

Département d'informatique et de génie logiciel
Compression de données
IFT-4003/IFT-7023

Notes de cours
Codage adaptatif de Huffman

Édition Hiver 2012

Mohamed Haj Taieb

Local: PLT 2113

Courriel: mohamed.haj-taieb.1@ulaval.ca

Faculté des sciences et de génie
Département de génie électrique et de
génie informatique



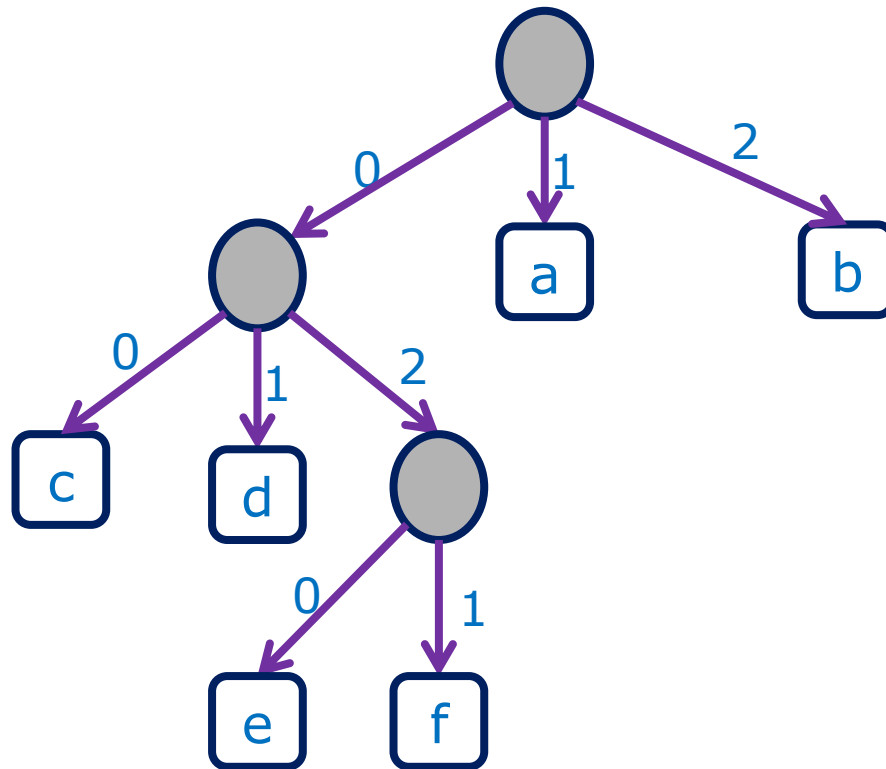
Plan de la présentation

□ Codage adaptatif de Huffman:

- Idée principale
- Arbre de Huffman
- Mise à jour de l'arbre
- Encodage adaptatif de Huffman
- Décodage adaptatif de Huffman

Construction d'un Code de Huffman ternaire [Arbre]

Lettres	a	b	c	d	e	f
Probabilité	0.25	0.20	0.20	0.20	0.10	0.05
Code	1	2	00	01	020	021



Idée de base du code adaptatif de Huffman (1)

❑ Problématique

- L'encodage de Huffman nécessite une connaissance à priori de la probabilité d'apparition des symboles.
- Il faut alors effectuer une étude statistique des données pour générer un code.
- Ensuite effectuer l'encodage de Huffman.
- Procédure d'encodage se fait en deux étapes.
- Ce qui n'est pas toujours pratique.

❑ Solution triviale

- Commencer par l'hypothèse d'équiprobabilité.
- Pour le $k^{\text{ème}}$ symbole construire un nouveau code de Huffman basé sur les symboles précédents.
- Complexité calculatoire élevée

Idée de base du code adaptatif de Huffman (2)

- Solution plus élaborée: Soit l'alphabet $A = \{a_1, \dots, a_n\}$
 1. Fixer un code initial pour l'alphabet.
 2. Arbre formée d'une seule feuille **NYT** (*not yet transmitted*).
 3. Parcourir les éléments de la source.
 - A. Si l'élément NYT
 - i. Envoyer NYT
 - ii. Mise à jour de l'arbre
 - B. Sinon
 - i. Envoyer le mot code de l'élément
 - ii. Mise à jour de l'arbre
 4. Condition d'arrêt: parcours de toute la source.
- Note 1: Code dépend de la fréquence des éléments.
- Note 2: Encodeur et décodeur doivent être synchrones.

