



وسائط متعددة المحاضرة الرابعة

أ.د. فادي غصنه



الترميز Coding والضغط Compression

- يقوم ضغط البيانات على فكرة تخفيض حجم البيانات باستخدام خصائص هذه البيانات (احتوائها على قيم مكرره) أو خصائص الإدراك البشري (وجود بعض البيانات الزائده لا يؤثر على قيمة هذه البيانات) أو الاثنين معا.
- يمكن ضغط الصوت الرقمي والصور والفيديو والنصوص.
- يؤدي هذا الضغط إلى:
 - ✓ توفير الوقت عند الإرسال.
 - ✓ تخفيض عرض النطاق الترددي اللازم.
 - ✓ التوفير في حجم التخزين.
- يجب أن يعاد فك ضغط البيانات المضغوطة ليتم عرضها (أو سماعها)، وهذا ما يتطلب معالجة إضافية.



➤ الضغط :

هو عملية تحويل لصيغة المعطيات (ترميز) لتحتوي على عدد أقل من البتات، مما يعطي إمكانية تخزين وإرسال معطيات أكبر.

➤ يعد الترميز والضغط المكونين الأساسيين لنجاح الملتيميديا وخاصةً من أجل تكامل الفيديو في تطبيقات الوسائط المتعددة

➤ يوجد نوعين من الضغط:

✓ الضغط دون خسارة (Lossless Compression)

✓ الضغط مع خسارة (Lossy Compression)

➤ يسمح الضغط مع وجود خسارة بمعدل ضغط أكبر من الضغط دون خسارة



طرائق ضغط البيانات

Data Compression Methods

Lossless Methods

TEXT

Lossy Methods

Image

Audio

Video



طرائق ضغط البيانات

□ الضغط دون خسارة (Lossless Compression):

- يحافظ على البيانات بعد الضغط، أي أن البيانات بعد الضغط هي نفسها قبل الضغط
- ✓ لا يضيع أي قسم من البيانات
- تكون خوارزميات الضغط وفك الضغط متعاكستين
- تحذف البيانات الزائدة في عملية الضغط وتعاد لتضاف في عملية فك الضغط
- تستخدم هذه الطريقة في حالة الحاجة للمحافظة على كامل البيانات دون فقد أي جزء منها، كما هو الحال في النصوص



طرائق ضغط البيانات

□ الضغط مع خسارة (Lossy Compression):

➤ تستخدم في الحالات التي تحوي تغيرات غير ملحوظة

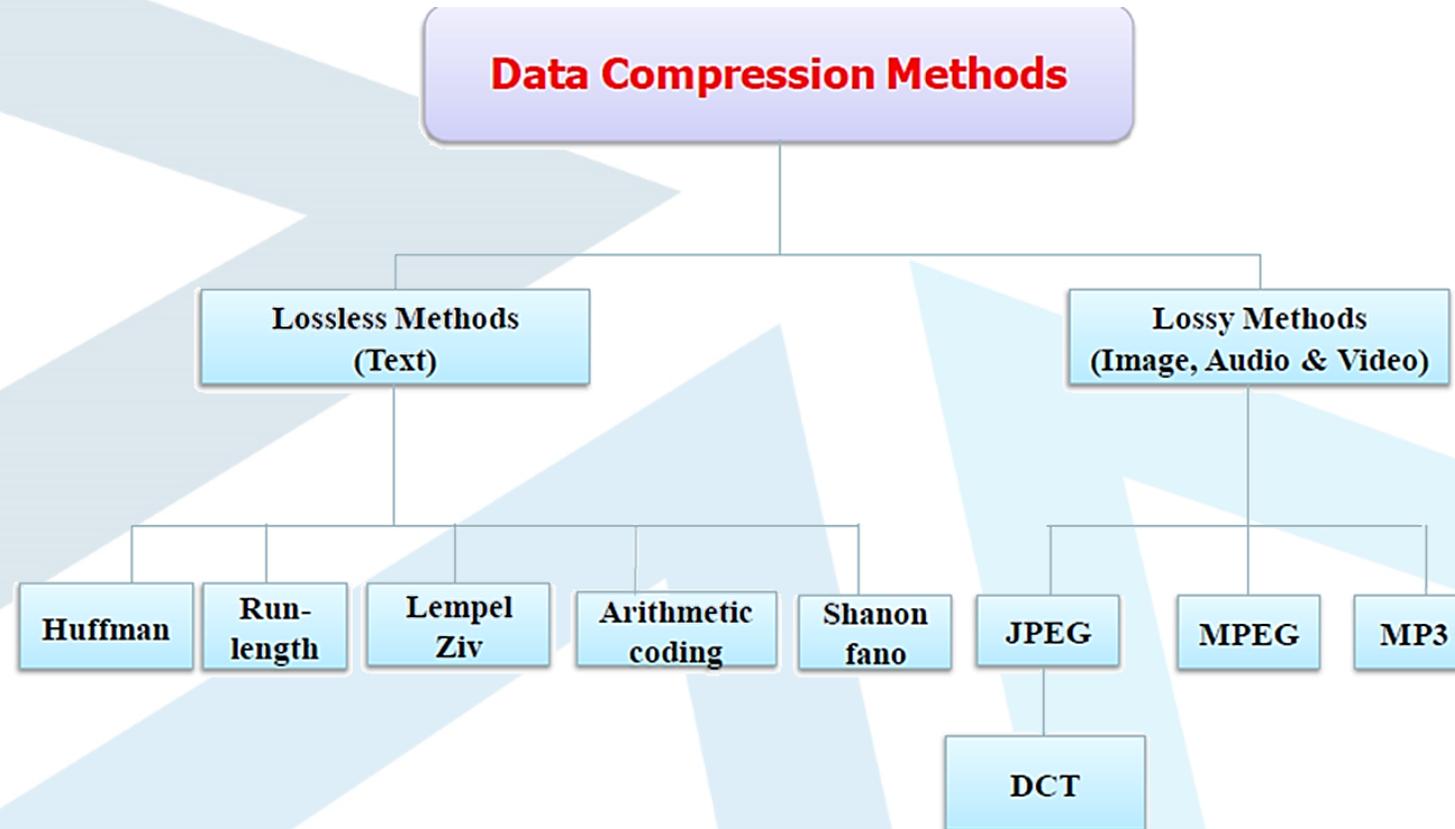
✓ حيث يتم استغلال التكرار والإدراك البشري

➤ هذه الطريقة أرخص وتحتاج إلى زمن ومساحة تخزين أقل من الضغط دون خسارة وتستخدم لإرسال الملايين من البتات في الصور والفيديو.

