

Codifica di Huffman

Claudio Mirolo

Dipartimento di Scienze Matematiche, Informatiche e Fisiche
Università di Udine, via delle Scienze 206 – Udine

claudio.mirolo@uniud.it

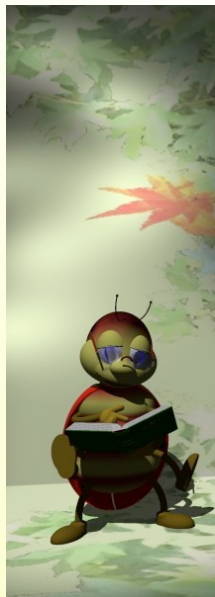


Corso di Programmazione

La definizione di informatica che prediligo è che si tratta dello studio degli algoritmi [...].

Un programma è una particolare rappresentazione di un algoritmo, esattamente nello stesso senso in cui il termine “dato” si riferisce a una particolare rappresentazione di “informazione”.

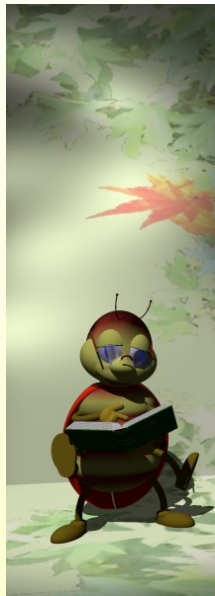
Forse la scoperta più significativa determinata dall'avvento del computer sarà che gli algoritmi, come oggetti di studio, sono straordinariamente ricchi di proprietà interessanti; inoltre, un punto di vista algoritmico costituisce un modo utile di organizzare la conoscenza in generale.



La definizione di informatica che prediligo è che si tratta dello studio degli algoritmi [...].

Un programma è una particolare rappresentazione di un algoritmo, esattamente nello stesso senso in cui il termine “dato” si riferisce a una particolare rappresentazione di “informazione”.

Forse la scoperta più significativa determinata dall'avvento del computer sarà che gli algoritmi, come oggetti di studio, sono straordinariamente ricchi di proprietà interessanti; inoltre, un punto di vista algoritmico costituisce un modo utile di organizzare la conoscenza in generale.

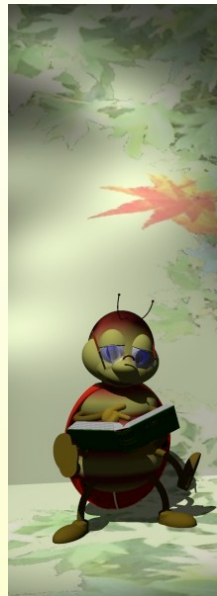


La definizione di informatica che prediligo è che si tratta dello studio degli algoritmi [...].

Un programma è una particolare rappresentazione di un algoritmo, esattamente nello stesso senso in cui il termine “dato” si riferisce a una particolare rappresentazione di “informazione”.

Forse la scoperta più significativa determinata dall'avvento del computer sarà che gli algoritmi, come oggetti di studio, sono straordinariamente ricchi di proprietà interessanti;

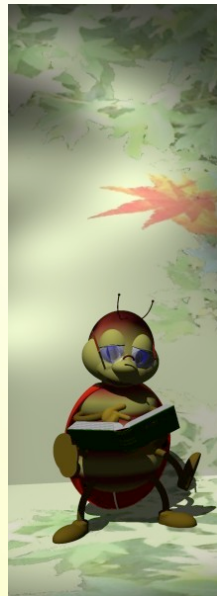
inoltre, un punto di vista algoritmico costituisce un modo utile di organizzare la conoscenza in generale.



La definizione di informatica che prediligo è che si tratta dello studio degli algoritmi [...].

Un programma è una particolare rappresentazione di un algoritmo, esattamente nello stesso senso in cui il termine “dato” si riferisce a una particolare rappresentazione di “informazione”.

Forse la scoperta più significativa determinata dall'avvento del computer sarà che gli algoritmi, come oggetti di studio, sono straordinariamente ricchi di proprietà interessanti; inoltre, un punto di vista algoritmico costituisce un modo utile di organizzare la conoscenza in generale.



Codifica di Huffman: Inventarsi nuovo spazio...



