos
"o"

(note que a
listagem abaixo
está sendo
montada por
**ordem de
frequência**)



Algoritmo huffman

passarinho_que_anda_com_morcego
_acaba_dormindo_de_ponta_cabeca

Temos 9 símbolos
" "

(note que a
listagem abaixo
está sendo
montada por
**ordem de
frequência**)

1	2	2	2	3	4	7	9	10
h	p	s	i	r	n	o		a



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeca

Temos 1 símbolo
"q"

(note que a
listagem abaixo
está sendo
montada por
**ordem de
frequência**)

1	1	2	2	2	3	4	7	9	10
h	q	p	s	i	r	n	o		a



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	1	1	2	2	2	3	4	7	9	10
h	q	u	p	s	i	r	n	o		a

Temos 1 símbolos
"u"

(note que a
listagem abaixo
está sendo
montada por
**ordem de
frequência**)



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	1	1	2	2	2	3	4	4	7	9	10
h	q	u	p	s	i	r	n	e	o		a

Temos 4 símbolos
"e"

(note que a
listagem abaixo
está sendo
montada por
**ordem de
frequência**)



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

Temos 4 símbolos
"d"

(note que a
listagem abaixo
está sendo
montada por
**ordem de
frequência**)

1	1	1	2	2	2	3	4	4	4	7	9	10
h	q	u	p	s	i	r	n	e	d	o		a



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

Temos 5 símbolos
"c"

(note que a
listagem abaixo
está sendo
montada por
**ordem de
frequência**)

1	1	1	2	2	2	3	4	4	4	5	7	9	10
h	q	u	p	s	i	r	n	e	d	c	o		a



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

Temos 3 símbolos
"m"

(note que a
listagem abaixo
está sendo
montada por
**ordem de
frequência**)

1	1	1	2	2	2	3	3	4	4	4	5	7	9	10
h	q	u	p	s	i	r	m	n	e	d	c	o		a



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

Temos 1 símbolo
"g"

(note que a
listagem abaixo
está sendo
montada por
**ordem de
frequência**)

1	1	1	1	2	2	2	3	3	4	4	4	5	7	9	10
h	q	u	g	p	s	i	r	m	n	e	d	c	o		a



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

Temos 2 símbolos
"b"

(note que a
listagem abaixo
está sendo
montada por
ordem de
frequência)

1	1	1	1	2	2	2	2	3	3	4	4	4	5	7	9	10
h	q	u	g	p	s	i	b	r	m	n	e	d	c	o		a



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de pont**a** cabeça

Temos 1 símbolo
"t"

(note que a
listagem abaixo
está sendo
montada por
ordem de
frequência)

1	1	1	1	1	2	2	2	2	3	3	4	4	4	5	7	9	10
h	q	u	t	g	p	s	i	b	r	m	n	e	d	c	o		a



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

Agora temos nossa
listagem de
frequências

1	1	1	1	1	2	2	2	2	3	3	4	4	4	5	7	9	10
h	q	u	t	g	p	s	i	b	r	m	n	e	d	c	o		a



Algoritmo huffman

passa
acab

Com a listagem de frequências, iremos
realizar o algoritmo de Huffman.

1	1	1	1	1	2	2	2	2	3	3	4	4	4	5	7	9	10
h	q	u	t	g	p	s	i	b	r	m	n	e	d	c	o		a



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

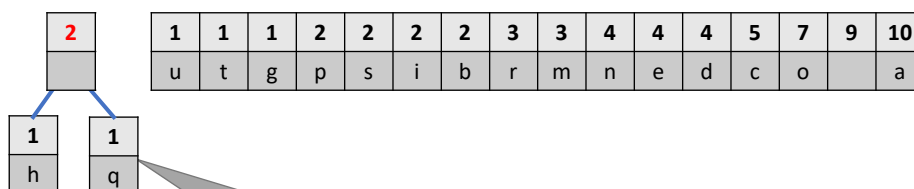
1	1	1	1	2	2	2	2	3	3	4	4	4	5	7	9	10
h	q	u	t	g	p	s	i	b	r	m	n	e	d	c	o	a

Com a listagem pronta, utilizamos sempre os **dois elementos de menores frequências**, retiramo-os da lista



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



Cria-se uma árvore, apontando para cada um dos elementos, sendo que o elemento de menor frequência fica a esquerda e o de maior, a direita. A raiz deve ser a soma das frequências



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	1	1	2	2	2	2	2	3	3	4	4	4	5	7	9	10
u	t	g	p	s	i	b		r	m	n	e	d	c	o		a

1	1
h	q

Depois, deve-se inserir a raiz
à listagem.

UNIVALI

Escola
Politécnica

Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	1	1	2	2	2	2	2	3	3	4	4	4	5	7	9	10
u	t	g	p	s	i	b		r	m	n	e	d	c	o		a

Deve-se repetir o
procedimento, até que
haja apenas uma raiz
na lista!