➤ لا تسمح باستعادة الإشارة بشكل دقيق، ولكن يمكن استخدامها من أجل المعطيات التي لا تكون حساسة جداً لفقد بعض المعطيات مثل الصورة والفيديو والصوت.



طرائق ضغط البيانات



ترميز هوفمان (Huffman code)

- يستخدم ترميز هوفمان لإرسال النصوص بطريقة مضغوطة
- يعتمد على عدد مرات تكرار كل حرف من الحروف
- يعد أساسيا في العديد من نظم التشغيل الشائعة، حيث أن بعض البرامج تستخدم ترميز هوفمان كطريقة أما البعض الآخر فيستخدمه كخطوة من عدة مراحل.
- يقسم ترميز هوفمان إلى قسمين هما:
 - هوفمان الساكن
 - هوفمان الديناميكي (المتكيف)



ترميز هوفمان الساكن

✓ يستخدم في حالة النصوص التي تعرف المحارف المستخدمة فيها مع تكرارات حصولها، حيث ترمز المحارف بكلمات ترميز متغيرة الطول بحيث تمثل كلمات الترميز الأقصر المحارف ذات تكرار الحصول الأكبر.

✓ للحصول على ترميز هوفمان يتم اتباع الخطوات الآتية:

(١) يتم ترتيب الاحتمالات من الأكبر (الأعلى) إلى الأصغر (الأدنى).

(٢) نجمع أصغر احتمالين في كل خطوة ونقوم بإعادة ترتيب الاحتمالات من جديد حتى الانتهاء، أي الوصول إلى مجموع نهائي يساوي الواحد .

(٣) كل احتمالين نقوم بجمعهما نرمزهما ب ١,٠ بشكل متناوب .

(٤) نحدد ترميز كل محرف بحيث نتبع ترميز الاحتمال الممثل للمحرف بحيث يبدأ ترتيب الخانات من اليمين إلى اليسار.



مثال (١) على ترميز هوفمان

- ليكن لدينا الحروف الآتية A,B,C,D,E واحتمالات ورودها مبينة بالجدول أدناه

المحرف	A	B	C	D	E
احتمال تكراره	0.25	0.25	0.2	0.15	0.15

- والمطلوب :

١ . ترميز هذه الحروف بترميز هوفمان

٢ . حساب نسبة الضغط



الحل:

A	0.25
B	0.25
C	0.2
D	0.15
E	0.15

١. يتم ترتيب الاحتمالات من الأكبر (الأعلى) إلى الأصغر (الأدنى)

٢. نجمع أصغر احتمالين في كل خطوة ونعيد الترتيب حتى الانتهاء.

بمقارنة مجموع أصغر احتمالين (٠.٣) مع الاحتمالات المتبقية

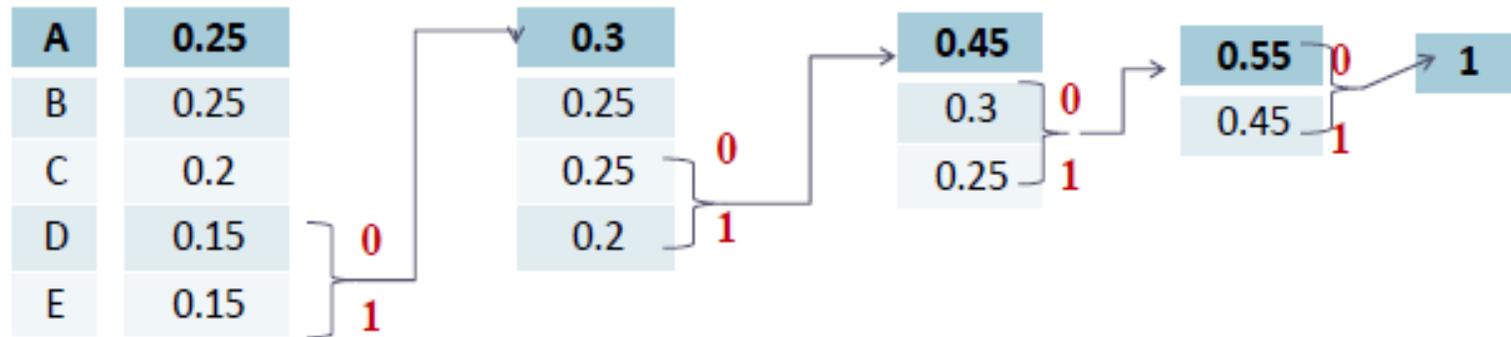
(٠.٢٥, ٠.٢٥, ٠.٢) نلاحظ أنها أكبر فيكون لدينا ترتيب جديد



وهكذا



٣- كل احتمالين نقوم بجمعهما نرمزهما ب ٠, ١ بشكل متناوب.



٤- نرمز كل محرف من المحارف

A	0.25	0.3	0.45	0.55	0	1
B	0.25	0.25	0.3	0	0.45	1
C	0.2	0.25	0	0.25	1	
D	0.15	0	0.2	1		
E	0.15	1				

فيكون:

نرمز المحرف A ب 01



A	0.25	0.3	0.45	0.55	0	1
B	0.25	0.25	0.3	0.45	0	1
C	0.2	0.25	0	0.25	1	
D	0.15	0.2	1			
E	0.15					

نرمز المحرف B بـ 10

A	0.25	0.3	0.45	0.55	0	1
B	0.25	0.25	0.3	0.45	0	1
C	0.2	0.25	0	0.25	1	
D	0.15	0.2	1			
E	0.15					

نرمز المحرف C بـ 11



A	0.25	0.3	0.45	0.55	0	1
B	0.25	0.25	0.3	0	0.45	1
C	0.2	0.25	0	0.25	1	
D	0.15	0	0.2	1		
E	0.15	1				

نرمز المحرف D بـ 000

A	0.25	0.3	0.45	0.55	0	1
B	0.25	0.25	0.3	0	0.45	1
C	0.2	0.25	0	0.25	1	
D	0.15	0	0.2	1		
E	0.15	1				

نرمز المحرف E بـ 001



➤ بالنتيجة يكون:

codeword	X	P				
01	A	0.25	0.3	0.45	0.55	1
10	B	0.25	0.25	0.3	0.45	
11	C	0.2	0.25	0.25		
000	D	0.15	0.2			
001	E	0.15				

٢. لحساب نسبة الضغط :

نسبة الضغط = (متوسط عدد بتات ترميز هوفمان $\times 100$) / متوسط عدد بتات ترميز الآسكي

✓ نعلم أن متوسط عدد البتات اللازمة لترميز كل حرف بالآسكي هو ٧ بتات

✓ نقوم بحساب متوسط عدد البتات الناتج عن استخدام ترميز هوفمان الساكن = مجموع ناتج ضرب عدد البتات

الممثلة لترميز كل حرف باحتمال تكرار هذا الحرف.