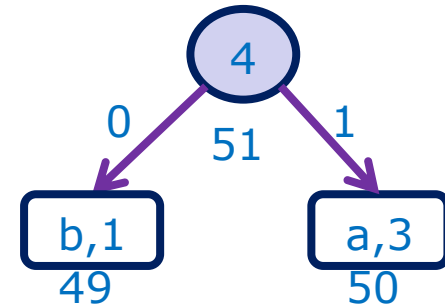
Arbre de Huffman

□ Alphabet de n éléments

- Un arbre contient $2n-1$ nœuds (racines+feuilles)
- Les attributs d'un nœuds sont: symboles, gauche=0, droite=1, parent, frère(a et b), feuille F, poids=4, identificateur id=19, symbole a.

□ Propriété des poids

- Si x feuille: poids=fréquence d'occurrence
- Si x racine: poids=poids[Frère(droite)]+poids[Frère(gauche)]



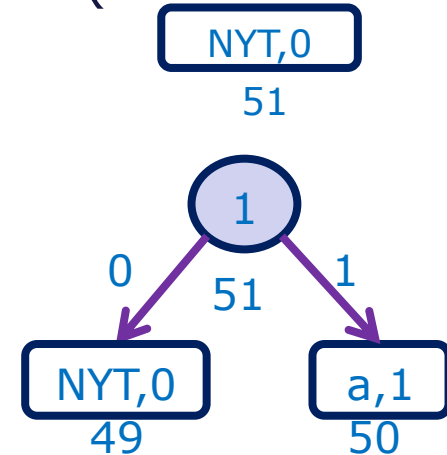
□ Propriété de l'identificateur

- Si $\text{poids}(b) \leq \text{poids}(a) \Rightarrow \text{id}(b) \leq \text{id}(a)$
- Propriété de fratrie: $\text{Id}(\text{frère gauche}) = 2k-1$ et $\text{Id}(\text{frère droite}) = 2k$

Mise à jour de l'arbre de Huffman

□ Algorithme

- Point de départ: arbre avec un seul élément (feuille sans racine): $id=2n-1$, $poids=0$, $symbole=NYT$.
- Nouveau symbole: on éclate le nœud NYT
- Préserver la propriété de fratrie.
 - Quant un nœud x est mis à jour
 - Répéter
 - Si $poids(x) < poids(frère(x))$ [cas binaire: 1 seul frère]
 - $poids(x)++$, exit.
 - Sinon
 - Inter-changer les frères, $poids(x)++$,
 - $x=parent(x)$, réitérer.
 - Jusqu'à: $x=racine$



Exemple encodage adaptatif de Huffman

□ Énoncé de l'exemple

- Alphabet: lettres minuscules de la langue anglaise. La longueur de l'alphabet=26.
- Message=[a a r d v a k]
- Code initial:

a	00000	i	01000	o	01110	u	10100
b	00001	j	01001	p	01111	v	10101
c	00010	k	01010	q	10000	w	10110
d	00011	l	01011	r	10001	X	10111

- Nombre total des nœuds dans l'arbre: $2 \times 26 - 1 = 51$.
- Identificateur de la racine de l'arbre est 51.

Encodage adaptatif de Huffman (1)

- Message: a a r d v a k
- Code:

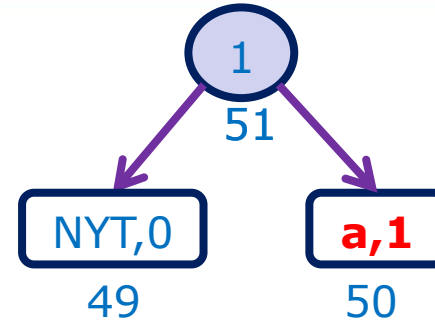
NYT,0
51

Symbole	Code
NYT	
a	
d	
k	
r	
v	
a	00000
d	00011
k	01010
r	10001
v	10101

Encodage adaptatif de Huffman (2)