UNIVALI

Escola
Politécnica

Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	1
u	t

1	2	2	2	2	2	3	3	4	4	4	5	7	9	10
g	p	s	i	b		r	m	n	e	d	c	o		a

1	1
h	q



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

2	
1	1
u	t

1	2	2	2	2	2	3	3	4	4	4	5	7	9	10
g	p	s	i	b		r	m	n	e	d	c	o		a

1	1
h	q



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	2	2	2	2	2	2	3	3	4	4	4	5	7	9	10
g	p	s	i	b			r	m	n	e	d	c	o		a

1	1	1	1
h	q	u	t



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	2
g	p

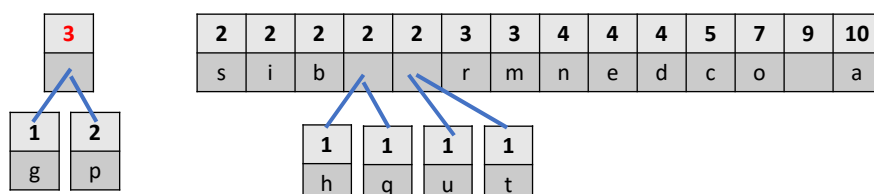
2	2	2	2	2	2	3	3	4	4	4	5	7	9	10
s	i	b				r	m	n	e	d	c	o		a

1	1	1	1
h	q	u	t



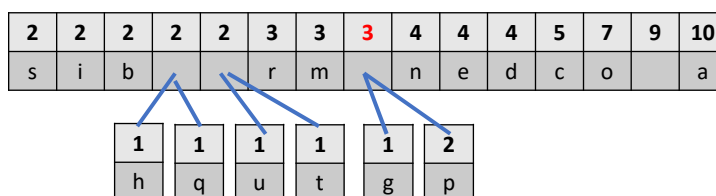
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

2	2
s	i

2	2	2	3	3	3	4	4	4	5	7	9	10
b			r	m		n	e	d	c	o		a

1	1	1	1	1	2
h	q	u	t	g	p



Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça

4	
2	2
s	i

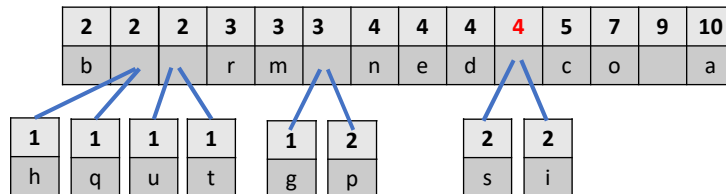
2	2	2	3	3	3	4	4	4	5	7	9	10
b			r	m		n	e	d	c	o		a

1	1	1	1	1	2
h	q	u	t	g	p



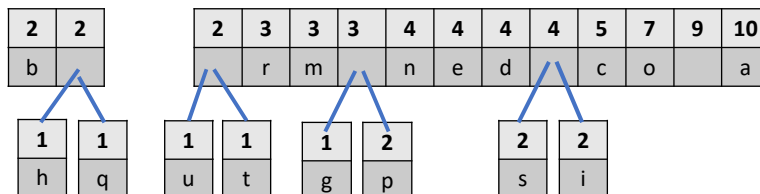
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



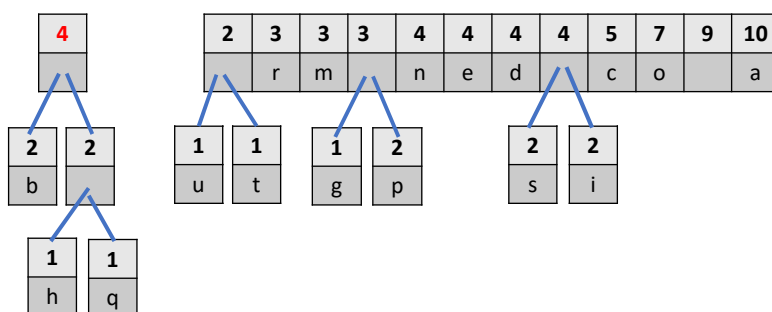
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



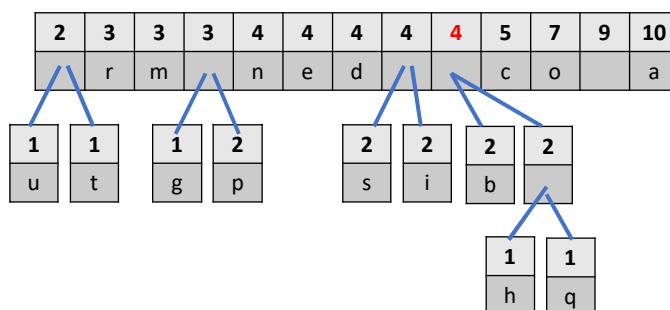
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



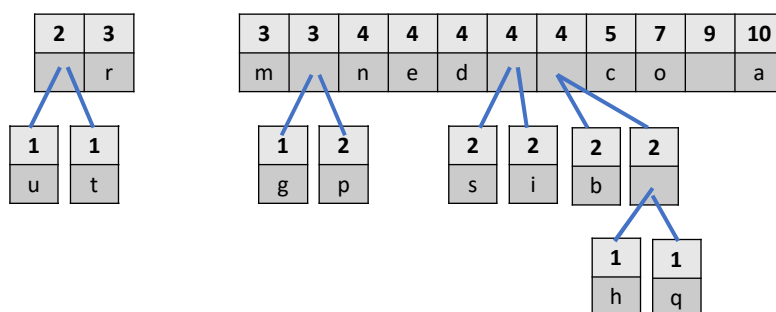
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



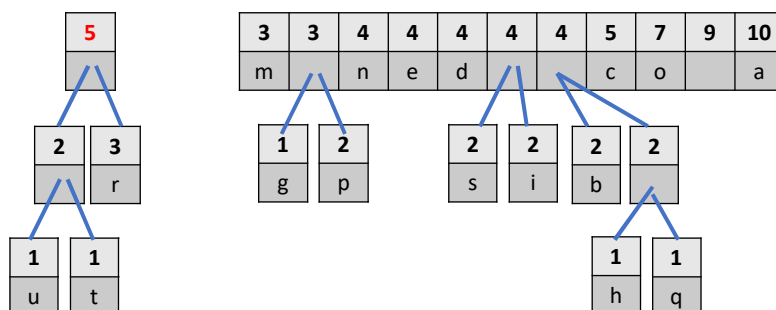
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