المحرف	ترميزه	احتمال تكراره	عدد بتات الترميز
A	01	0.25	2
B	10	0.25	2
C	11	0.2	2
D	000	0.15	3
E	001	0.15	3

متوسط عدد البتات في ترميز هوفمان :

$$2 \times 0.25 + 2 \times 0.25 + 2 \times 0.2 + 3 \times 0.15 + 3 \times 0.15 = 2.3 \text{ bits}$$

$$32.85\% = 7 / (2.3 * 100) = \text{نسبة الضغط}$$



مثال (٢) على ترميز هوفمان

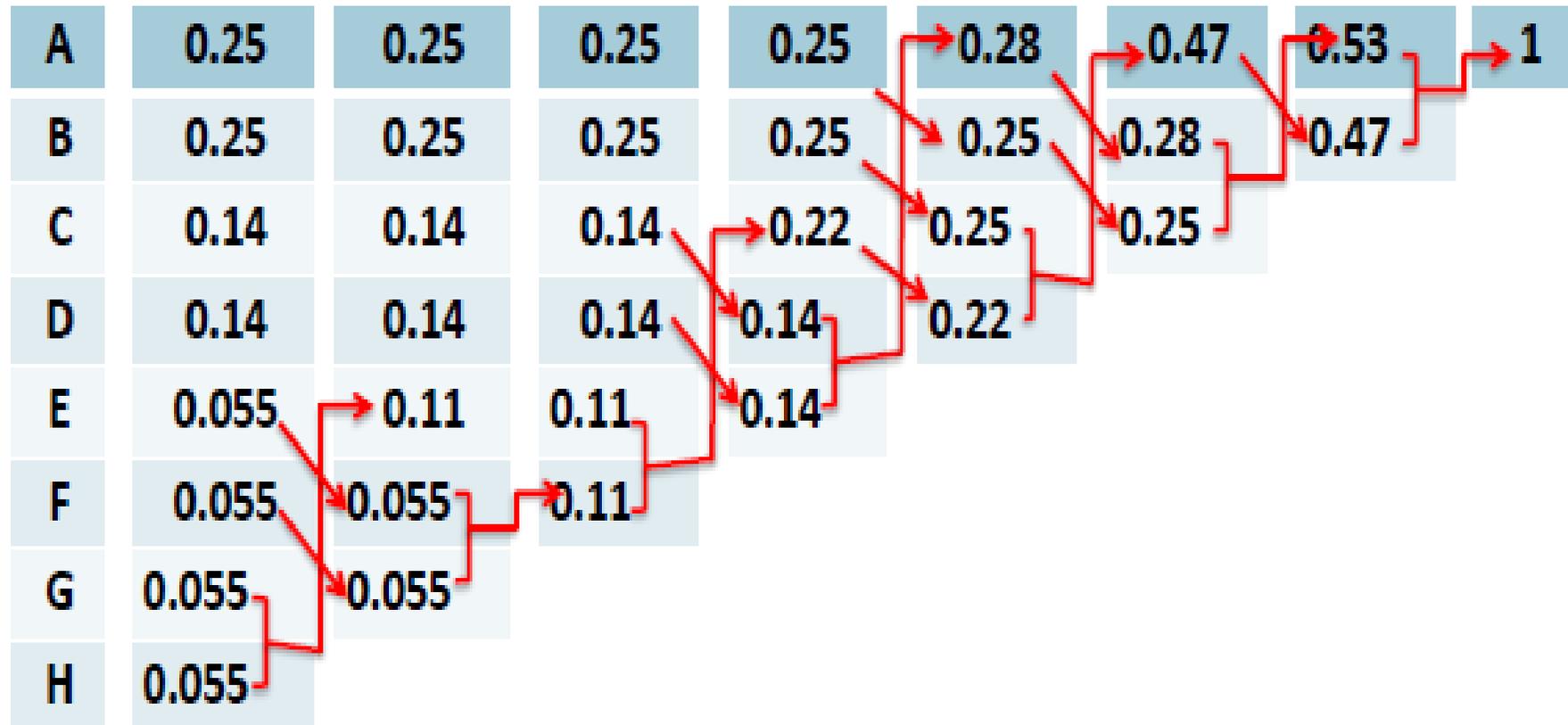
ليكن لدينا المحارف الآتية واحتمالات ورودها ظاهرة في الجدول الآتي

المحرف	A	B	C	D	E	F	G	H
الاحتمال	0.25	0.25	0.14	0.14	0.055	0.055	0.055	0.055

والمطلوب ترميز هذه المحارف بترميز هوفمان الساكن وحساب نسبة الضغط.



الحل



A	0.25	0.25	0.25	0.25	0.28	0.47	0.53	0	1
B	0.25	0.25	0.25	0.25	0.25	0.28	0.47	0	1
C	0.14	0.14	0.14	0.22	0.25	0.25	0	1	
D	0.14	0.14	0.14	0.14	0.22	0	1		
E	0.055	0.11	0.11	0.14	0	1			
F	0.055	0.055	0.11	0	1				
G	0.055	0.055	0	1					
H	0.055	0	1						

الحل



فيكون

المحرف	الاحتمال	الترميز	عدد بتات الترميز
A	0.25	01	2
B	0.25	10	2
C	0.14	000	3
D	0.14	001	3
E	0.055	1110	4
F	0.055	1111	4
G	0.055	1100	4
H	0.055	1101	4



٢- حساب نسبة الضغط:

✓ لدينا متوسط عدد البتات المستخدمة في الترميز الآسكي: ٧ بت

✓ متوسط عدد البتات في ترميز هوفمان:

$$2*0.25+2*0.25+3*0.14+3*0.14+4*0.055+4*0.055+4*0.055+4*0.055=2.72 \text{ bits}$$

فتكون نسبة الضغط:

$$(2.72*100)/7=38.86\%$$



شجرة هوفمان (Huffman Tree)

- نكوّن أوراقا بعدد الأحرف التي لدينا، ووضع في كل ورقة تكرار الحرف الذي تمثله.
- اختيار عقدتان لهما أصغر تكرارين لتكونا ورقتان لعقدة جديدة
- نضع في هذه العقدة الجديدة مجموع التكرارين (تكراري العقدين صاحبتنا أقل تكرارين).
- طبق هذه العملية حتى تحصل على عقدة واحدة وهي الجذر.
- نرقّم كل عقده وفقاً لموضعها، إذا كانت يمين عقدة أخرى تأخذ الرقم ١ وإذا كانت شمال عقدة أخرى تأخذ الرقم ٠ .