- Message: **a** a r d v a k
- Code: 00000

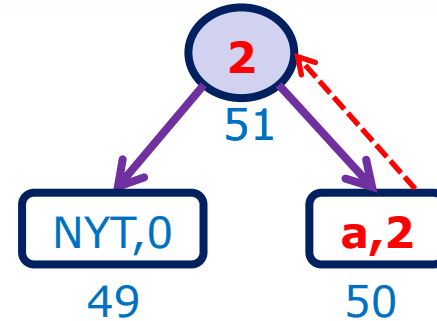
Symbole	Code
NYT	0
a	1
d	
k	
r	
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (3)

- Message: a **a** r d v a k
- Code: 00000**1**

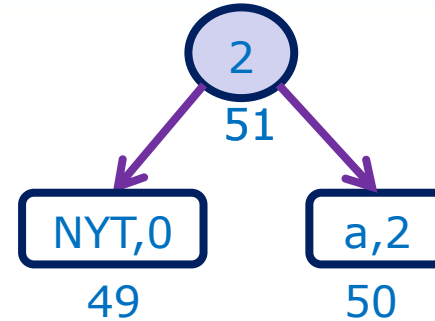
Symbole	Code
NYT	0
a	1
d	
k	
r	
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (4)

- Message: a a **r** d v a k
- Code: 000001010001

Symbole	Code
NYT	0
a	1
d	
k	
r	
v	
a	00000
d	00011
k	01010
r	10001
v	10101

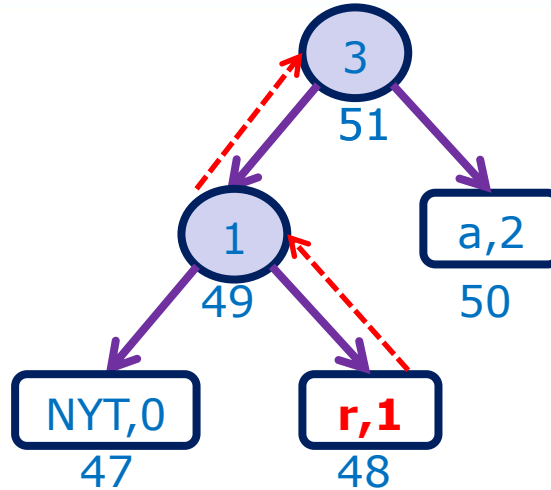


Encodage adaptatif de Huffman (5)

- Message: a a **r** d v a k
- Code: 000001010001

Symbole	Code
NYT	00
a	1
d	
k	
r	01
v	

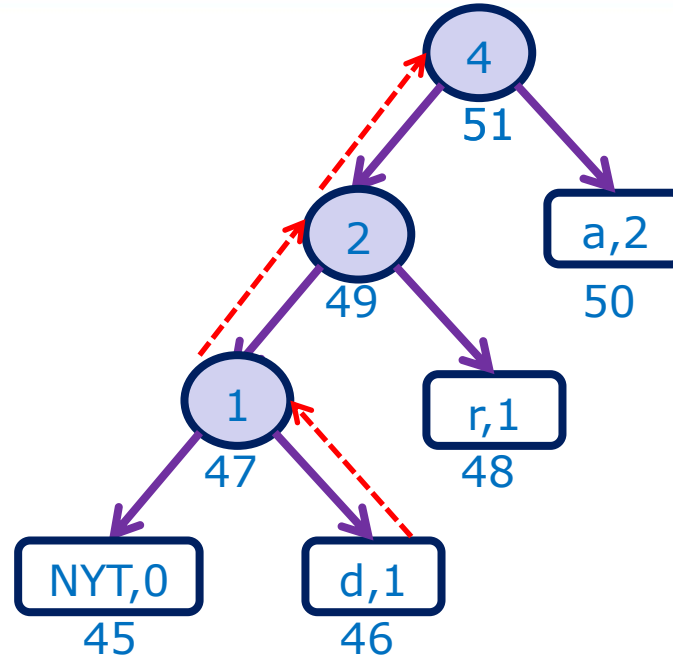
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (6)