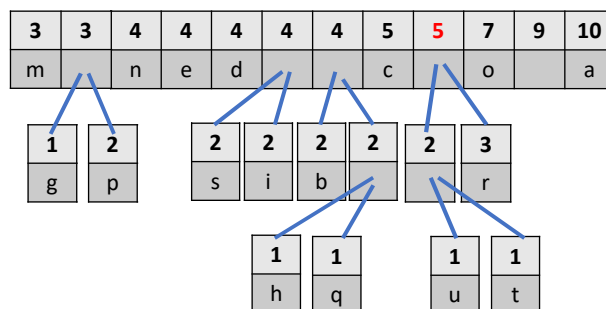
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



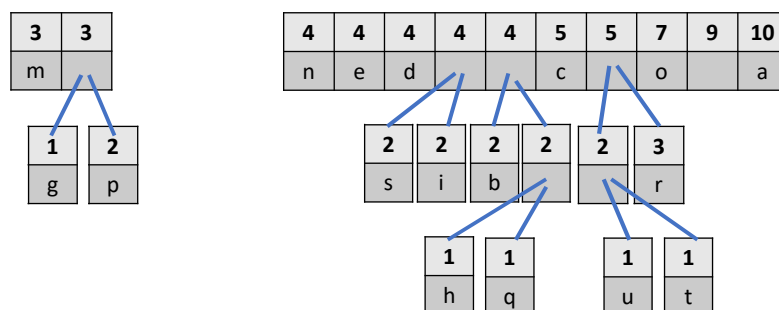
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



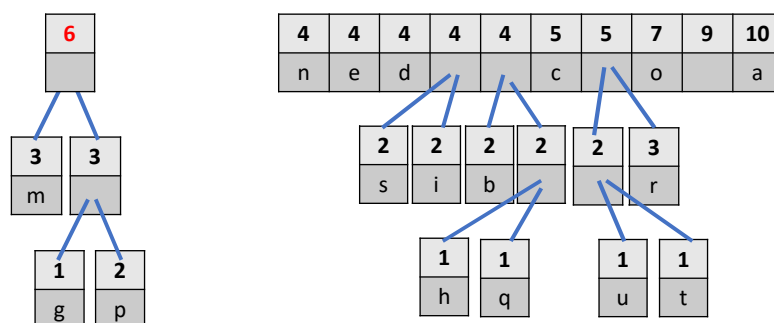
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



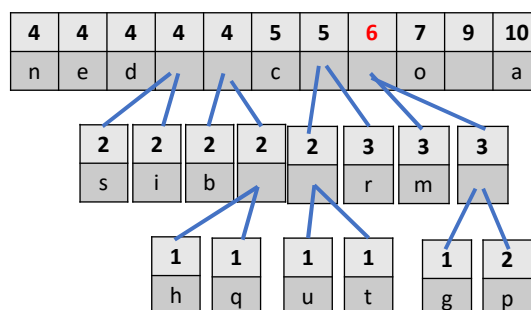
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



Algoritmo huffman

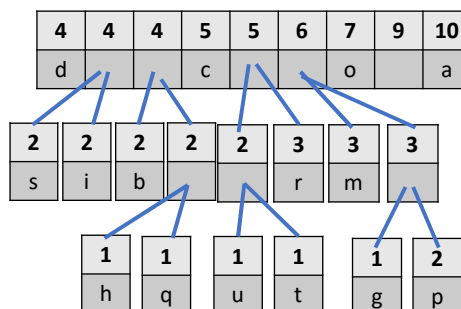
passarinho que anda com morcego
acaba dormindo de ponta cabeça



Algoritmo huffman

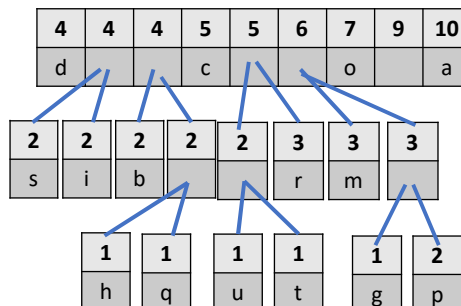
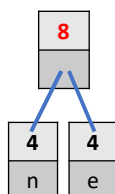
passarinho que anda com morcego
acaba dormindo de ponta cabeça

4	4
n	e



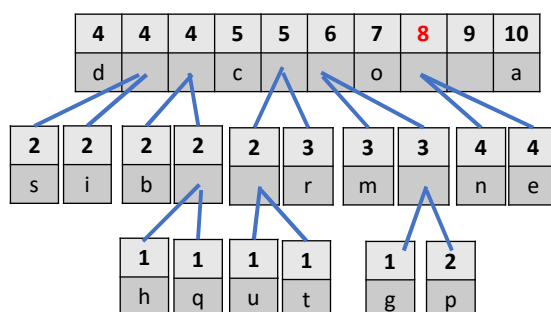
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



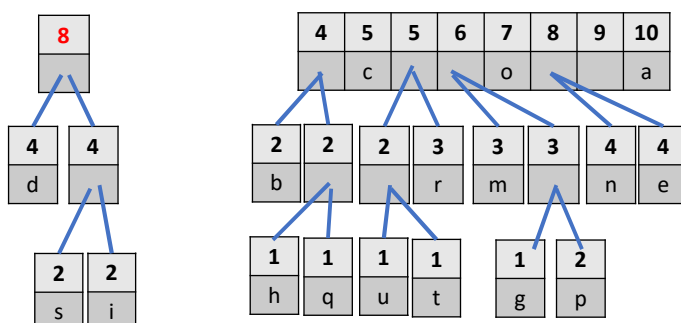
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