مثال (٣) على ترميز هوفمان

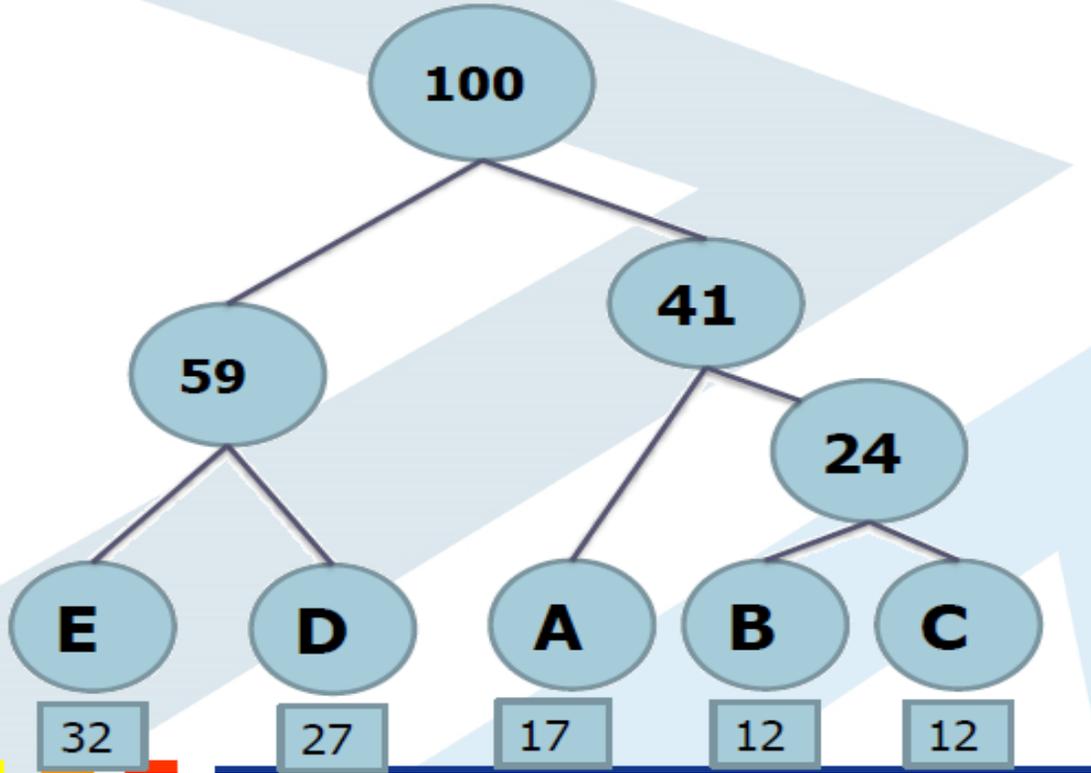
ليكن لدينا نص مكون من مجموعة من الحروف مع ترددات تكرارها كما هو موضح بالجدول، والمطلوب نسبة ضغط النص باستخدام ترميز هوفمان الساكن .

Character	A	B	C	D	E
Frequency	17	12	12	27	32



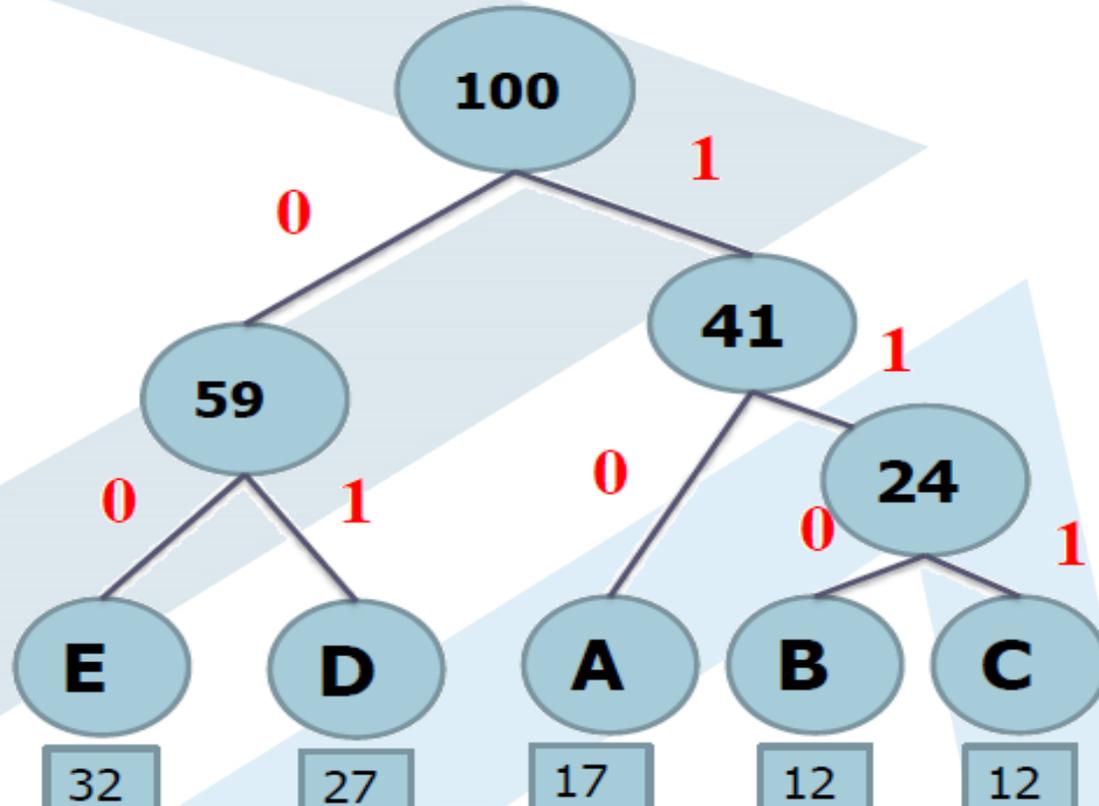
الحل:

- ١ . نقوم أولاً بترتيب الحروف حسب تردد التكرار تنازلياً من اليسار إلى اليمين
- ٢ . ثم في كل خطوة بجمع أصغر ترددين في المرحلة:



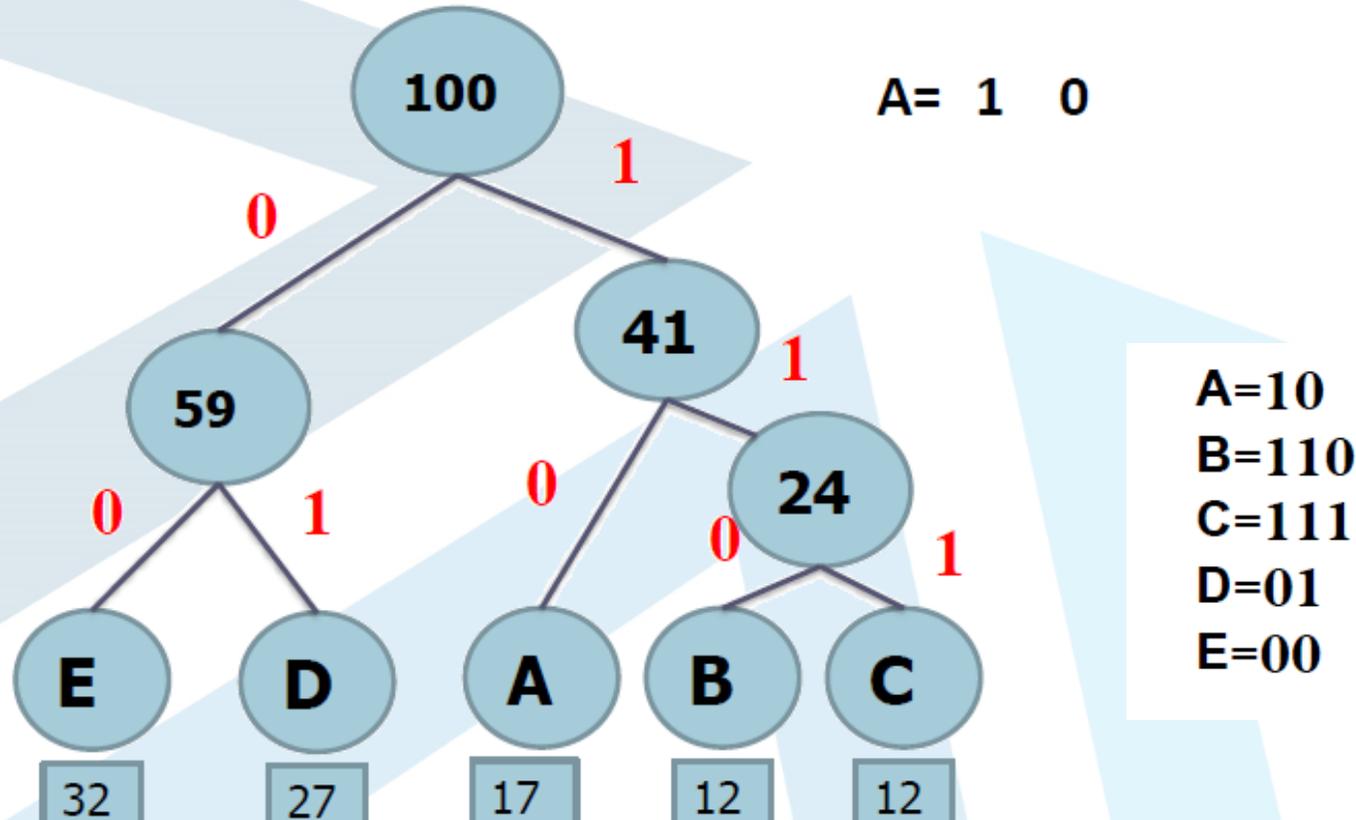
الحل:

٣- نرّمز أفرع الشجرة ب ٠ و ١ بحيث الفرع اليميني يأخذ ترميز الواحد والفرع اليساري يأخذ ترميز الصفر وهكذا على جميع فروع الشجرة.



الحل:

٤- نقوم بترميز الحروف: من الأعلى (من اليسار) إلى الأسفل (إلى اليمين)
فيكون:



٢. حساب نسبة الضغط :

عدد البتات في ترميز هوفمان :

المحرف	ترميزه	عدد بتات الترميز	عدد مرات التكرار
A	10	2	17
B	110	3	12
C	111	3	12
D	01	2	27
E	00	2	32

$$17 \times 2 + 12 \times 3 + 12 \times 3 + 27 \times 2 + 32 \times 2 = 224 \text{bits}$$

بينما يكون لدينا عدد البتات المستخدمة في ترميز الآسكي:

$$(32 + 27 + 12 + 12 + 17 \times 7) = 700 \text{bits}$$

فتكون نسبة الضغط:

$$224 * 100 / 700 = 32\%$$



مثال (٤) على ترميز هوفمان

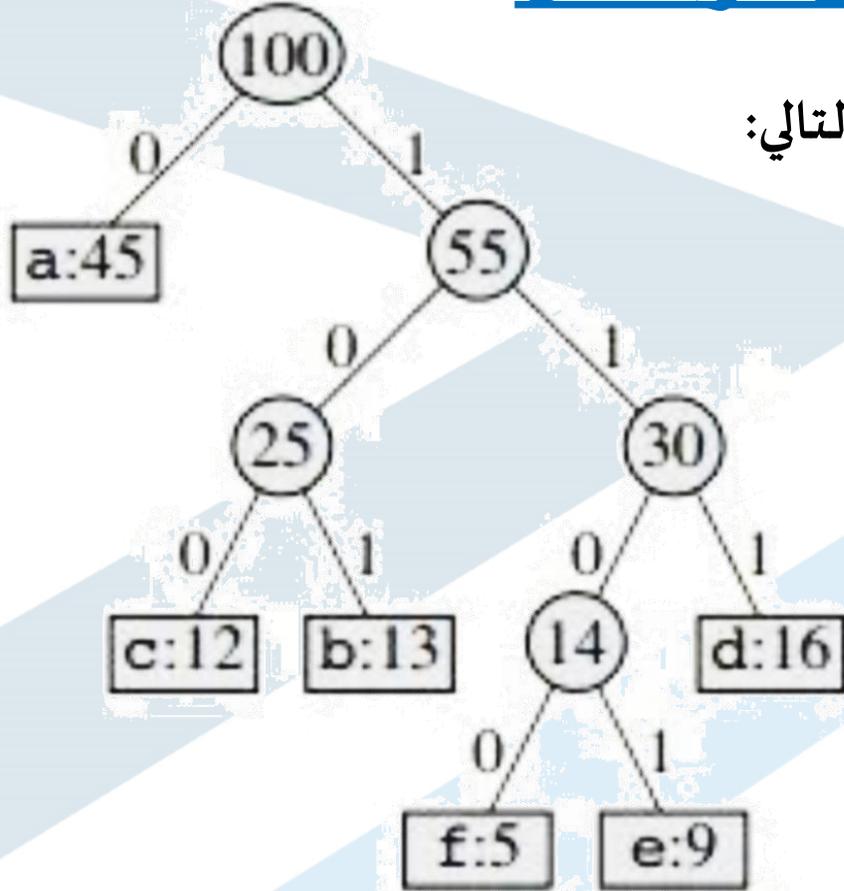
ليكن لدينا نص مكون من مجموعة من الحروف مع ترددات تكرارها كما هو موضح بالجدول،
والمطلوب نسبة ضغط النص باستخدام ترميز هوفمان الساكن .

