

# Information theory and coding

## نظرية المعلومات والترميز

مدرسة المقرر

د. بشرى علي معلا

## مفردات المحاضرة

➤ مقدمة إلى الترميز

➤ طرائق الضغط

➤ ترميز هوفمان

✓ ترميز هوفمان الساكن

✓ ترميز هوفمان الديناميكي

## مقدمة إلى الترميز

➤ الترميز هو: تحويل معلومات الدخل إلى إشارات رقمية حتى يتمكن النظام من نقلها

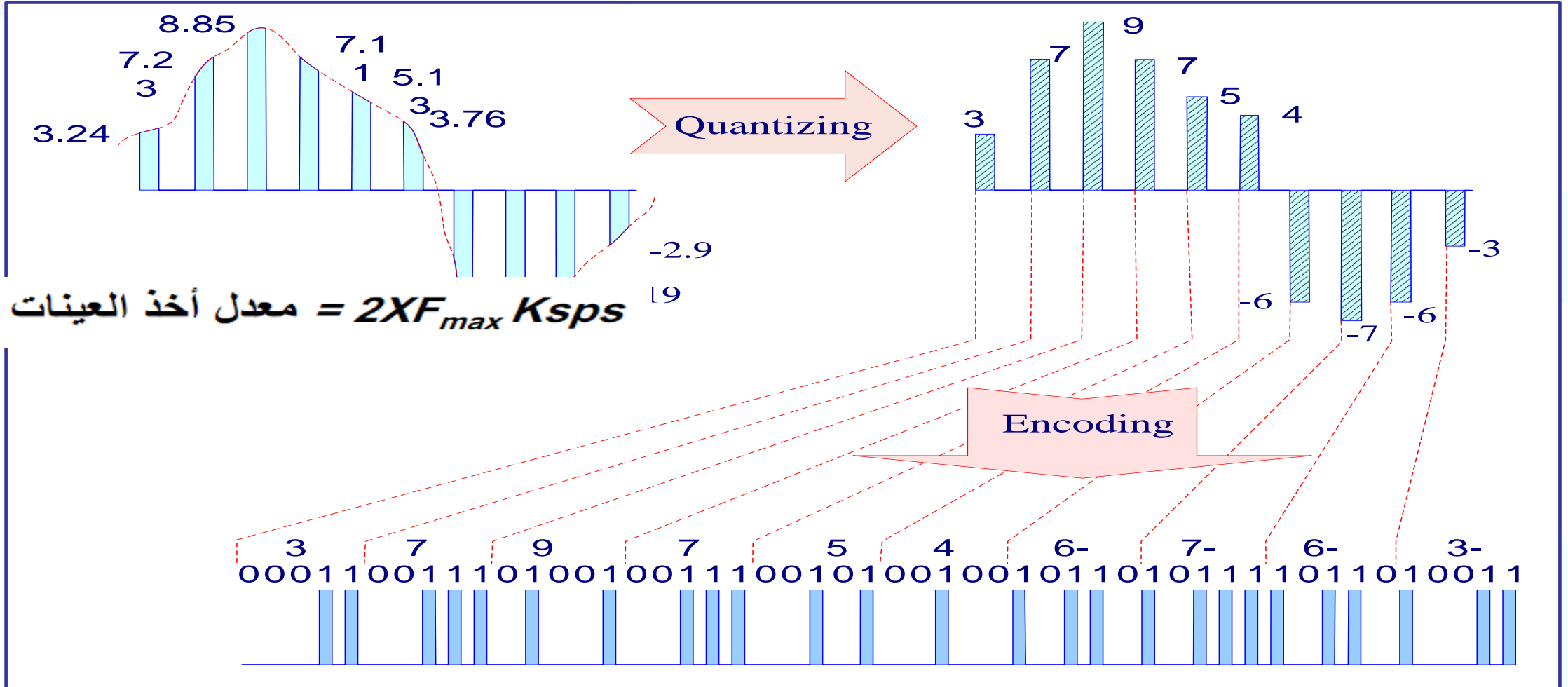
➤ يوجد للترميز صنفين أساسيين:

### ١. ترميز المصدر (Resource Coding):

الهدف منه جعل معدل معلومات المصدر (معدل انتروبيا المصدر) قريباً من سعة القناة مثلاً: ضغط البيانات  
(Data Compression)

### ٢. ترميز القناة (Channel Coding):

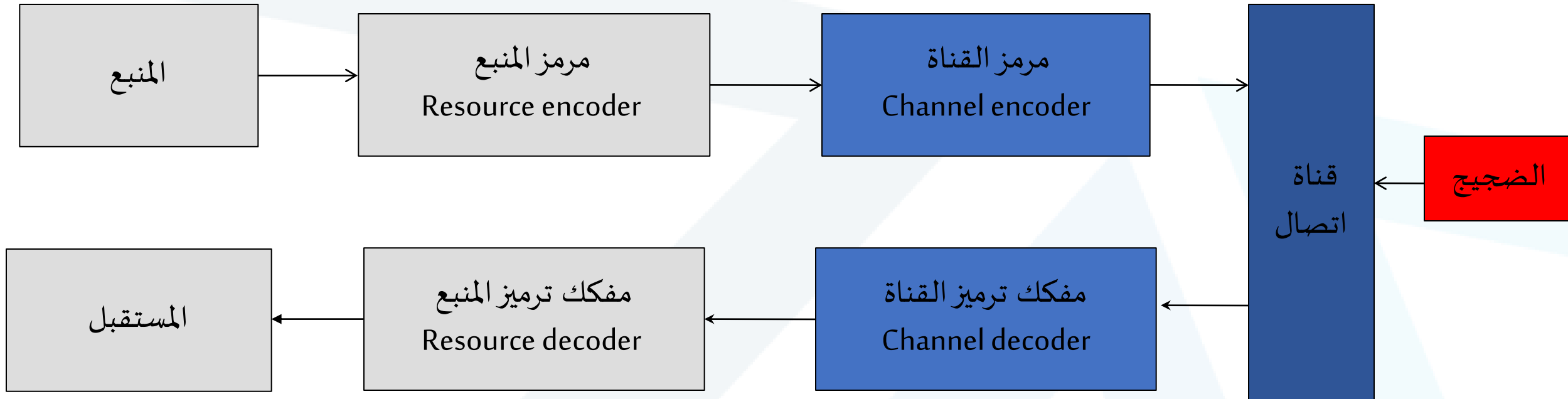
الهدف منه تصحيح الأخطاء و ذلك لتحويل القناة غير الموثوقة إلى قناة موثوقة





جامعة  
المنارة  
MANARA UNIVERSITY

## هدف الترميز



## نظام اتصال: ترميز القناة و المنبع

## ترميز المنبع

➤ الضغط :

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

➤ يوجد نوعان 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 الضغط (1/3)

### Data Compression Methods

**Lossless Methods**

**TEXT**

**Lossy Methods**

**Image**

**Audio**

**Video**

## طرائق الضغط (2/3)

### □ 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 الضغط (3/3)

□ 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.

## مفاهيم أساسية

➤ الطول الوسطي لكلمة الترميز بالعلاقة:

$$L_{average} = \sum_{i=1}^n L_i \cdot P_i$$

$L_i$  عدد بتات كلمة الترميز.