- Message: a a r **d** v a k
- Code: 0000010100010000011

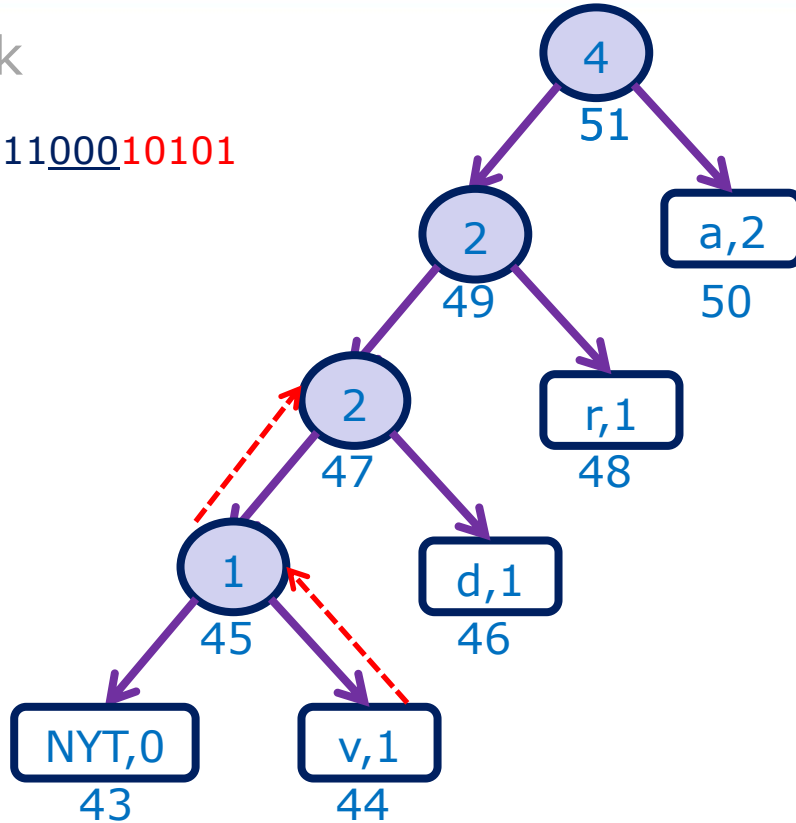
Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (7)

- Message: a a r d **v** a k
- Code: 0000010100010000011000**10101**

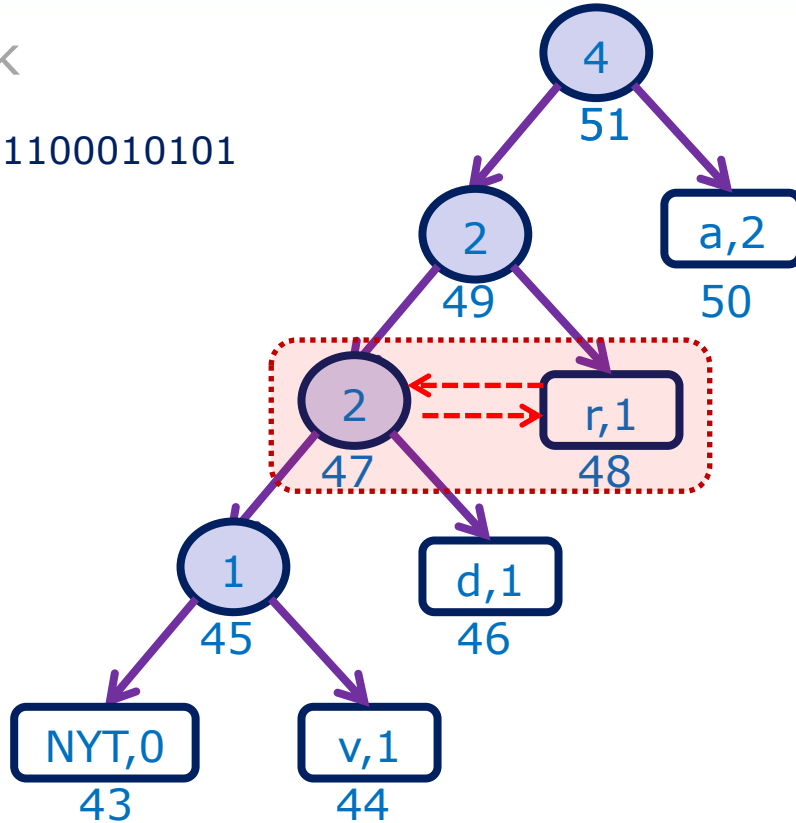
Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (8)