Character	a	b	c	d	e	f
Frequency	45	13	12	16	9	5



مثال (٤) على ترميز هوفمان

نطبق خطوات طريقة شجرة هوفمان فنحصل على الشكل التالي:



مثال (٤) على ترميز هوفمان

بعد أن نقوم بترميز الحروف: من الأعلى (من اليسار) إلى الأسفل (إلى اليمين) يكون لدينا الترميز التالي:

Character	Coding
a	0
b	101
c	100
d	111
e	1101
f	1100



مثال (٤) على ترميز هوفمان

المحرف	ترميزه	عدد بتات الترميز	عدد مرات التكرار
a	0	1	45
b	101	3	13
c	100	3	12
d	111	3	16
e	1101	4	9
f	1100	4	5

٢. حساب نسبة الضغط :

✓ عدد البتات في ترميز هوفمان :

$$45 \times 1 + 13 \times 3 + 12 \times 3 + 16 \times 3 + 9 \times 4 + 5 \times 4 = 224 \text{bits}$$

✓ بينما يكون لدينا عدد البتات المستخدمة في ترميز الآسكي :

$$(45 + 13 + 12 + 16 + 9 + 5) \times 7 = 700 \text{bits}$$

فتكون نسبة الضغط:

$$224 * 100 / 700 = 32\%$$



مثال (٥) على ترميز هوفمان

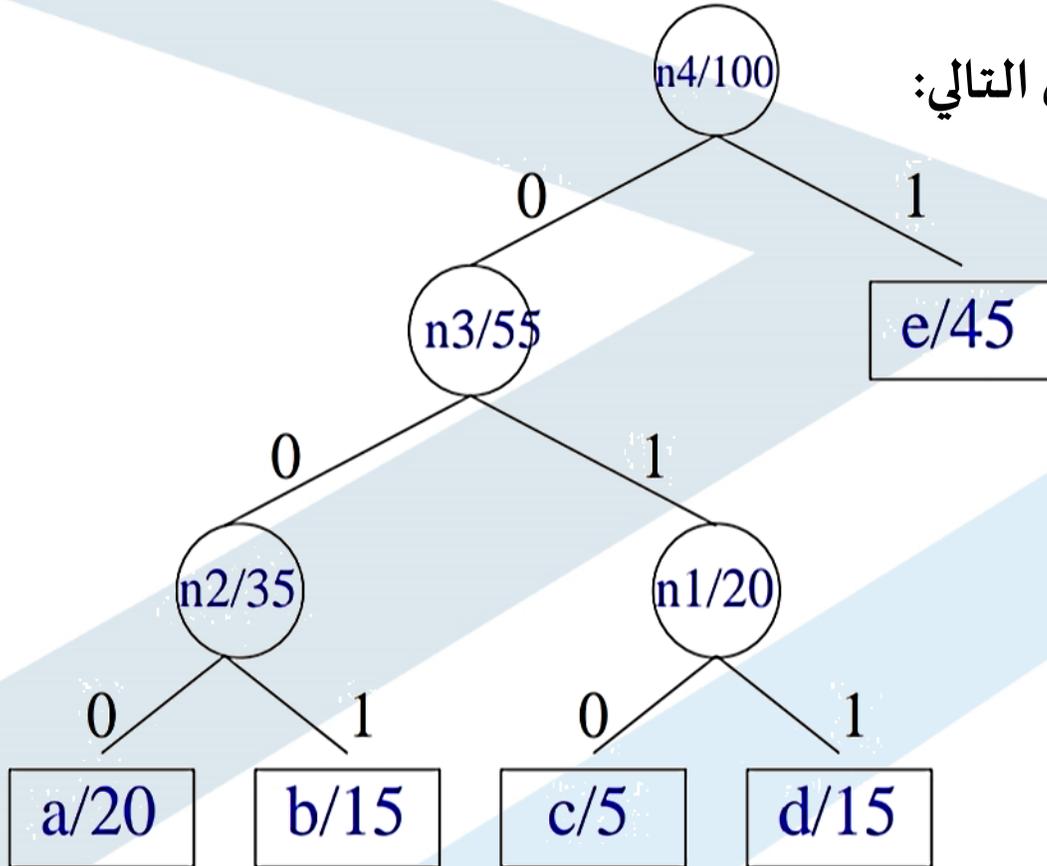
ليكن لدينا نص مكون من مجموعة من الحروف مع ترددات تكرارها كما هو موضح بالجدول،
والمطلوب نسبة ضغط النص باستخدام ترميز هوفمان الساكن .

Character	a	b	c	d	e
Frequency	20	15	5	15	45



مثال (٥) على ترميز هوفمان

نطبق خطوات طريقة شجرة هوفمان فنحصل على الشكل التالي:



مثال (٥) على ترميز هوفمان

بعد أن نقوم بترميز الحروف: من الأعلى (من اليسار) إلى الأسفل (إلى اليمين) يكون لدينا الترميز التالي:

Character	Coding
a	000
b	001
c	010
d	011
e	1



مثال (٥) على ترميز هوفمان

المحرف	ترميزه	عدد بتات الترميز	عدد مرات التكرار
a	000	3	20
b	001	3	15
c	010	3	5
d	011	3	15
e	1	1	45

٢. حساب نسبة الضغط :

✓ عدد البتات في ترميز هوفمان :

$$20 \times 3 + 15 \times 3 + 5 \times 3 + 15 \times 3 + 45 \times 1 = 210 \text{ bits}$$