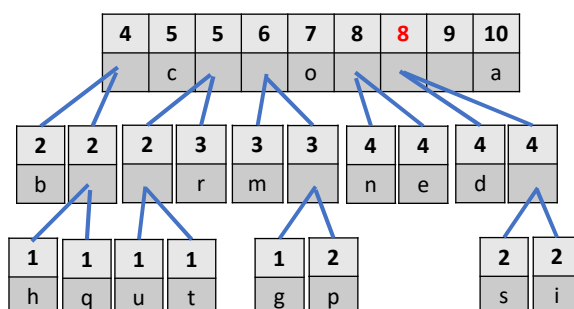
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



Algoritmo huffman

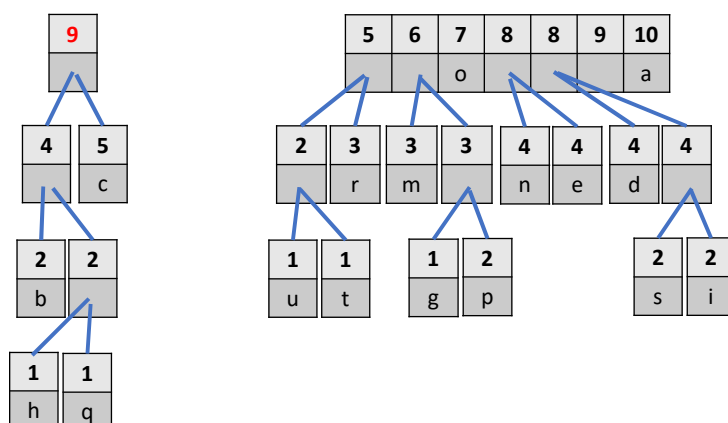
passarinho que anda com morcego
acaba dormindo de ponta cabeça



Escola
Politécnica

Algoritmo huffman

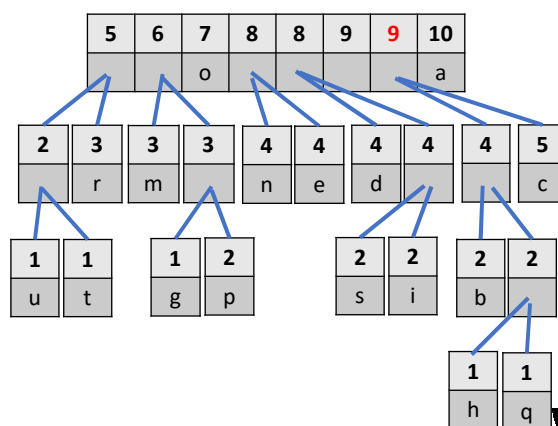
passarinho que anda com morcego
acaba dormindo de ponta cabeça



Escola
Politécnica

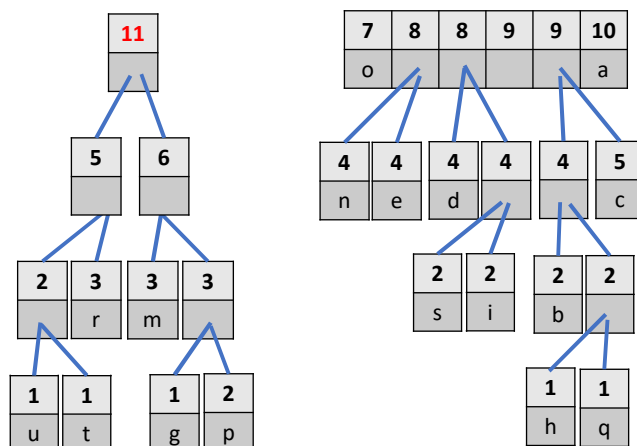
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



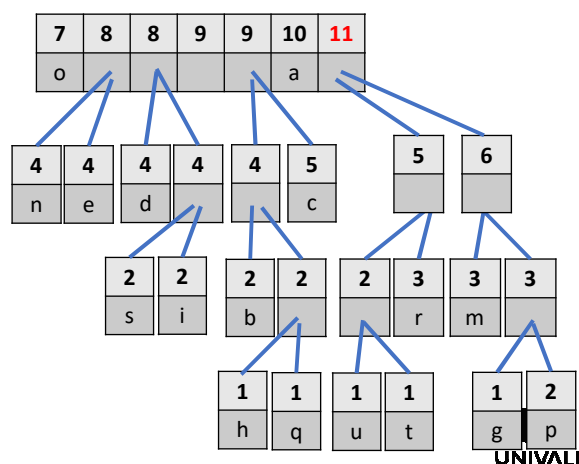
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



Algoritmo huffman

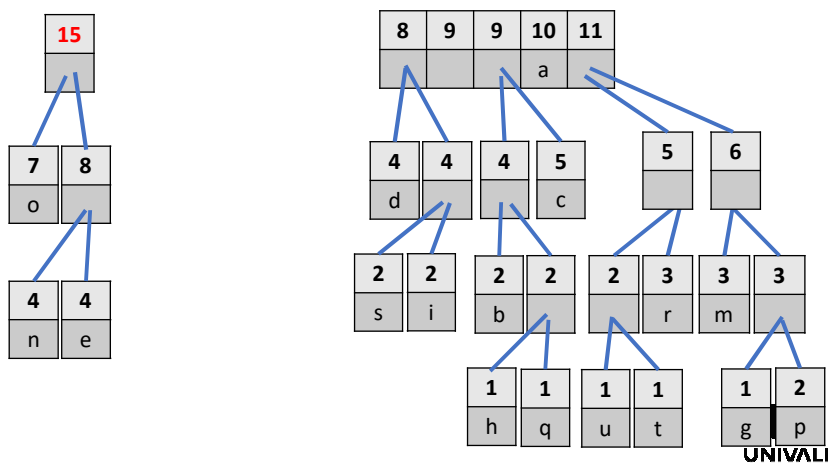
passarinho que anda com morcego
acaba dormindo de ponta cabeça