- Message: a a r d v a k
- Code: 000001010001000001100010101

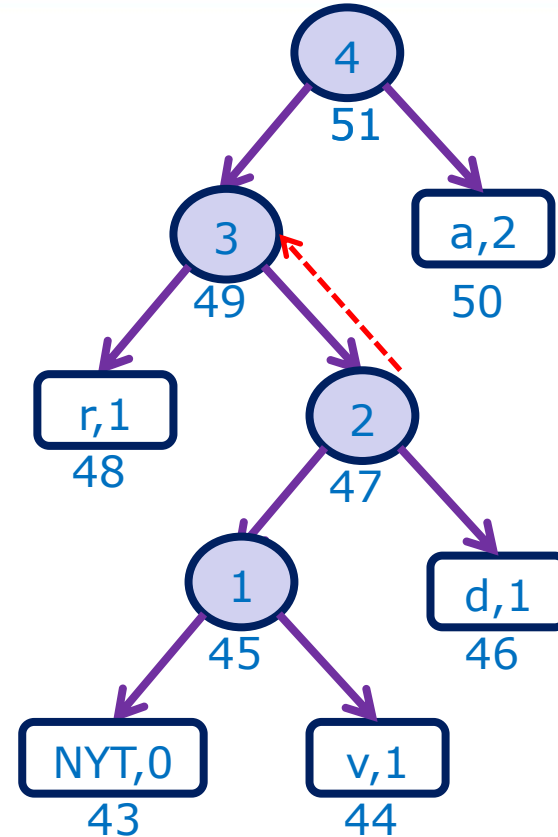
Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (9)

- Message: a a r d v a k
- Code: 000001010001000001100010101

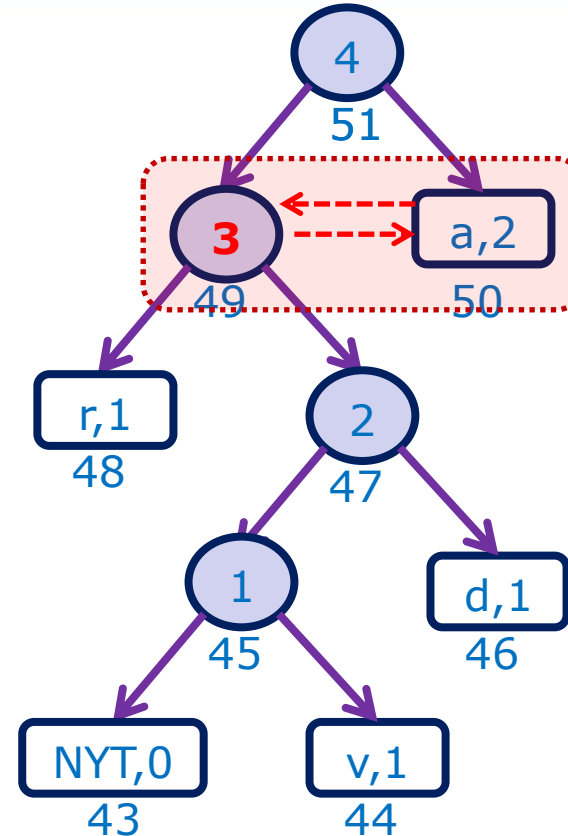
Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (10)