Codifica di Huffman: Esempio

- Sequenza di 12 basi nucleiche (pretesto):

A T T C T A C C T T G T

- Codifica binaria standard (stile ASCII)
— due bit per distinguere fra quattro simboli:

A	00
T	01
C	10
G	11

- Rappresentazione della sequenza (24 bit):

0 0 0 1 0 1 1 0 0 1 0 0 1 0 1 0 0 1 0 1 1 1 0 1

Codifica di Huffman: Esempio

- Sequenza di 12 basi nucleiche (pretesto):

A T T C T A C C T T G T

- Codifica binaria standard (stile ASCII)
— due bit per distinguere fra quattro simboli:

A	00
T	01
C	10
G	11

- Rappresentazione della sequenza (24 bit):

0 0 0 1 0 1 1 0 0 1 0 0 1 0 1 0 0 1 0 1 1 1 0 1

Codifica di Huffman: Esempio

- Sequenza di 12 basi nucleiche (pretesto):

A T T C T A C C T T G T

- Codifica binaria standard (stile ASCII)
— due bit per distinguere fra quattro simboli:

A	00
T	01
C	10
G	11

- Rappresentazione della sequenza (24 bit):

0 0 0 1 0 1 1 0 0 1 0 0 1 0 1 0 0 1 0 1 1 1 0 1

Codifica di Huffman: Esempio

- Sequenza di 12 basi nucleiche (pretesto):

A T T C T A C C T T G T

- Codifica binaria standard (stile ASCII)
— due bit per distinguere fra quattro simboli:

A	00
T	01
C	10
G	11

- Rappresentazione della sequenza (24 bit):

0 0 0 1 0 1 1 0 0 1 0 0 1 0 1 0 0 1 0 1 1 1 0 1

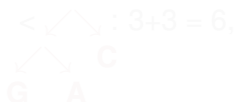
Codifica di Huffman: Esempio

- *Peso* dei simboli = numero di occorrenze (o frequenza)
coppie simbolo/peso ordinate per peso crescente:

< **G** : 1, **A** : 2, **C** : 3, **T** : 6 >

- Costruzione dell'albero di Huffman — aggregazione dei 2 elementi di minor peso e riordinamento:

<  : 1+2 = 3, **C** : 3, **T** : 6 >

<  : 3+3 = 6, **T** : 6 >

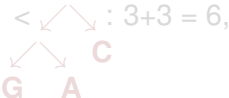
Codifica di Huffman: Esempio

- *Peso* dei simboli = numero di occorrenze (o frequenza)
coppie simbolo/peso ordinate per peso crescente:

< **G** : 1, **A** : 2, **C** : 3, **T** : 6 >

- Costruzione dell'albero di Huffman — aggregazione dei 2 elementi di minor peso e riordinamento:

<  : 1+2 = 3, **C** : 3, **T** : 6 >

<  : 3+3 = 6, **T** : 6 >

Codifica di Huffman: Esempio

- *Peso* dei simboli = numero di occorrenze (o frequenza)
coppie simbolo/peso ordinate per peso crescente:

< **G** : 1, **A** : 2, **C** : 3, **T** : 6 >

- Costruzione dell'albero di Huffman — aggregazione dei 2 elementi di minor peso e riordinamento:

<  : 1+2 = 3, **C** : 3, **T** : 6 >

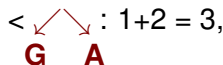
<  : 3+3 = 6, **T** : 6 >

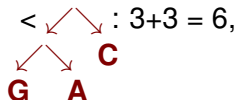
Codifica di Huffman: Esempio

- *Peso* dei simboli = numero di occorrenze (o frequenza)
coppie simbolo/peso ordinate per peso crescente:

< **G** : 1, **A** : 2, **C** : 3, **T** : 6 >

- Costruzione dell'albero di Huffman — aggregazione dei 2 elementi di minor peso e riordinamento:

<  : 1+2 = 3, **C** : 3, **T** : 6 >

<  : 3+3 = 6, **T** : 6 >

Codifica di Huffman: Esempio

- Albero di Huffman:

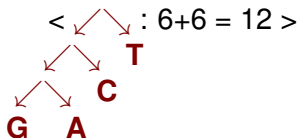


- Codice definito dal percorso attraverso l'albero per raggiungere il simbolo (0 = sinistra, 1 = destra):

A	001
T	1
C	01
G	000

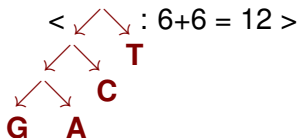
Codifica di Huffman: Esempio

- Albero di Huffman:



Codifica di Huffman: Esempio

- Albero di Huffman:

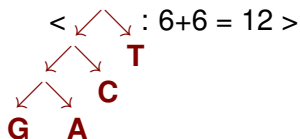


- Codice definito dal percorso attraverso l'albero per raggiungere il simbolo (0 = sinistra, 1 = destra):

A	001
T	1
C	01
G	000

Codifica di Huffman: Esempio

- Albero di Huffman:

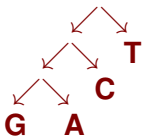


- Codice definito dal percorso attraverso l'albero per raggiungere il simbolo (0 = sinistra, 1 = destra):

A	001
T	1
C	01
G	000

Codifica di Huffman: Esempio

- Albero di Huffman:



A	001
T	1
C	01
G	000

- Rappresentazione “compressa” della sequenza (21 bit):

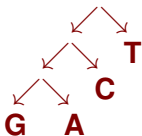
A T T C T A C C T T G T

0 0 1 1 1 0 1 1 0 0 1 0 1 0 1 1 1 0 0 0 1

- Decodifica: interpretazione della sequenza binaria come istruzioni per scendere attraverso l'albero...

Codifica di Huffman: Esempio

- Albero di Huffman:



A	001
T	1
C	01
G	000

- Rappresentazione “compressa” della sequenza (21 bit):

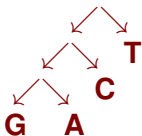
A T T C T A C C T T G T

0 0 1 1 1 0 1 1 0 0 1 0 1 0 1 1 1 0 0 0 1

- Decodifica: interpretazione della sequenza binaria come istruzioni per scendere attraverso l'albero...

Codifica di Huffman: Esempio

- Albero di Huffman:



A	001
T	1
C	01
G	000

- Rappresentazione “compressa” della sequenza (21 bit):

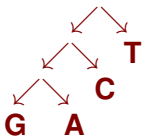
A T T C T A C C T T G T

001110110010101110001

- Decodifica: interpretazione della sequenza binaria come istruzioni per scendere attraverso l'albero...

Codifica di Huffman: Esempio

- Albero di Huffman:



A	001
T	1
C	01
G	000

- Rappresentazione “compressa” della sequenza (21 bit):

A T T C T A C C T T G T

001110110010101110001

- Decodifica: interpretazione della sequenza binaria come istruzioni per scendere attraverso l'albero...

Codifica di Huffman: Ancora un esempio?



Codifica di Huffman: Ancora un esempio

- Testo con 4 vocali e 4 consonanti:

ILTEATROALLIETALASERATA

- Codifica binaria standard (stile ASCII)
— tre bit per distinguere fra otto lettere:

A	000
E	001
I	010
O	011

L	100
R	101
S	110
T	111

- Rappresentazione della sequenza (69 bit):

010100111001000111101011000100100010
001111000100000110001101000111000

Codifica di Huffman: Ancora un esempio

- Testo con 4 vocali e 4 consonanti:

ILTEATROALLIETALASERATA

- Codifica binaria standard (stile ASCII)
— tre bit per distinguere fra otto lettere:

A	000
E	001
I	010
O	011

L	100
R	101
S	110
T	111

- Rappresentazione della sequenza (69 bit):

010100111001000111101011000100100010
001111000100000110001101000111000

Codifica di Huffman: Ancora un esempio

- Testo con 4 vocali e 4 consonanti:

I L T E A T R O A L L I E T A L A S E R A T A

- Codifica binaria standard (stile ASCII)
— tre bit per distinguere fra otto lettere:

A	000
E	001
I	010
O	011

L	100
R	101
S	110
T	111

- Rappresentazione della sequenza (69 bit):

010100111001000111101011000100100010
001111000100000110001101000111000

Codifica di Huffman: Ancora un esempio

- Testo con 4 vocali e 4 consonanti:

I L T E A T R O A L L I E T A L A S E R A T A

- Codifica binaria standard (stile ASCII)
— tre bit per distinguere fra otto lettere:

A	000
E	001
I	010
O	011

L	100
R	101
S	110
T	111

- Rappresentazione della sequenza (69 bit):

010100111001000111101011000100100010
001111000100000110001101000111000

Codifica di Huffman: Ancora un esempio

- Lettere e pesi = numeri di occorrenze:

< **O** : ?, **S** : ?, **I** : ?, **R** : ?, **E** : ?, **L** : ?, **T** : ?, **A** : ? >

Codifica di Huffman: Ancora un esempio

- Lettere e pesi = numeri di occorrenze:

< **O** : ?, **S** : ?, **I** : ?, **R** : ?, **E** : ?, **L** : ?, **T** : ?, **A** : ? >

ILTEATROALLIETALASERATA

Codifica di Huffman: Ancora un esempio

- Lettere e pesi = numeri di occorrenze:

< **O** : 1, **S** : 1, **I** : 2, **R** : 2, **E** : 3, **L** : 4, **T** : 4, **A** : 6 >

ILTEATROALLIETALASERATA

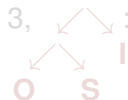
Codifica di Huffman: Ancora un esempio

- Lettere e pesi = numeri di occorrenze:

< **O** : 1, **S** : 1, **I** : 2, **R** : 2, **E** : 3, **L** : 4, **T** : 4, **A** : 6 >

- Costruzione dell'albero di Huffman:

<  : 2, **I** : 2, **R** : 2, **E** : 3, **L** : 4, **T** : 4, **A** : 6 >

< **R** : 2, **E** : 3,  : 4, **L** : 4, **T** : 4, **A** : 6 >

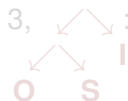
Codifica di Huffman: Ancora un esempio

- Lettere e pesi = numeri di occorrenze:

< **O** : 1, **S** : 1, **I** : 2, **R** : 2, **E** : 3, **L** : 4, **T** : 4, **A** : 6 >

- Costruzione dell'albero di Huffman:

<  : 2, **I** : 2, **R** : 2, **E** : 3, **L** : 4, **T** : 4, **A** : 6 >

< **R** : 2, **E** : 3,  : 4, **L** : 4, **T** : 4, **A** : 6 >

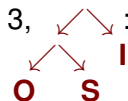
Codifica di Huffman: Ancora un esempio

- Lettere e pesi = numeri di occorrenze:

< **O** : 1, **S** : 1, **I** : 2, **R** : 2, **E** : 3, **L** : 4, **T** : 4, **A** : 6 >

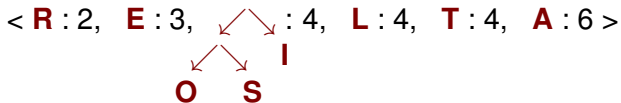
- Costruzione dell'albero di Huffman:

<  : 2, **I** : 2, **R** : 2, **E** : 3, **L** : 4, **T** : 4, **A** : 6 >

< **R** : 2, **E** : 3,  : 4, **L** : 4, **T** : 4, **A** : 6 >

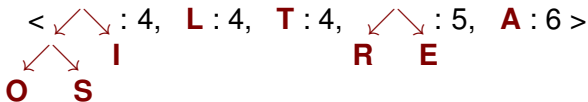
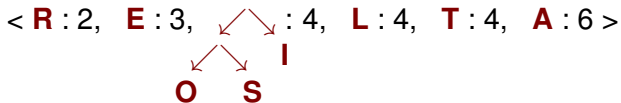
Codifica di Huffman: Ancora un esempio

- Costruzione dell'albero di Huffman:



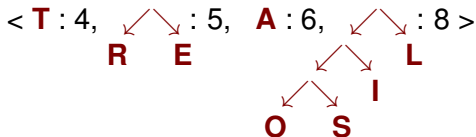
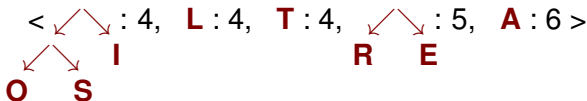
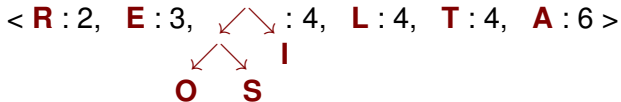
Codifica di Huffman: Ancora un esempio

- Costruzione dell'albero di Huffman:



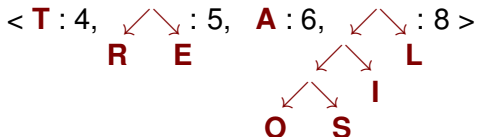
Codifica di Huffman: Ancora un esempio

- Costruzione dell'albero di Huffman:



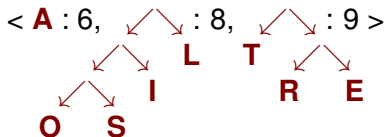
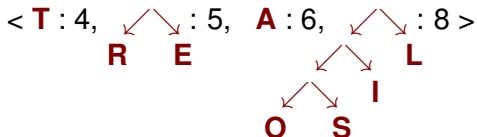
Codifica di Huffman: Ancora un esempio

- Costruzione dell'albero di Huffman:



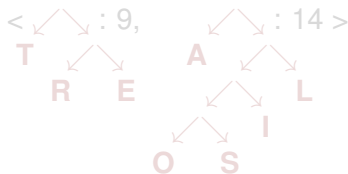
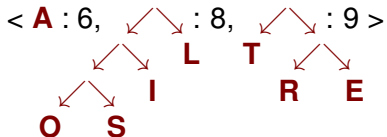
Codifica di Huffman: Ancora un esempio

- Costruzione dell'albero di Huffman:



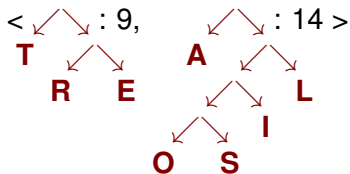
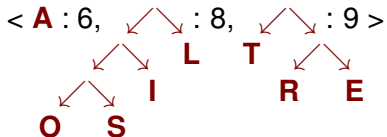
Codifica di Huffman: Ancora un esempio

- Costruzione dell'albero di Huffman:



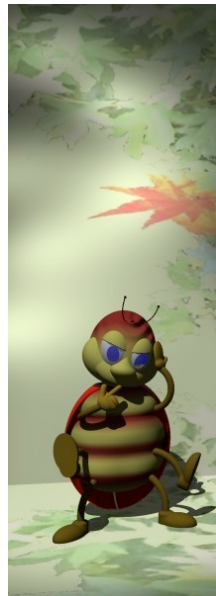
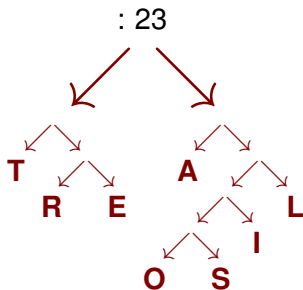
Codifica di Huffman: Ancora un esempio

- Costruzione dell'albero di Huffman:



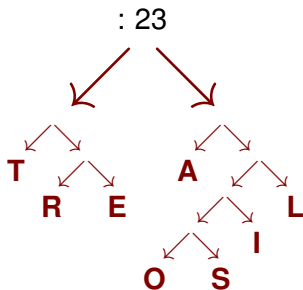
Codifica di Huffman: Ancora un esempio

- Albero di Huffman:



Codifica di Huffman: Ancora un esempio

- Albero di Huffman:

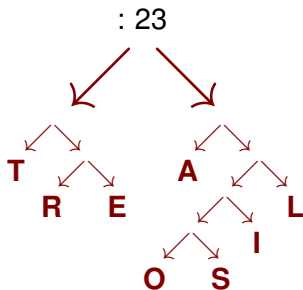


A	
E	
I	
O	

L	
R	
S	
T	

Codifica di Huffman: Ancora un esempio

- Albero di Huffman:

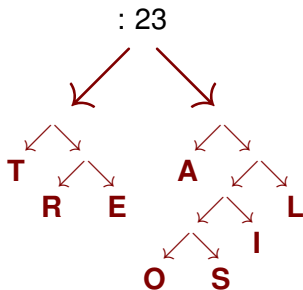


A	10
E	
I	
O	

L	
R	
S	
T	

Codifica di Huffman: Ancora un esempio

- Albero di Huffman:

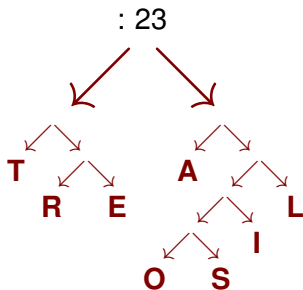


A	10
E	011
I	
O	

L	
R	
S	
T	

Codifica di Huffman: Ancora un esempio

- Albero di Huffman:

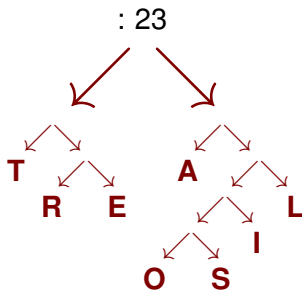


A	10
E	011
I	1101
O	

L	
R	
S	
T	

Codifica di Huffman: Ancora un esempio

- Albero di Huffman:

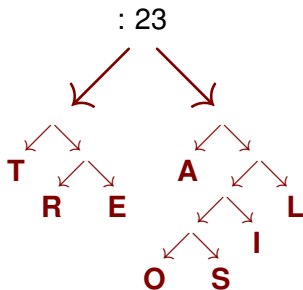


A	10
E	011
I	1101
O	11000

L	
R	
S	
T	

Codifica di Huffman: Ancora un esempio

- Albero di Huffman:

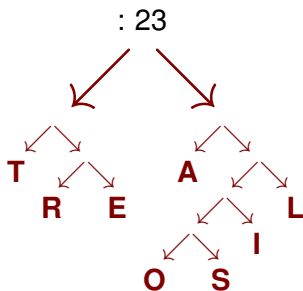


A	10
E	011
I	1101
O	11000

L	111
R	
S	
T	

Codifica di Huffman: Ancora un esempio

- Albero di Huffman:

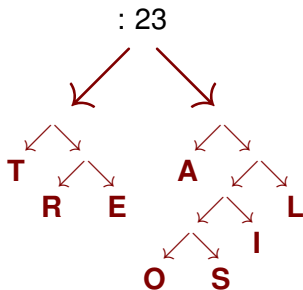


A	10
E	011
I	1101
O	11000

L	111
R	010
S	
T	

Codifica di Huffman: Ancora un esempio

- Albero di Huffman:

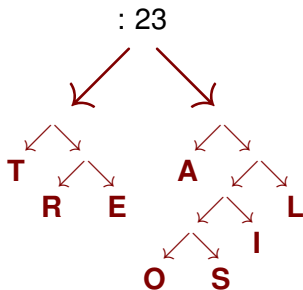


A	10
E	011
I	1101
O	11000

L	111
R	010
S	11001
T	

Codifica di Huffman: Ancora un esempio

- Albero di Huffman:



A	10
E	011
I	1101
O	11000

L	111
R	010
S	11001
T	00

Codifica di Huffman: Ancora un esempio

- Codifica del testo:

ILTEATROALLIETALASERATA

- Codifica di Huffman:

A	10
E	011
I	1101
O	11000

L	111
R	010
S	11001
T	00

- Rappresentazione della sequenza (65 bit):

110111100011100001011000101111111101
01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Codifica del testo:

ILTEATROALLIETALASERATA

- Codifica di Huffman:

A	10
E	011
I	1101
O	11000

L	111
R	010
S	11001
T	00

- Rappresentazione della sequenza (65 bit):

110111100011100001011000101111111101
01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Codifica del testo:

ILTEATROALLIETALASERATA

- Codifica di Huffman:

A	10
E	011
I	1101
O	11000

L	111
R	010
S	11001
T	00

- Rappresentazione della sequenza (65 bit):

110111100011100001011000101111111101
01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Codifica del testo:

ILTEATROALLIETALASERATA

- Codifica di Huffman:

A	10
E	011
I	1101
O	11000

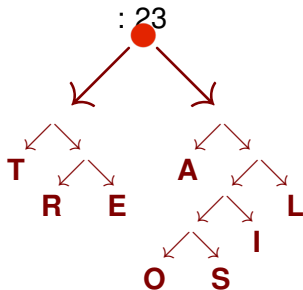
L	111
R	010
S	11001
T	00

- Rappresentazione della sequenza (65 bit):

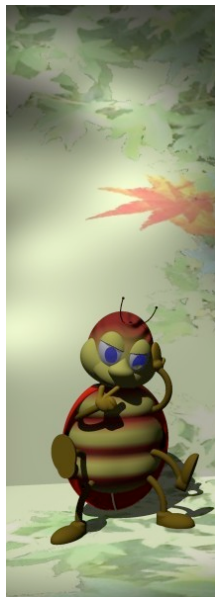
110111100011100001011000101111111101
01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

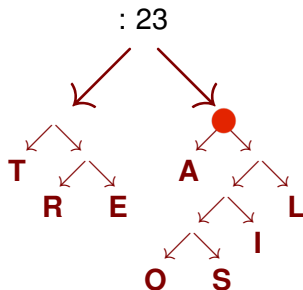


110111100011100001011000101111111101
0110010111101100101...



Codifica di Huffman: Ancora un esempio

- Decodifica?