Escola
Politécnica

Algoritmo huffman

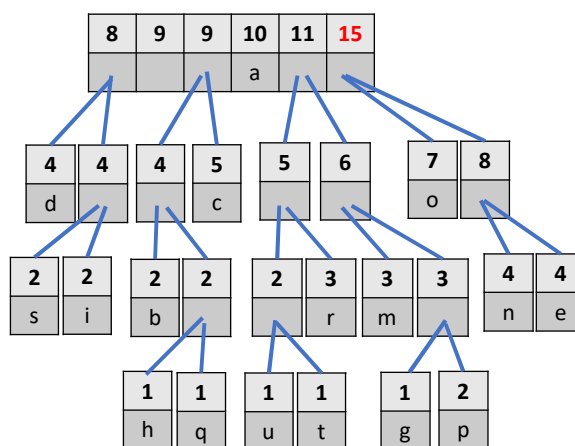
passarinho que anda com morcego
acaba dormindo de ponta cabeça



Escola
Politécnica

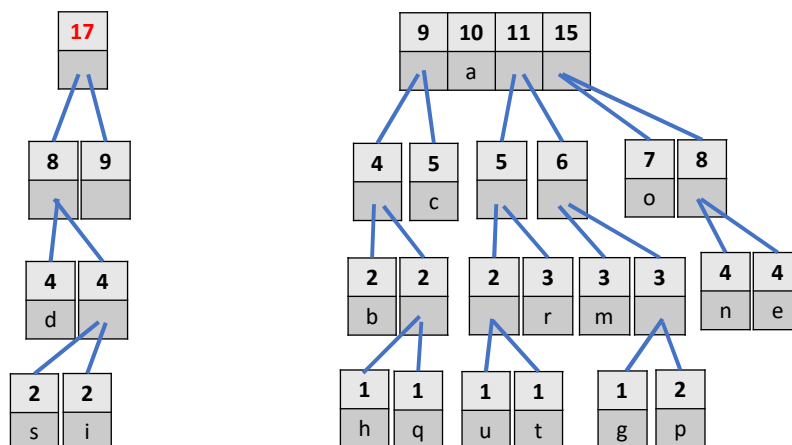
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



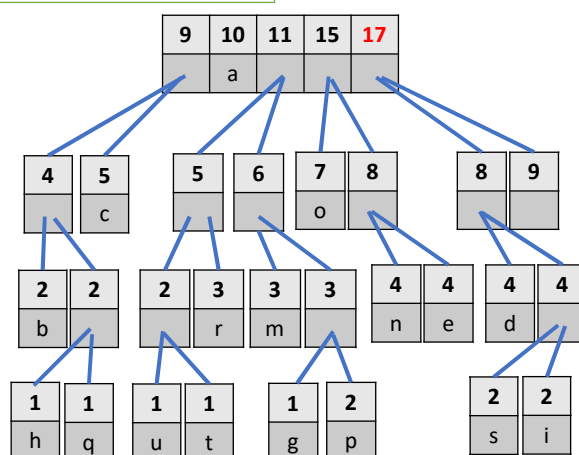
Algoritmo huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



Algoritmo huffman

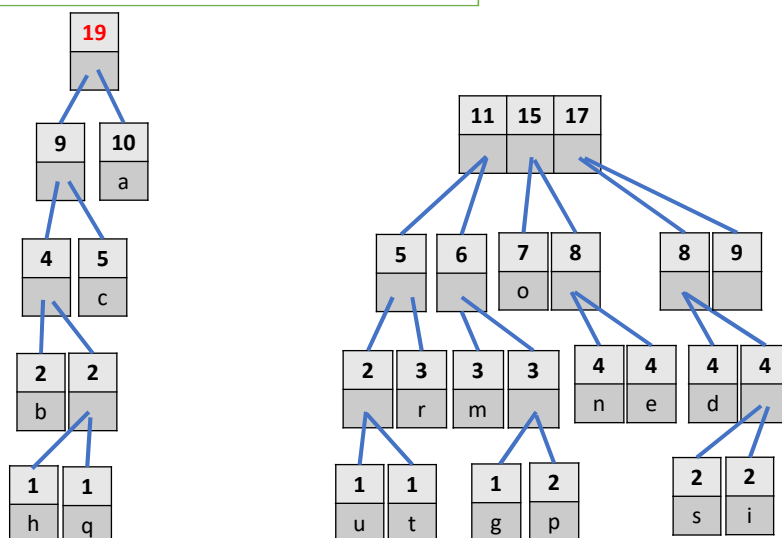
passarinho que anda com morcego
acaba dormindo de ponta cabeça



Escola
Politécnica

Algoritmo huffman

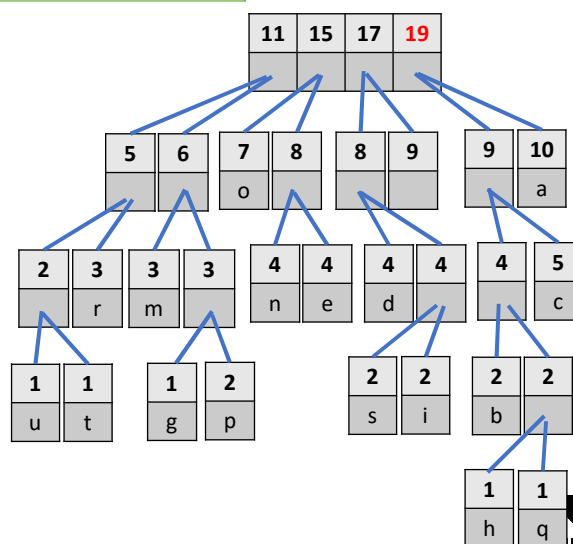
passarinho que anda com morcego
acaba dormindo de ponta cabeça