- Message: a a r d v a k
- Code: 000001010001000001100010101

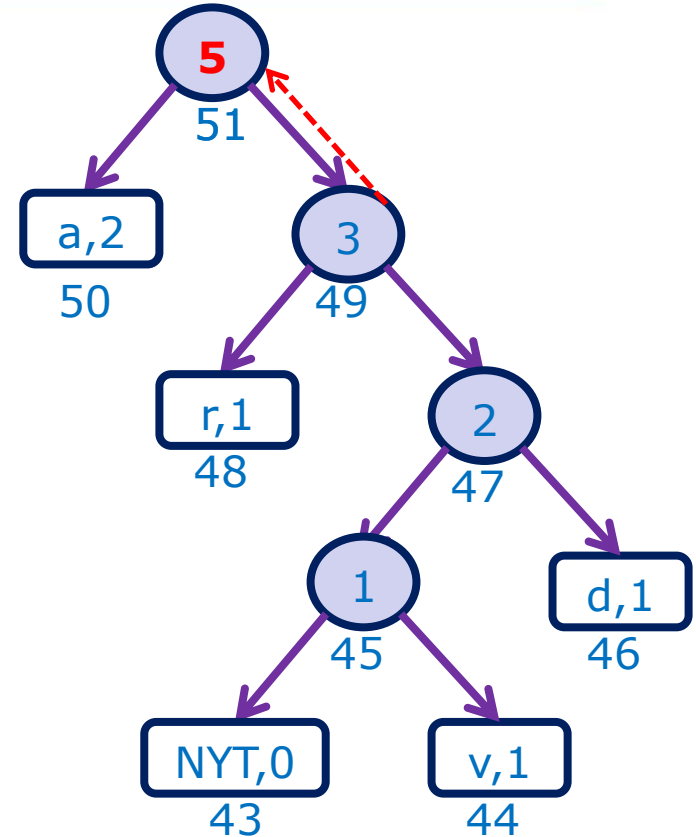
Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (11)

- Message: a a r d v a k
- Code: 000001010001000001100010101

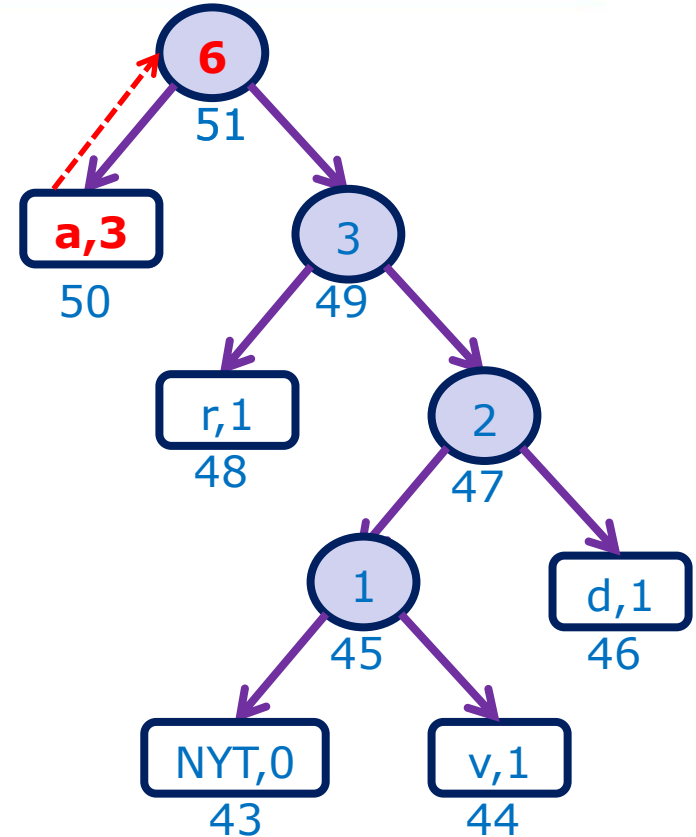
Symbole	Code
NYT	1100
a	0
d	111
k	
r	10
v	1101
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (12)