$P_i$  احتمال ظهور هذه الكلمة الترميز  $i$

➤ مردود الترميز:

$$\eta = \frac{L_{min}}{L_{average}}$$

حيث  $L_{min}$  طول الأصغري لكلمة الترميز

➤ الفائض:

$$\rho = 1 - \eta$$

$$l_{min} = \frac{H(S)}{\log_2 m}$$

## طرائق ضغط البيانات

### Data Compression Methods

#### Lossless Methods (Text)

#### Lossy Methods (Image, Audio & Video)

Huffman

Run-length

Lempel Ziv

Arithmetic  
coding

Shanon  
fano

JPEG

MPEG

MP3

## ترميز هوفمان (Huffman coding)

➤ يستخدم ترميز هوفمان لإرسال النصوص بطريقة مضغوطة

➤ يعتمد على عدد مرات تكرار كل حرف من الحروف

➤ يعد أساسياً 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 بـ 0,1 بشكل متناوب.
٤. نحدد ترميز كل محرف بحيث نتبع ترميز الاحتمال الممثل للمحرف بحيث يبدأ ترتيب الخانات من اليمين إلى اليسار.

## مثال على ترميز هوفمان (١/٩)

ليكن لدينا الحارف 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 الحارف بترميز هوفمان

٢. حساب نسبة الضغط

## مثال (١) على ترميز هوفمان (٢/٩)

➤ الحل:

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

<b>A</b>	<b>0.25</b>
B	0.25
C	0.2
D	0.15
E	0.15

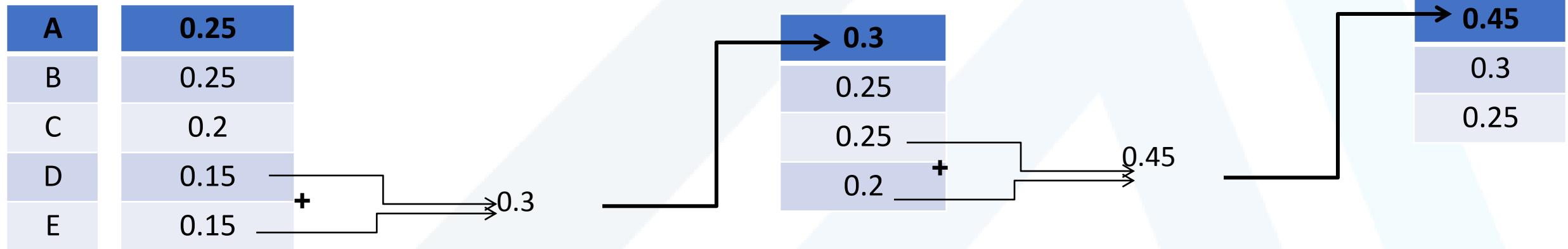
## مثال (١) على ترميز هوفمان (٣/٩)

➤ الحل:

١. يتم ترتيب الاحتمالات من الأكبر إلى الأصغر.

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

بمقارنة مجموع أصغر احتمالين (0.3) مع الاحتمالات المتبقية (0.25, 0.25, 0.2) نلاحظ أنها أكبر فيكون لدينا ترتيب جديد



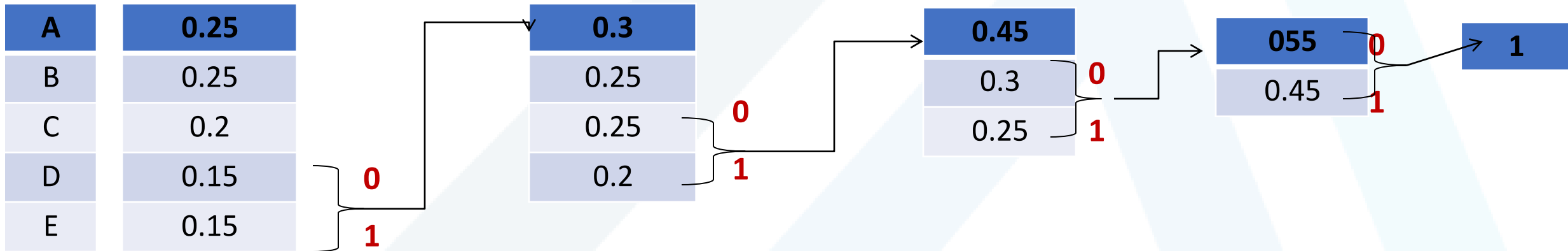
وهكذا



## مثال (١) على ترميز هوفمان (٤/٩)

➤ الحل:

١. يتم ترتيب الاحتمالات من الأكبر إلى الأصغر.
٢. نجمع أصغر احتمالين في كل خطوة ونعيد الترتيب حتى الانتهاء.
٣. كل احتمالين نقوم بجمعهما نرمزهما بـ 0,1 بشكل متناوب.



## مثال (١) على ترميز هوفمان (٥/٩)

➤ الحل:

١. يتم ترتيب الاحتمالات من الأكبر إلى الأصغر.
٢. نجمع أصغر احتمالين في كل خطوة ونعيد الترتيب حتى الانتهاء.
٣. كل احتمالين نقوم بجمعهما نرمزهما بـ 0,1 بشكل متناوب.
٤. نرمز كل محرف من المحارف

فيكون:

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	0.25	0.25	0	1
D	0.15	0.2			0	1
E	0.15				1	

نرمز المحرف A بـ 01

## مثال (١) على ترميز هوفمان (٦/٩)

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	0.25	0.25	0	1
D	0.15	0.2			0	1
E	0.15				1	

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	0.25	0.25	0	1
D	0.15	0.2			0	1
E	0.15				1	

نرمز المحرف C بـ 11

## مثال (١) على ترميز هوفمان (٧/٩)

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

Diagram illustrating the Huffman coding process for Example 1. The table shows the frequencies of characters A, B, C, D, and E. The process involves repeatedly merging the two smallest frequencies until a single root node is reached. The final tree structure is shown with the root node 1. The internal nodes are 0.55, 0.45, 0.3, 0.25, and 0.2. The leaf nodes are A, B, C, D, and E. The final tree structure is shown with the root node 1. The internal nodes are 0.55, 0.45, 0.3, 0.25, and 0.2. The leaf nodes are A, B, C, D, and E. The final tree structure is shown with the root node 1. The internal nodes are 0.55, 0.45, 0.3, 0.25, and 0.2. The leaf nodes are A, B, C, D, and E.