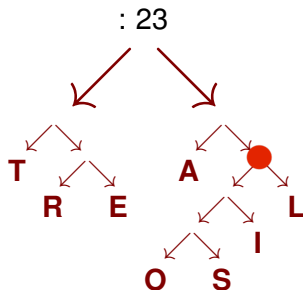


110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

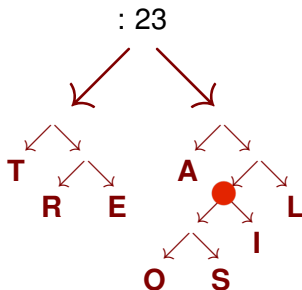


110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

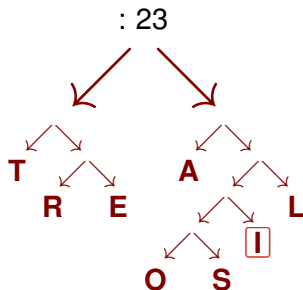


110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?



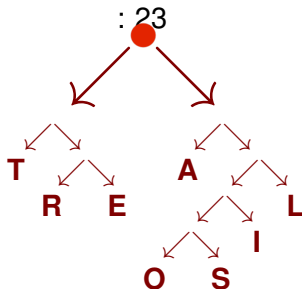
110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

I



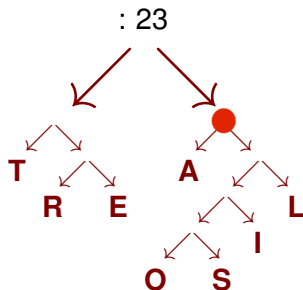
110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

I



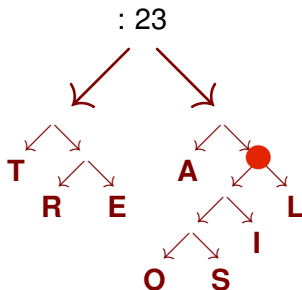
1101110001110000101100010111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

I



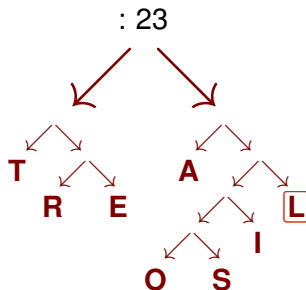
110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

I



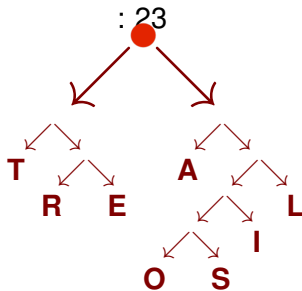
110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

I L



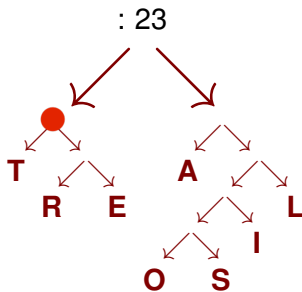
110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

I L



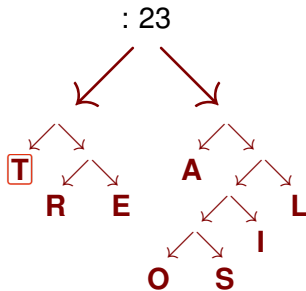
110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

I L



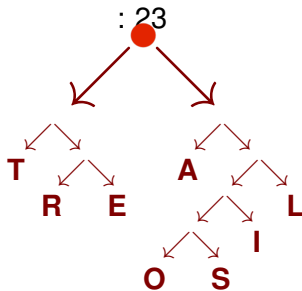
110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

I L T



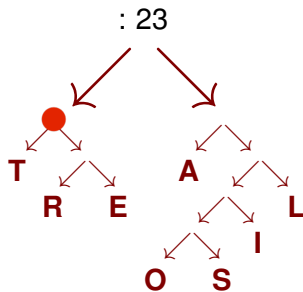
110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

I L T



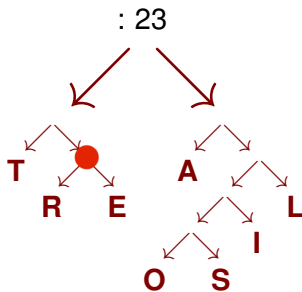
110111100011100001011000101111111101

01100101111011001011010100010

Codifica di Huffman: Ancora un esempio

- Decodifica?

I L T ...



1101111000**1**1100001011000101111111101

01100101111011001011010100010

Domande?