Escola
Politécnica

Algoritmo de Huffman

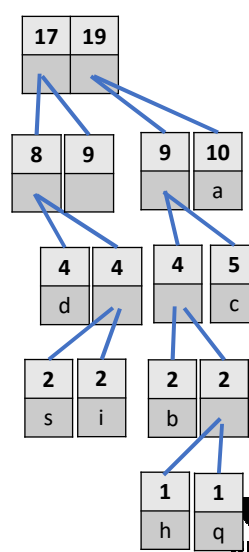
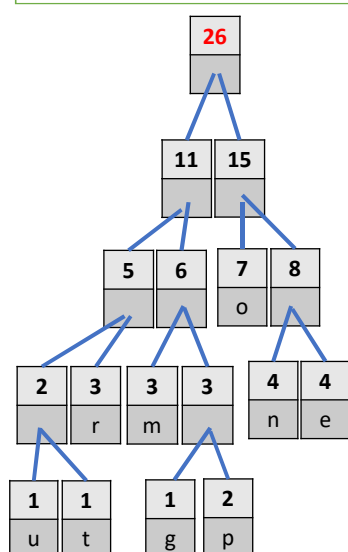
passarinho que anda com morcego
acaba dormindo de ponta cabeça



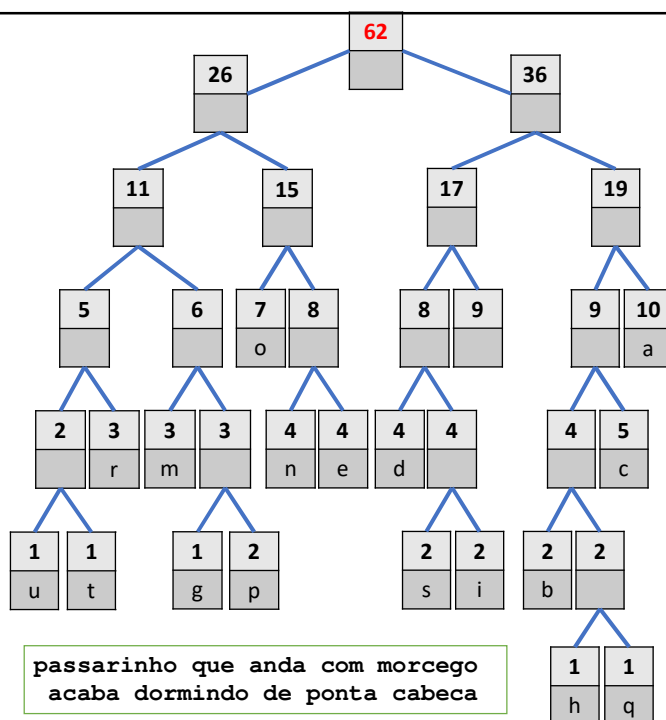
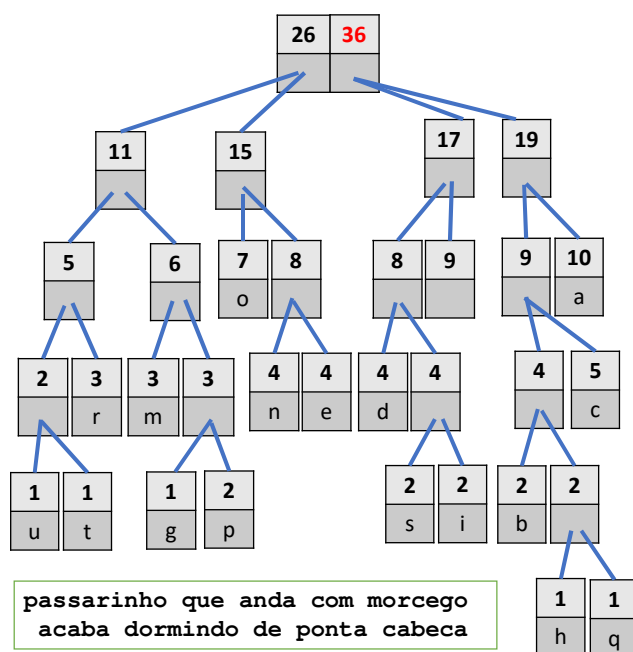
Escola
Politécnica