نرمز المحرف D بـ 000

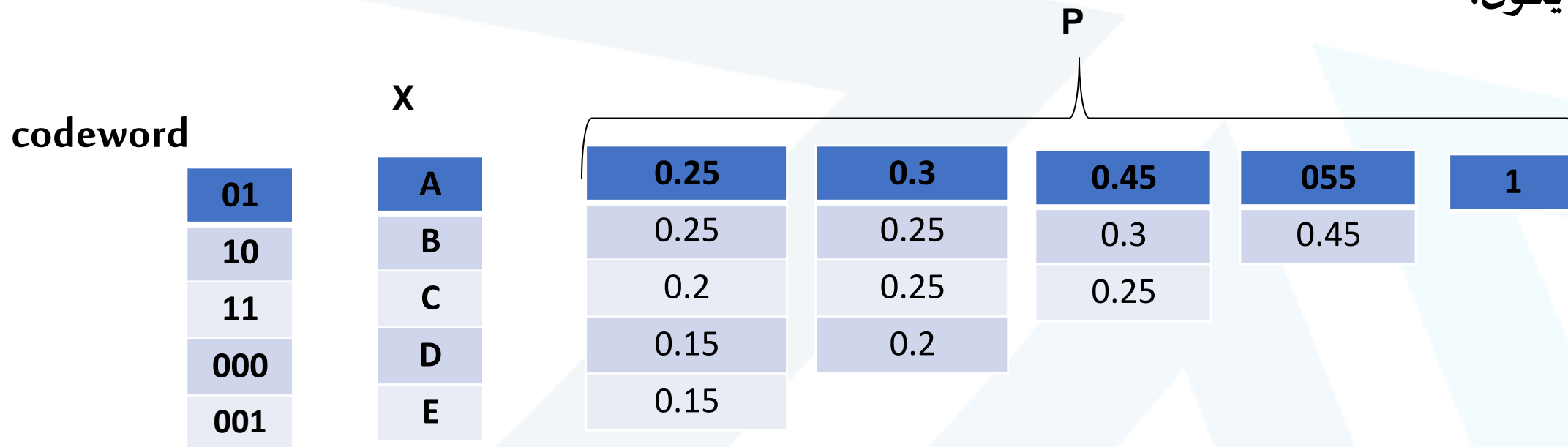
A	0.25	0.3	0.45	0.55	1
B	0.25	0.25	0.3	0.45	
C	0.2	0.25	0.25		
D	0.15	0.2			
E	0.15				

Diagram illustrating the Huffman coding process for Example 1. The table shows the frequencies of characters A, B, C, D, and E. The process involves repeatedly merging the two smallest frequencies until a single root node is reached. The final tree structure is shown with the root node 1. The internal nodes are 0.55, 0.45, 0.3, 0.25, and 0.2. The leaf nodes are A, B, C, D, and E. The final tree structure is shown with the root node 1. The internal nodes are 0.55, 0.45, 0.3, 0.25, and 0.2. The leaf nodes are A, B, C, D, and E.

نرمز المحرف E بـ 001

## مثال (١) على ترميز هوفمان (٨/٩)

➤ بالنتيجة يكون:



## مثال (١) على ترميز هوفمان (٩/٩)

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 7 بتات

✓ نقوم بحساب متوسط عدد البتات الناتج عن استخدام ترميز هوفمان الساكن ومجموع ناتج ضرب عدد البتات الممثلة لترميز كل حرف باحتمال تكرار هذا الحرف

متوسط عدد البتات في ترميز هوفمان :

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

$$\text{نسبة الضغط} = 2.3 \times 100 / 7 = 32.85\%$$

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

## مثال (٢) على ترميز هوفمان (5/1)

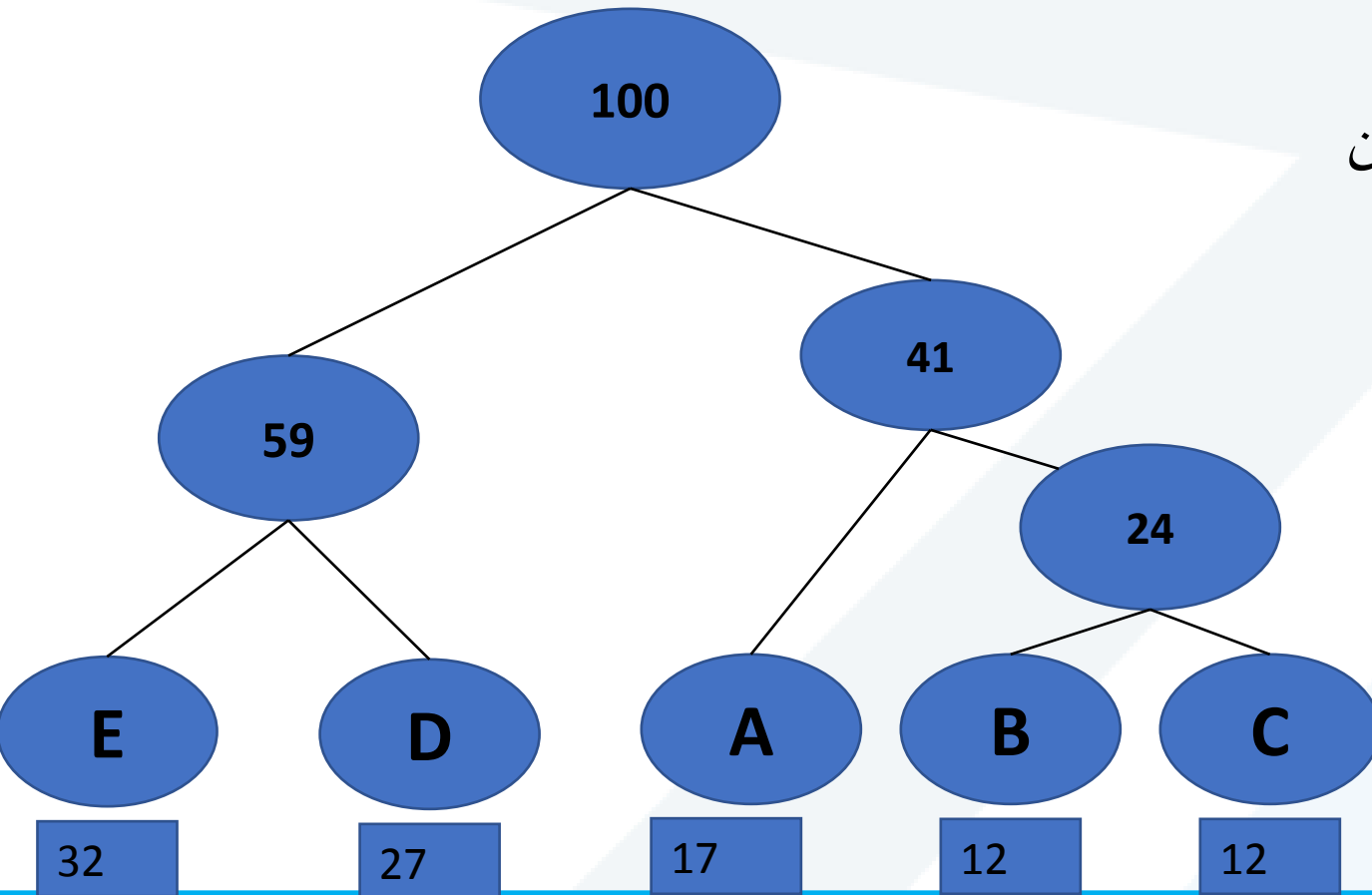
ليكن لدينا نص مكون من مجموعة من الحروف مع ترددات تكرارها كما هو موضح بالجدول، والمطلوب ضغط النص باستخدام ترميز هوفمان الساكن.