✓ بينما يكون لدينا عدد البتات المستخدمة في ترميز الآسكي:

$$(20 + 15 + 5 + 15 + 45) \times 7 = 700 \text{ bits}$$

فتكون نسبة الضغط:

$$210 \times 100 / 700 = 30\%$$



مثال (٦) على ترميز هوفمان

ليكن لدينا النص : MISSISSIPPI RIVER

والمطلوب ضغط النص باستخدام ترميز هوفمان الساكن . احسب نسبة الضغط

الحل:

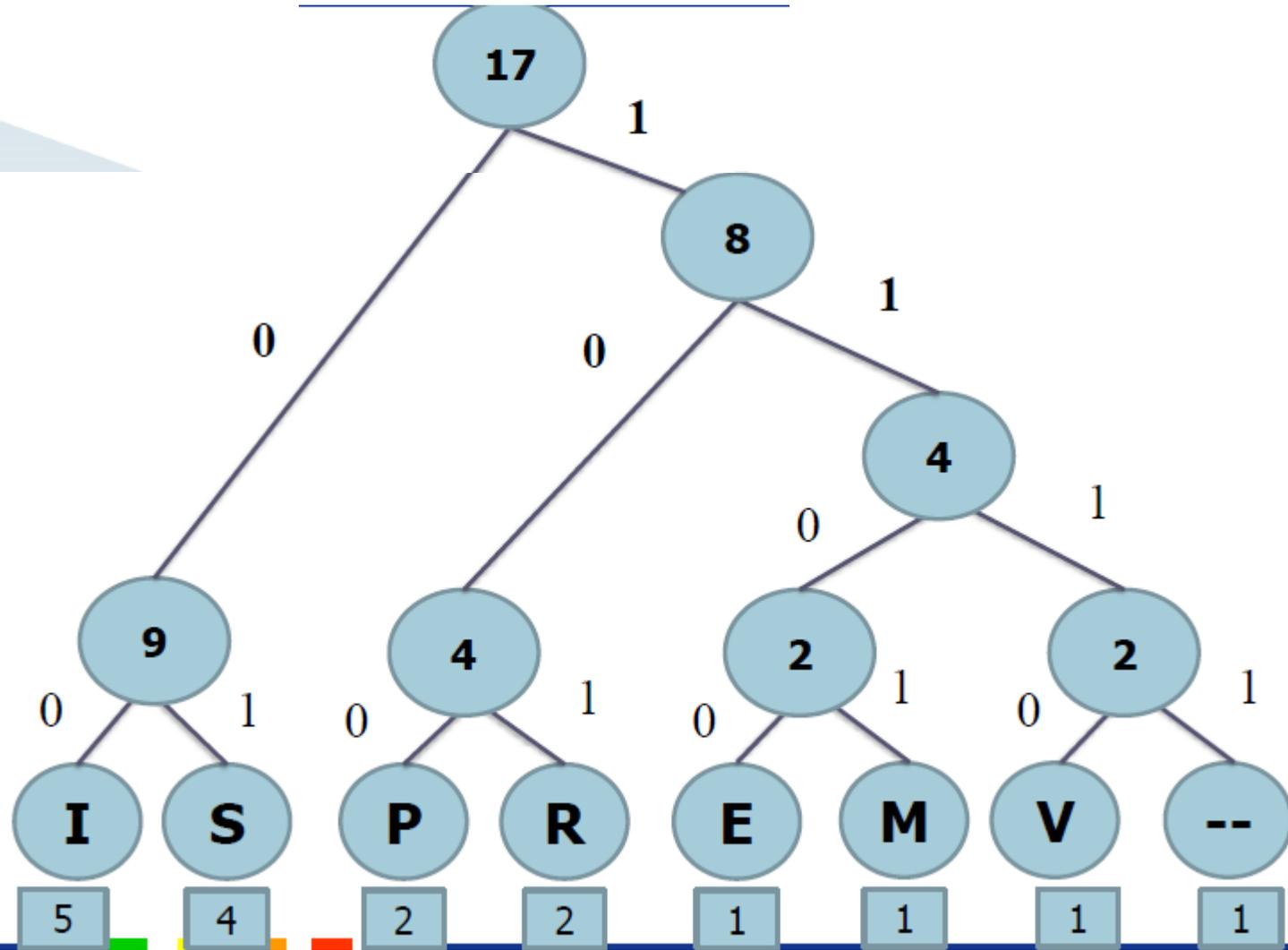
نأخذ من النص كل محرف وعدد مرات تكراره

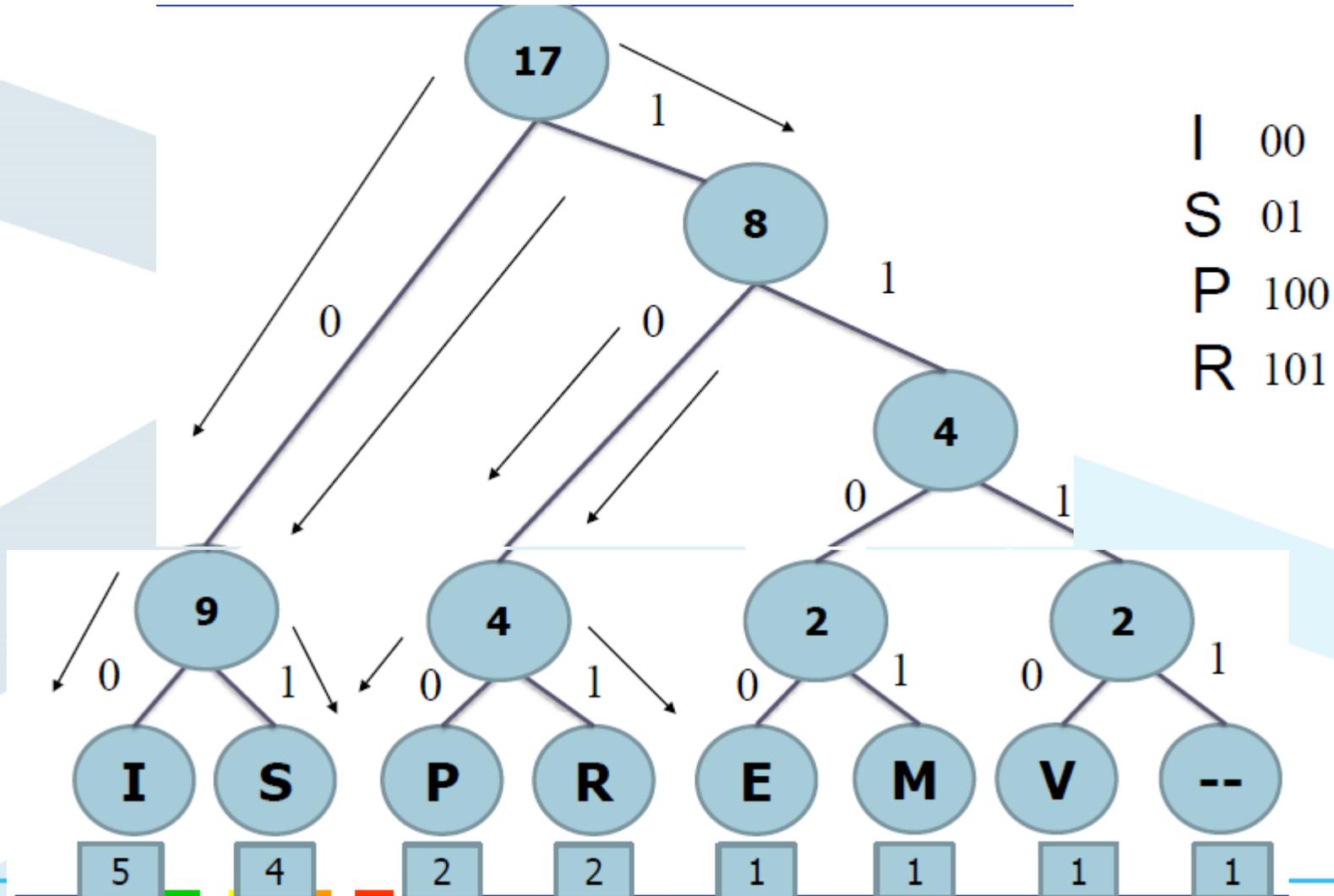
المحرف الأكثر تكراراً ترميزه أقل طولاً

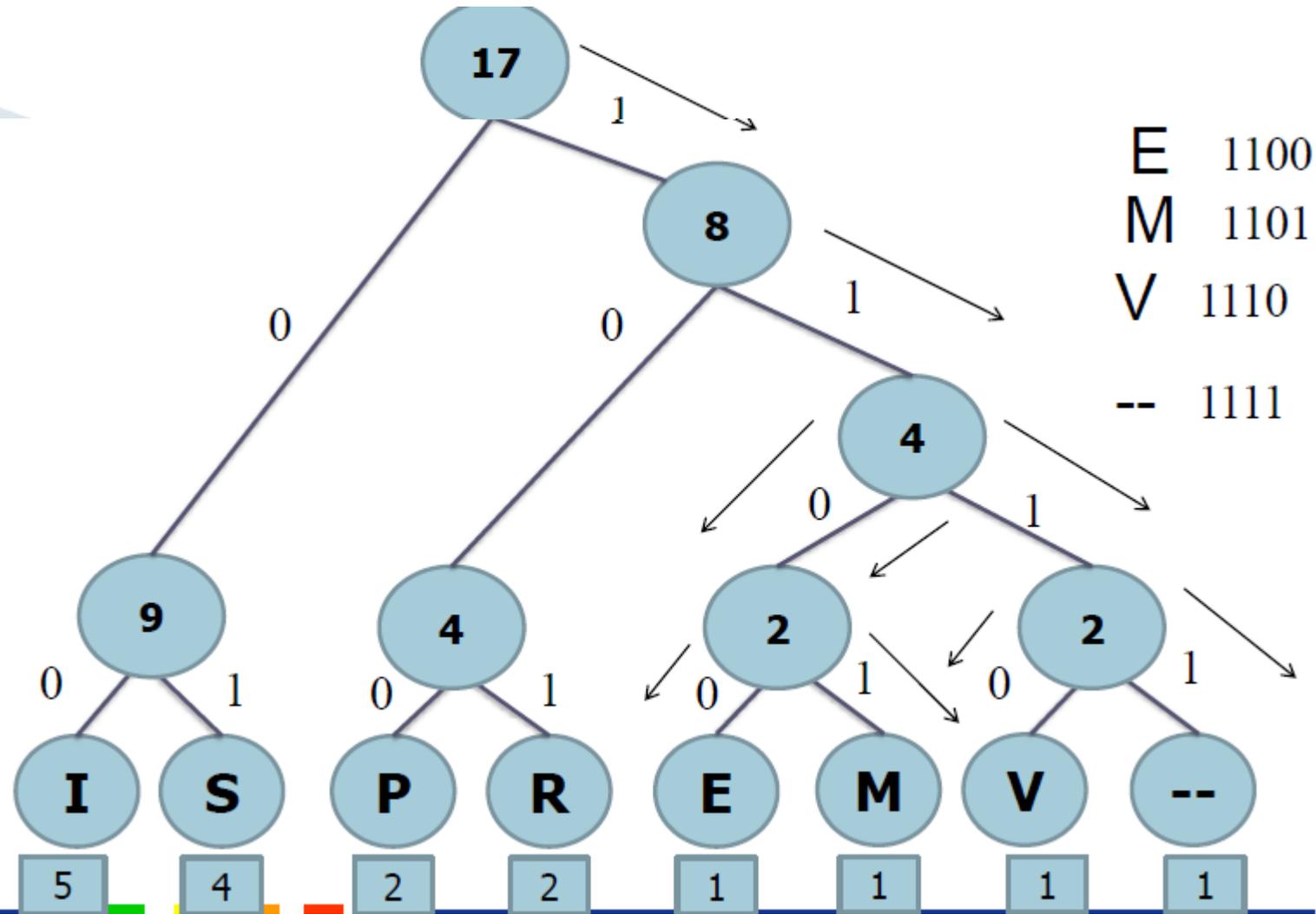
نقوم برسم الشجرة

M	1
I	5
S	4
P	2
R	2
V	1
E	1
--	1









I	00	V	1110
S	01	M	1101
P	100	E	1100
R	101	--	1111

اعتماداً على ترميز الحروف ➤

يكون ترميز النص المعطى: ➤

MISSISSIPPI RIVER

1101000101000101001001000011111010011101100101

حساب نسبة الضغط: ➤

✓ عدد البتات في ترميز هوفمان:

$$5 \times 2 + 4 \times 2 + 2 \times 3 + 2 \times 3 + 1 \times 4 + 1 \times 4 + 1 \times 4 + 1 \times 4 = 46 \text{ bits}$$

✓ عدد البتات المستخدمة لترميز النص بالأسكي: $17 \times 7 = 119 \text{ bits}$

✓ نسبة الضغط: $46 \times 100 / 119 = 38.6 \%$