- Message: a a r d v **a** k
- Code: 0000010100010000011000101010

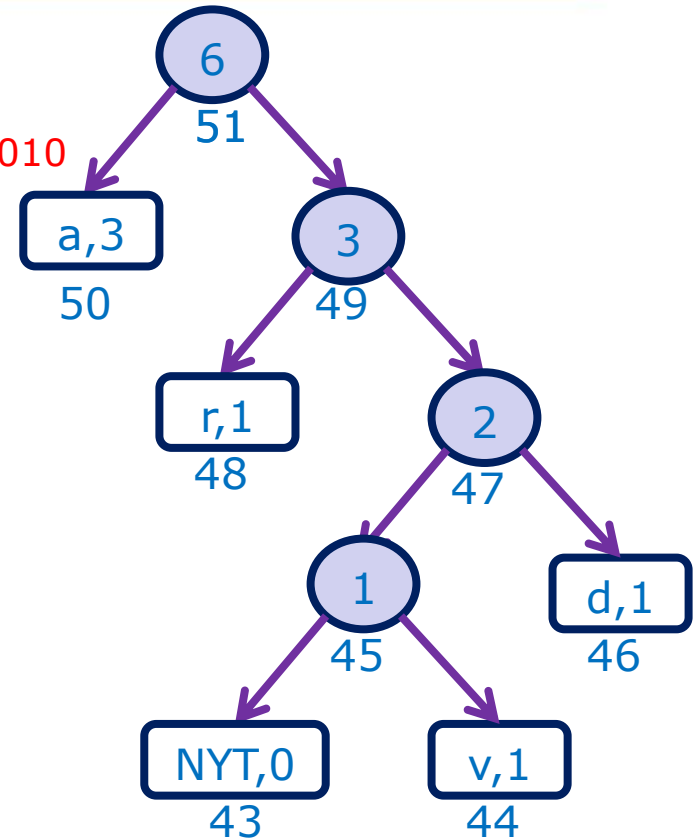
Symbole	Code
NYT	1100
a	0
d	111
k	
r	10
v	1101
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (12)

- Message: a a r d v a **k**
- Code: 000001010001000001100101010110001010

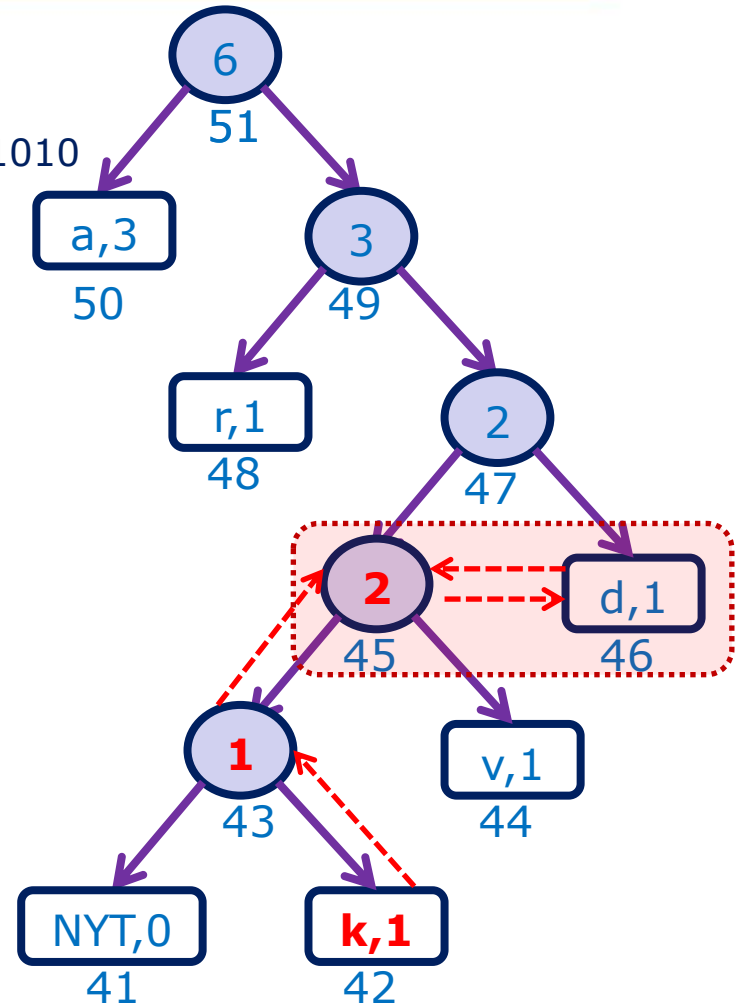
Symbole	Code
NYT	1100
a	0
d	111
k	
r	10
v	1101
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (13)