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

## مثال (٢) على ترميز هوفمان (5/2)

الحل:

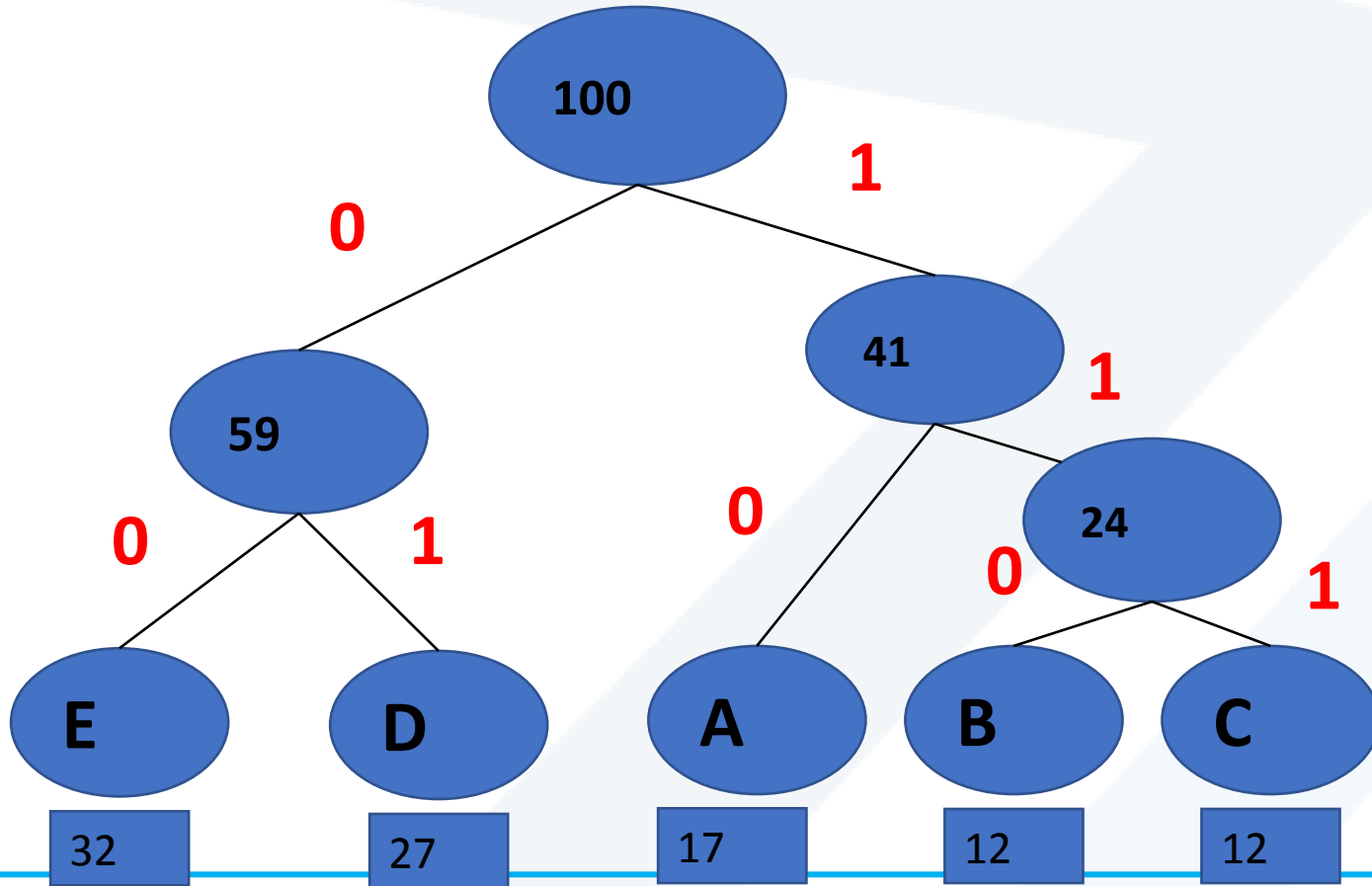
١. نقوم أولاً بترتيب المحارف حسب تردد التكرار تنازلياً من اليسار إلى اليمين
٢. ثم في كل خطوة بجمع أصغر ترددين في المرحلة:





## مثال (٢) على ترميز هوفمان (5/3)

٣. نرسم أفرع الشجرة بـ 0 و 1 بحيث الفرع اليميني يأخذ ترميز الواحد والفرع اليساري يأخذ ترميز الصفر وهكذا على جميع فروع الشجرة.



## مثال (٢) على ترميز هوفمان (5/4)

٤- نقوم بترميز المحارف: (نبدأ بالترميز لكل محرف من الأسفل إلى الأعلى وتكتب من اليمين إلى اليسار)

فيكون:

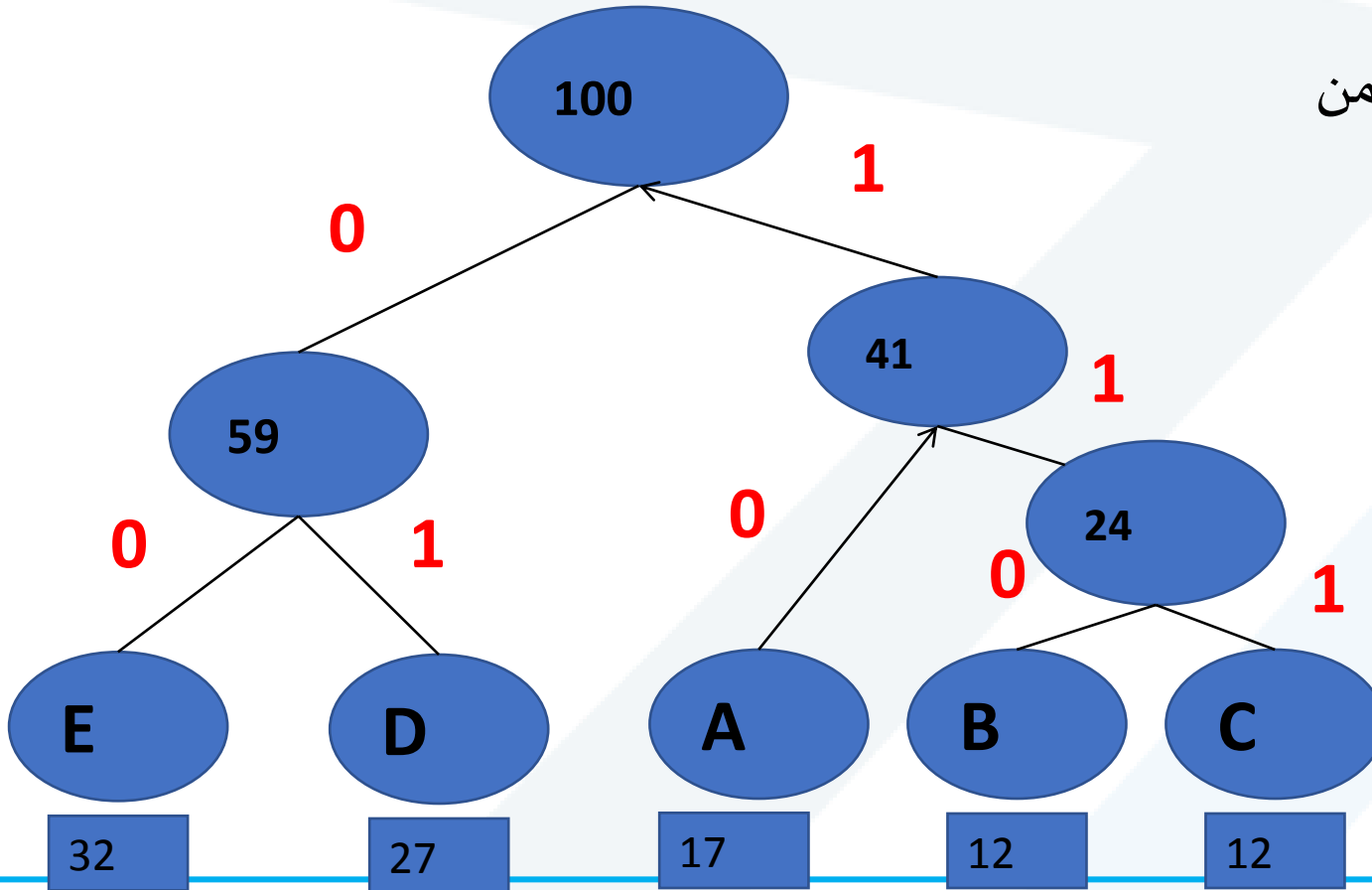
A=10

B=110

C=111

D=01

E=00



## مثال (٢) على ترميز هوفمان (5/5)

٢. حساب نسبة الضغط :