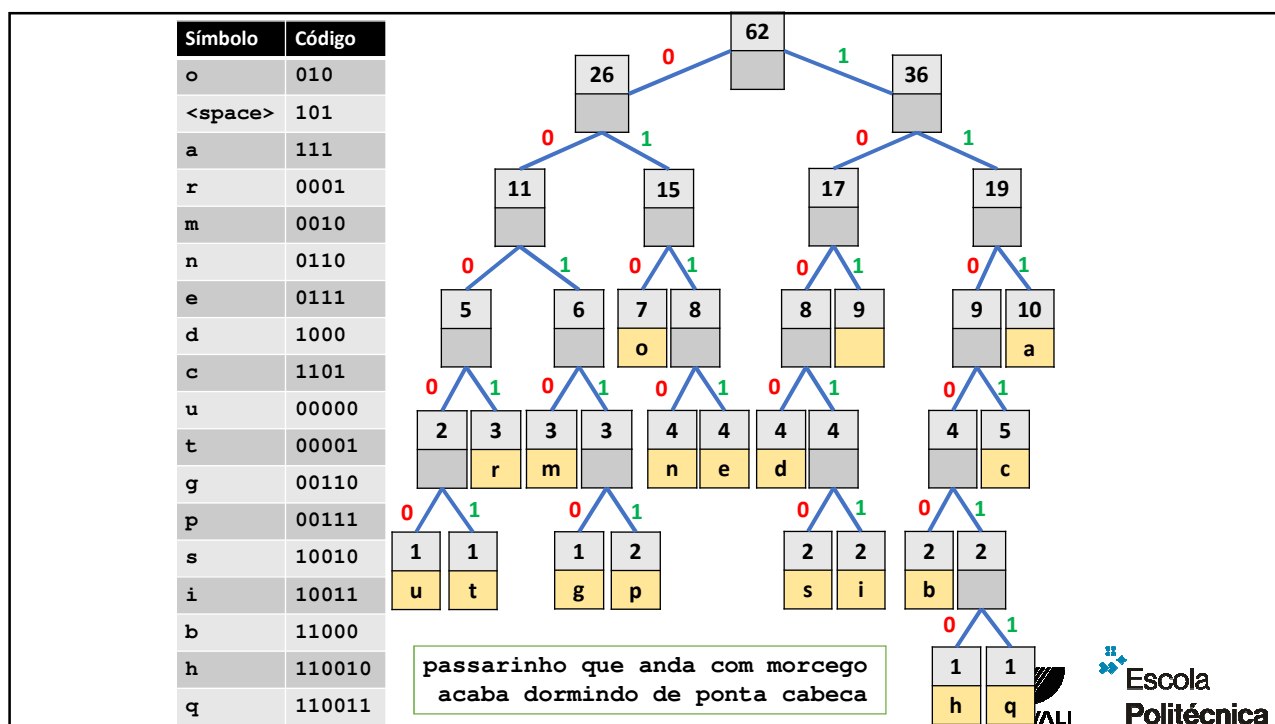
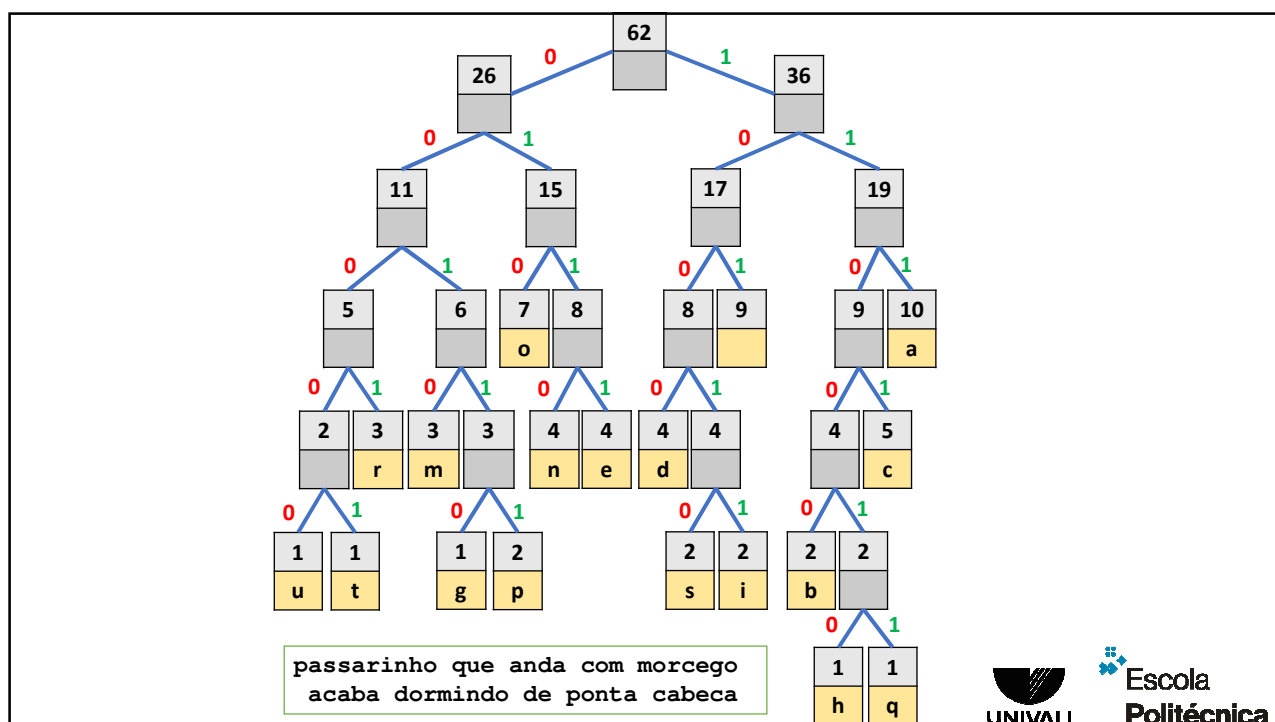
Algoritmo de Huffman

passarinho que anda com morcego
acaba dormindo de ponta cabeça



Escola
Politécnica





Símbolo	Código
o	010
<space>	101
a	111
r	0001
m	0010
n	0110
e	0111
d	1000
c	1101
u	00000
t	00001
g	00110
p	00111
s	10010
i	10011
b	11000
h	110010
q	110011

passarinho que anda com morcego
acaba dormindo de ponta cabeça

```

0011111110010100101110001100110
110110010010101110011000001111
0111101101000111101110101000101
0100100100001110101110011001010
111110111111000111101100001000
0100101001101101000010101100001
111010011101001100000111101110
111110000111101111