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

✓ عدد البتات في ترميز هوفمان :

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

✓ بينما يكون لدينا عدد البتات المستخدمة في ترميز الآسكي:

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

فتكون نسبة الضغط:  $224 \times 100 / 700 = 32\%$

## ترميز هوفمان الديناميكي

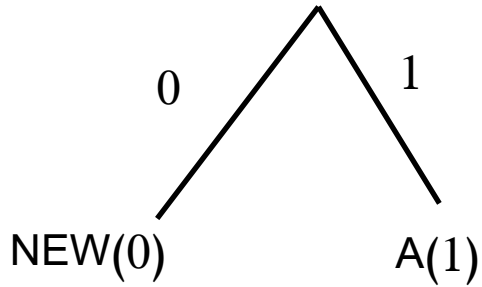
➤ يستخدم مع النصوص التي تتغير من تحويل إلى آخر عندها ستتغير مجموعة كلمات الترميز ديناميكياً، حيث أن المرسل والمستقبل يشكلان شجرة هوفمان بشكل ديناميكي

➤ آلية العمل:

- ✓ يرسل المحرف في المرة الأولى بطريقة غير مضغوطة أما في المرات التالية فيرسل بطريقة مضغوطة
- ✓ في كل مرة نرسل فيها محرف جديد نضيف NEW إلى الشجرة

## خطوات رسم الشجرة ترميز هوفمان الديناميكي (1/2)

١. يمثل أول محرف بالسلسلة على شكل فرعي شجرة :



- الفرع اليميني يمثل المحرف مع وزنه ويكون وزن الفرع 1
- الفرع اليساري يمثل NEW(0) والتي تعني أنه محرف جديد ويكون وزن الفرع 0
- ونقطة التقاء الفرعين تحمل قيمة مجموع الوزنين

٢. عند ورود محرف جديد يتم إضافة فرع جديد في أسفل الشجرة وإجراء تحديث على جميع قيم الشجرة

٣. عند ورود محرف سبق و وجد ضمن الشجرة يجب إجراء تحديث لوزنه الموجود بإضافة واحد إلى الوزن القديم و إجراء تحديث على جميع قيم الشجرة

## خطوات رسم الشجرة ترميز هوفمان الديناميكي (2/2)

٤. يؤخذ الترميز دائماً قبل الموازنة فمع كل عملية إضافة لفرع جديد أو تحديث لفرع موجود يجب إجراء موازنة للشجرة وفق الملاحظات الآتية:

- ضمن نفس المستوى:

وزن الفرع اليميني يجب أن يكون أكبر أو يساوي وزن الفرع اليساري وإلا تتم عملية التبديل بين اليمين واليسار

- ضمن نفس الجهة:

وزن الفرع الأعلى يجب أن يكون أكبر أو يساوي وزن الفرع الأدنى وإلا تتم عملية التبديل بين الأعلى والأدنى

- بين جهتين:

وزن الفرع الأدنى على اليمين يجب أن يكون **أصغر أو يساوي** وزن الفرع الأدنى على اليسار وإلا تتم عملية التبديل بين اليمين واليسار

**ملاحظة: الأولوية للتبديل بين اليمين واليسار في حال وجد فرعين وليس بين الأعلى والأسفل لنفس الفرع**

## مثال (١) (1/10)

ليكن لدينا السلسلة AADCCDD وفق الترميز الأولية الآتية :

NEW:0

A:00001

C:00011

D:00100

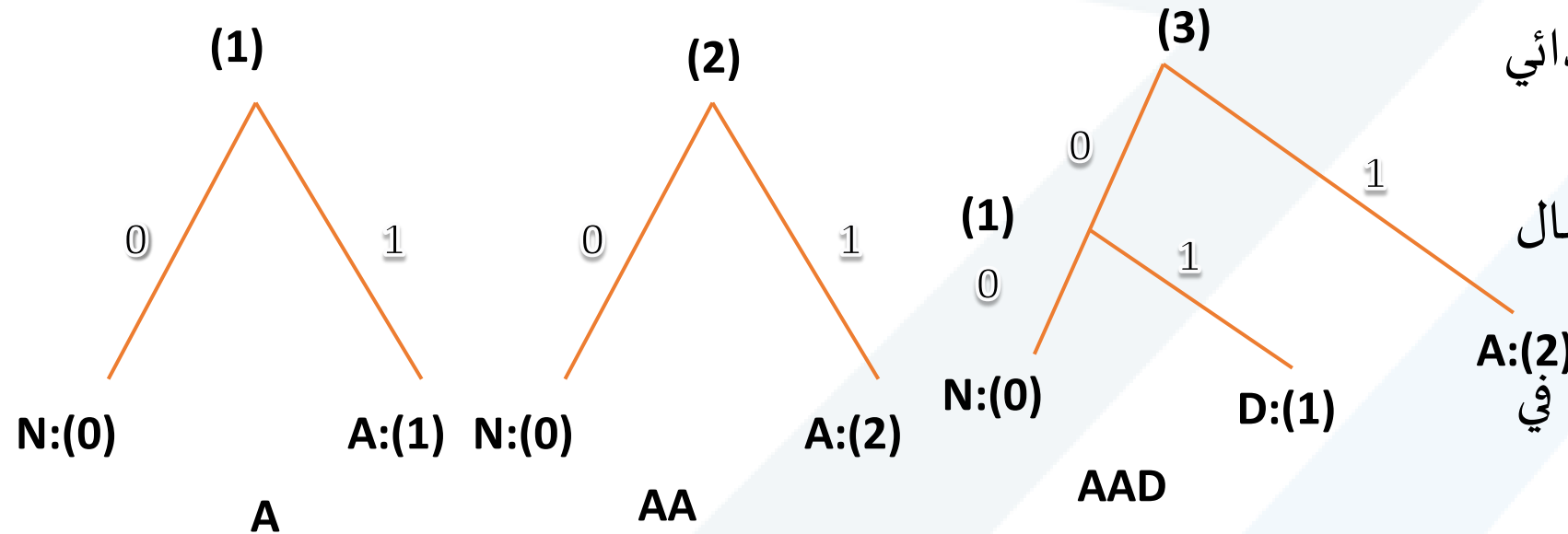
والمطلوب:

ترميز السلسلة وفق هوفمان الديناميكي وحساب نسبة الضغط

NEW:0  
A:00001  
C:00011  
D:00100

## مثال (١) (2/10)

الحل: إن فكرة الحل تتم وفق الخطوات الآتية: AADCCDD

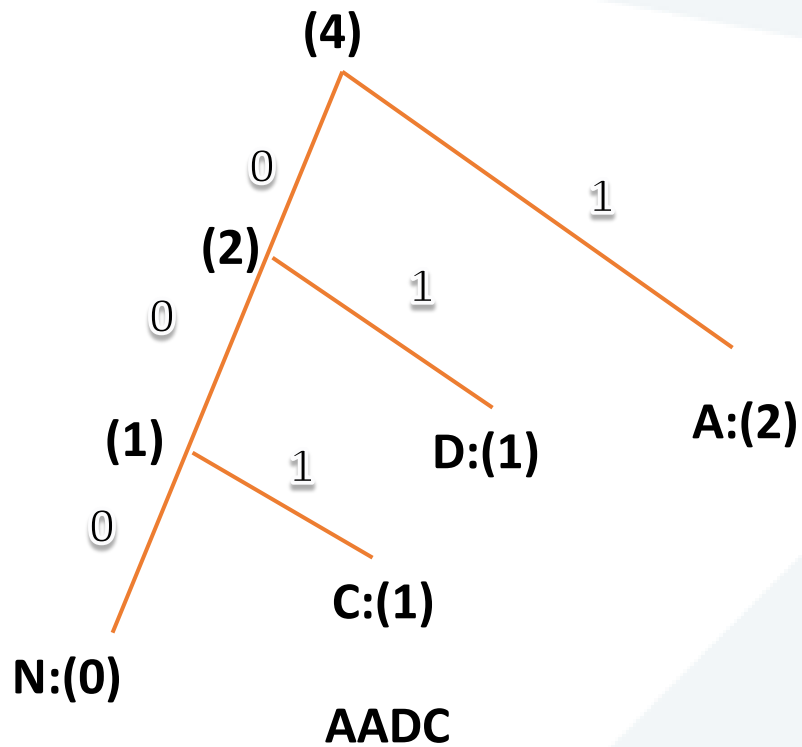


١. في البداية نرسل NEW والرميز البدائي A
٢. عند إرسال A للمرة الثانية يتم إرسال 1
٣. نرسل NEW بالرمز 0 لأن D ستأتي في المرة الأولى

Symb	N	A	A	N	D
code	0	00001	1	0	00100



NEW:0  
A:00001  
C:00011  
D:00100



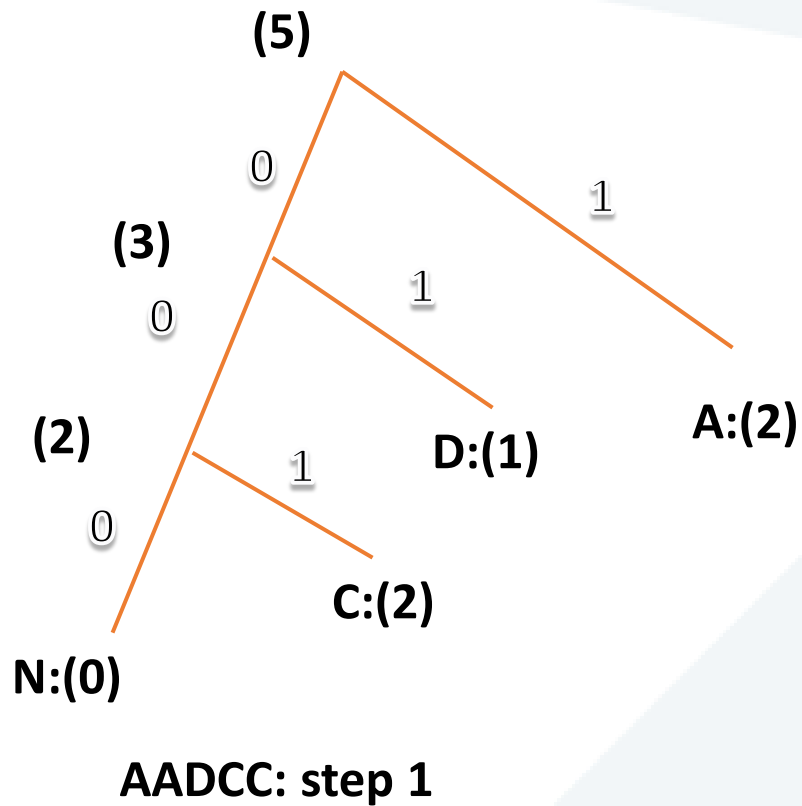
الحل: AADCCDD

٤. نرسل NEW بالرمز 0 لأن C ستأتي في المرة الأولى

٥. تأتي بعدها C أيضا نرسلها بالترميز 001

Symb	N	A	A	N	D	N	C	C
code	0	00001	1	0	00100	0	00011	001

## مثال (١) (4/10)



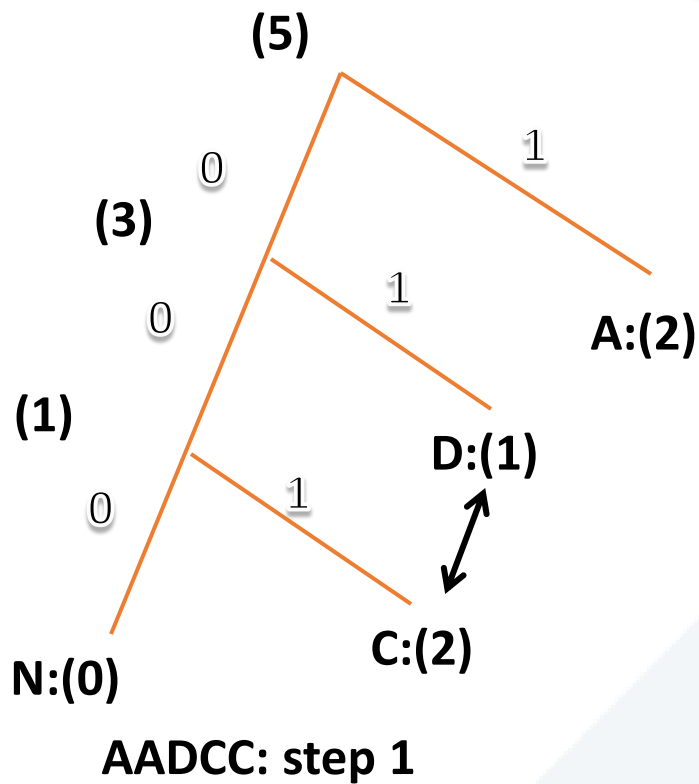
الحل: إن فكرة الحل تتم وفق الخطوات التالية: AADCCDD

٤. نرسل NEW بالرمز 0 لأن C ستأتي في المرة الأولى

٥. تأتي بعدها C أيضا نرسلها بالترميز 001

Symb	N	A	A	N	D	N	C	C
code	0	00001	1	0	00100	0	00011	001

## مثال (١) (5/10)



الحل: إن فكرة الحل تتم وفق الخطوات التالية: AADCCDD

ملاحظة هامة: في نفس الجهة يجب أن يكون دائماً وزن الفرع الأعلى أكبر من الأدنى وإلا يجب القيام بعملية التبديل بين الأعلى والأدنى

٦- نقوم بالتبديل بين موقع C و D وفق الشكل

وذلك لأن وزن الفرع C أكبر من وزن الفرع

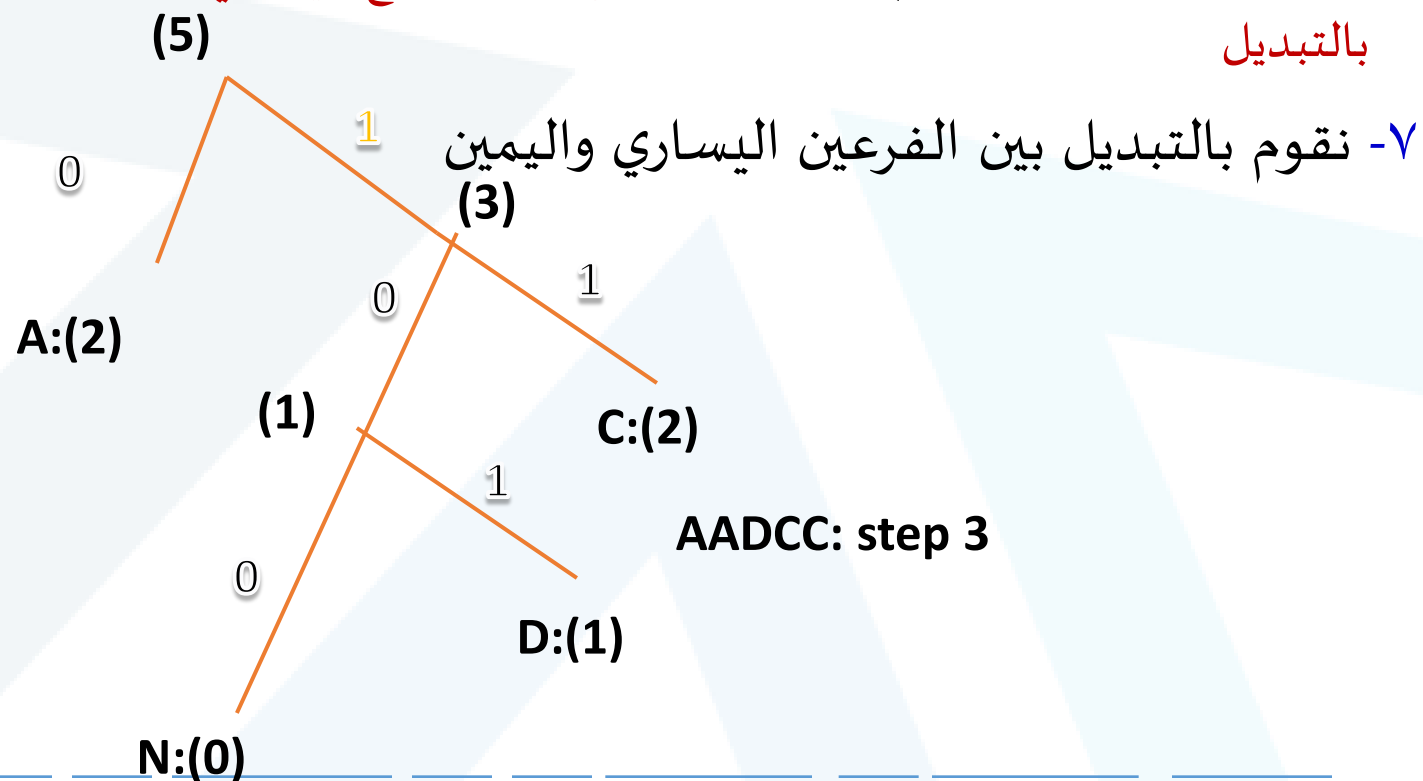
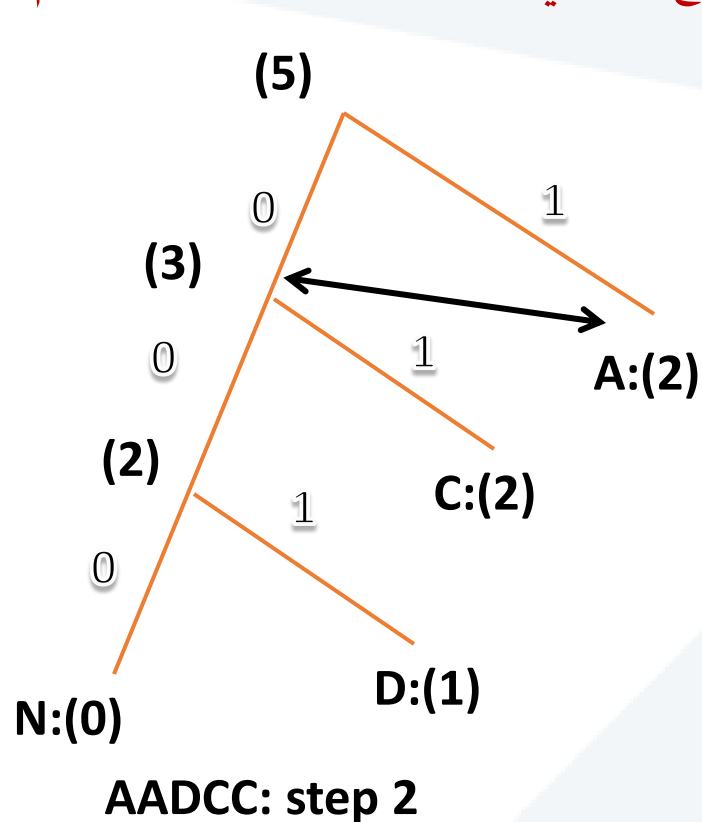
Symb	N	A	A	N	D	N	C	C
code	0	00001	1	0	00100	0	00011	001



جامعة  
المنصورة

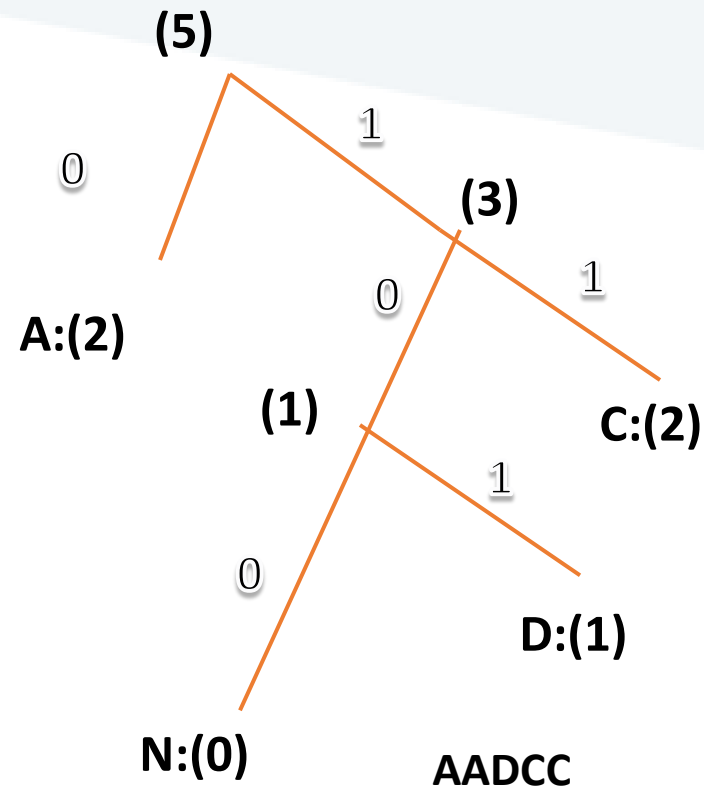
## مثال (١) (6/10)

ملاحظة هامة: من مقارنة الجهتين يجب أن يكون وزن الفرع اليساري دائما أصغر من وزن الفرع اليميني المقابل له. وإلا يجب القيام بالتبديل