```

237 bits - Huffman
versus
496 bits - ASCII

UNIVALI Escola Politécnica

Símbolo	Código
o	010
<space>	101
a	111
r	0001
m	0010
n	0110
e	0111
d	1000
c	1101
u	00000
t	00001
g	00110
p	00111
s	10010
i	10011
b	11000
h	110010
q	110011

passarinho que anda com morcego
acaba dormindo de ponta cabeça

```

0011111110010100101110001100110
110110010010101110011000001111
0111101101000111101110101000101
0100100100001110101110011001010
111110111111000111101100001000
0100101001101101000010101100001
111010011101001100000111101110
111110000111101111

```

237 bits - Huffman
versus
496 bits - ASCII

$\frac{237}{496} = 0,47\dots$

A codificação de Huffman é equivalente a cerca de 47% do tamanho em ASCII para a frase em questão.
Ou seja, utiliza menos das metade do espaço para armazenar/transmitir o mesmo conteúdo.

UNIVALI Escola Politécnica

Árvore de Huffman → TAD!



TAD

- Vamos pensar no que precisaremos para criar uma árvore de Huffman!
- Vamos tentar classificar



TAD

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	1	1	1	1	2	2	2	2	3	3	3	4	4	5	7	9	10
h	q	u	t	g	p	s	i	b	r	d	m	n	e	c	o		a

É fato que precisamos de uma lista, que temos que manter ordenada de acordo com a quantidade de ocorrências!



TAD

passarinho que anda com morcego
acaba dormindo de ponta cabeça

1	1
h	q

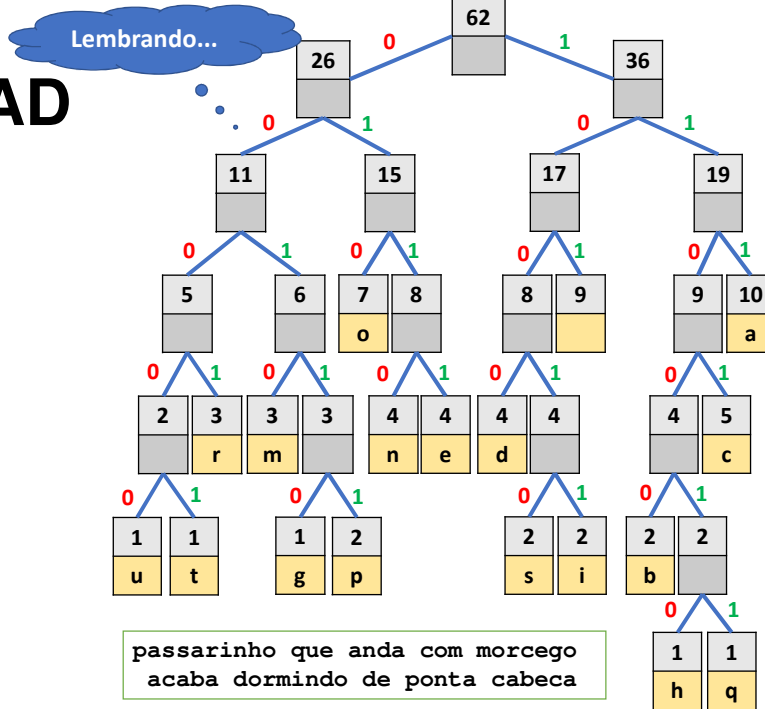
Como faremos os **elementos** nesta lista...

9	10
	a

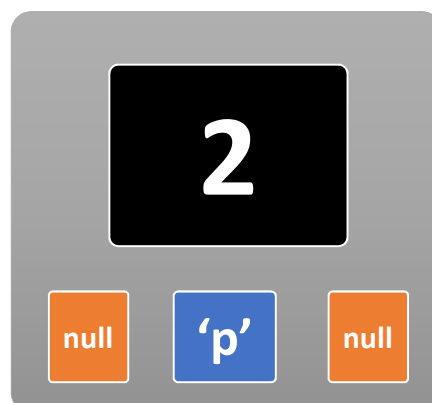
Importante considerarmos que a mesma lista deverá ser usada na montagem da árvore.



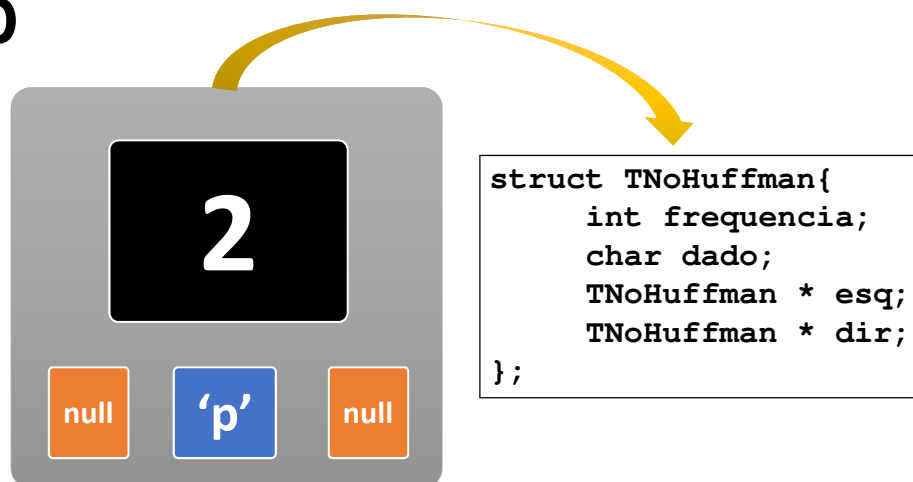
TAD



TAD



TAD



Escola
Politécnica



K-HARD



Escola
Politécnica