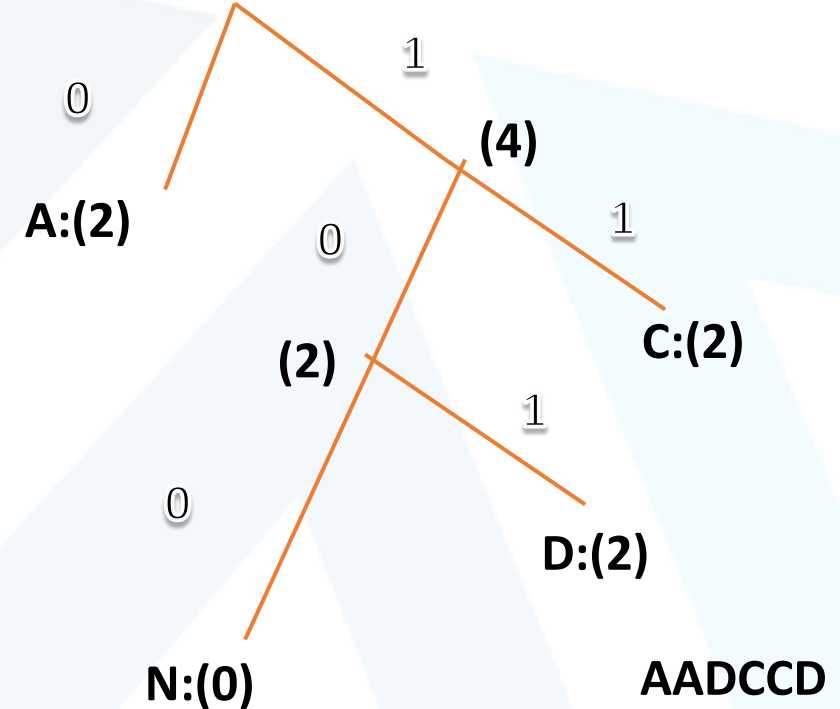


Symb	N	A	A	N	D	N	C	C
code	0	00001	1	0	00100	0	00011	001

## مثال (١) (7/10)

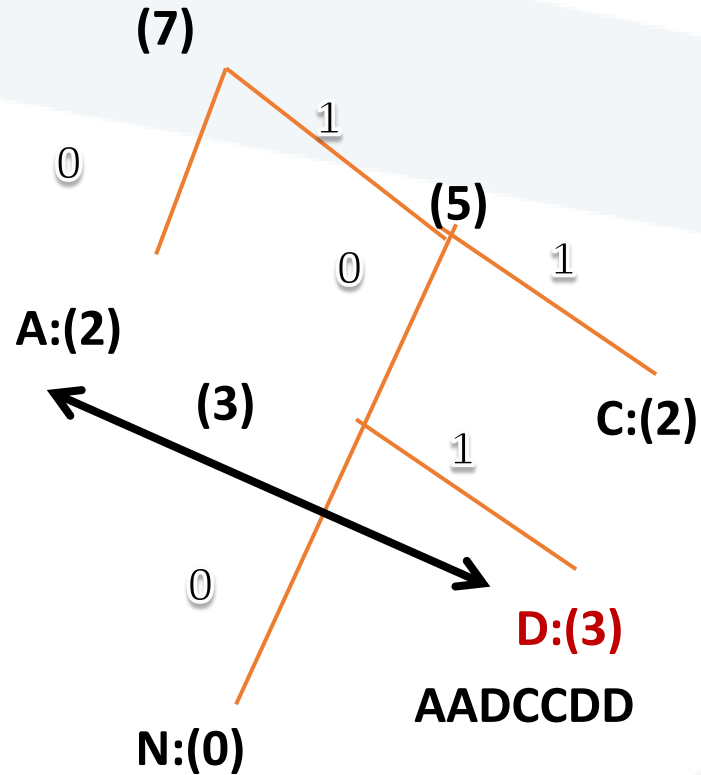


٨- نرسل D بالرمز 101 مرتين على التوالي (6)



Symb	N	A	A	N	D	N	C	C	D
code	0	00001	1	0	00100	0	00011	001	101

## مثال (١) (8/10)



٨- نرسل D بالرمز 101 مرتين على التوالي

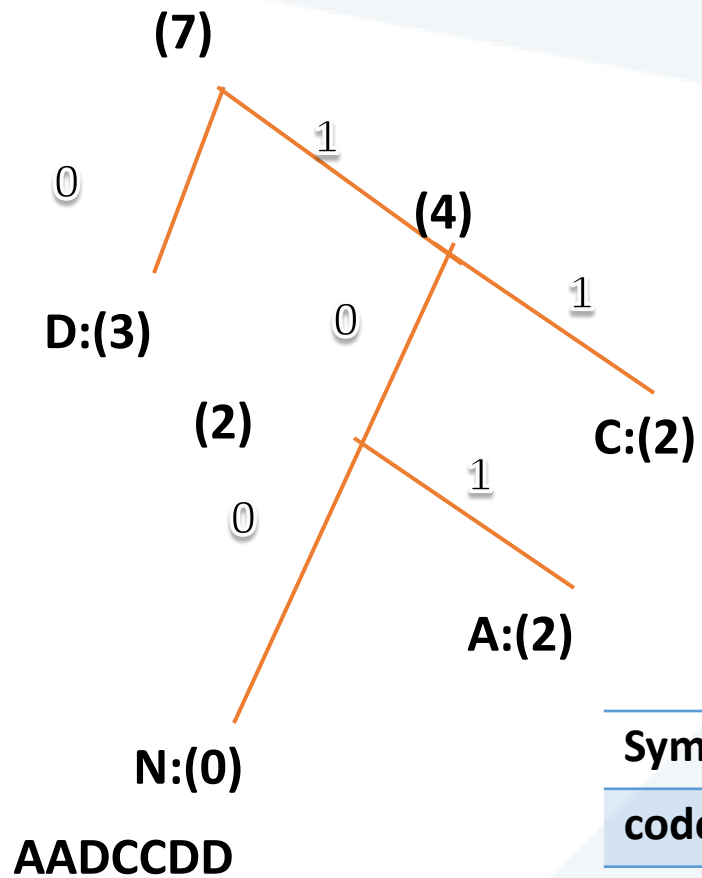
ملاحظة: يجب أن نوازن الشجرة الناتجة

ملاحظة: في حال وجد فرع يساري، يجب أن يكون وزن أي الفرع في جهة اليسار أعلى من وزن أدنى فرع في جهة اليمين وإلا يجب التبديل بينهما

Symb	N	A	A	N	D	N	C	C	D	D
code	0	00001	1	0	00100	0	00011	001	101	101

## مثال (١) (9/10)

٩- إذا لموازنة الشجرة نقوم بالتبديل بين موقع D و A



Symb	N	A	A	N	D	N	C	C	D	D
code	0	00001	1	0	00100	0	00011	001	101	101

## مثال (١) (10/10)

الحل: ٢. حساب نسبة الضغط

Symb	N	A	A	N	D	N	C	C	D	D
code	0	00001	1	0	00100	0	00011	001	101	101

✓ عدد البتات في ترميز هوفمان :

$$1+5+1+1+5+1+5+3+3+3=28 \text{ bits}$$

✓ بينما يكون لدينا عدد البتات المستخدمة في ترميز الآسكي:  $7*7=49 \text{ bits}$

✓ فتكون نسبة الضغط:  $28*100/49=57.14\%$



## نهاية المحاضرة الرابعة