- Message: a a r d v a k
- Code: 000001010001000001100101010110001010

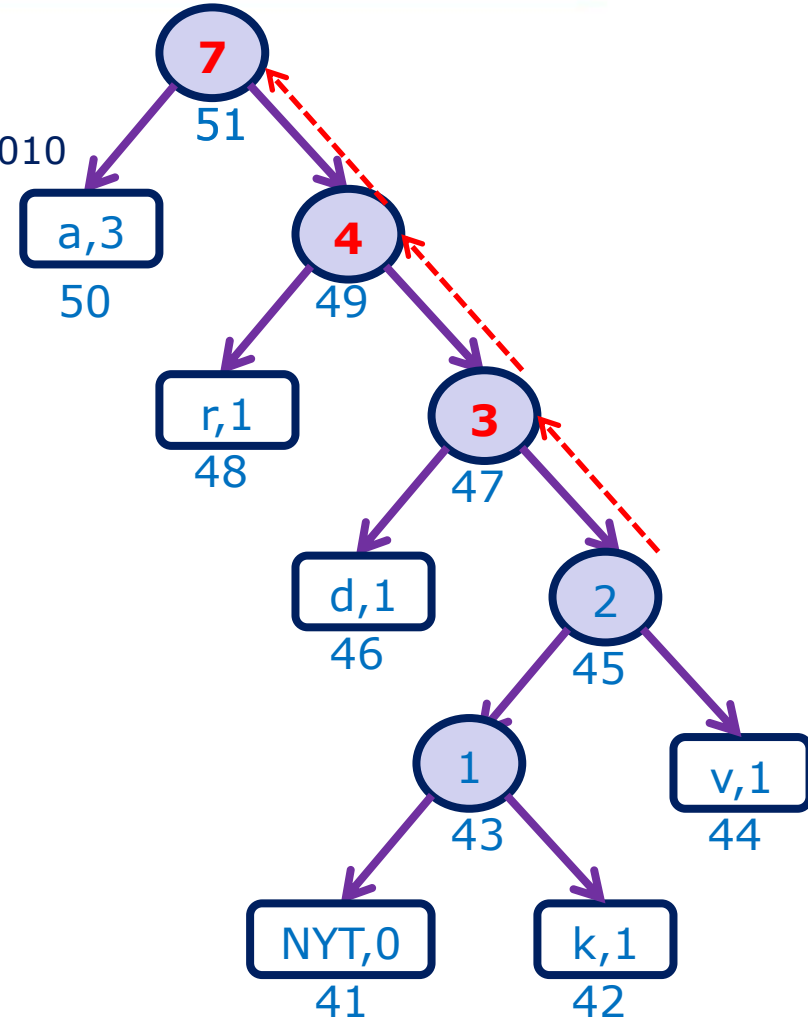
Symbole	Code
NYT	1100
a	0
d	111
k	
r	10
v	1101
a	00000
d	00011
k	01010
r	10001
v	10101



Encodage adaptatif de Huffman (14)

- Message: a a r d v a k
- Code: 000001010001000001100101010110001010

Symbole	Code
NYT	11100
a	0
d	110
k	11101
r	10
v	1111
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (1)

- Message décodé:
- Code: 000010100010000011000101010110001010

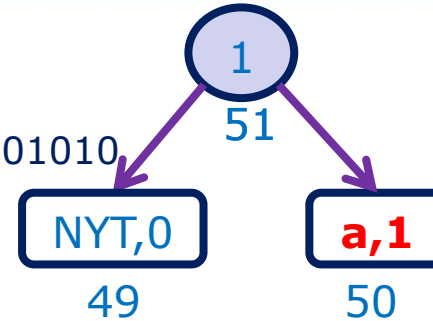
NYT,0
51

Symbole	Code
NYT	
a	
d	
k	
r	
v	
a	00000
d	00011
k	01010
r	10001
v	10101

Décodage adaptatif de Huffman (2)

- Message décodé: **a**
- Code: **0000**10100010000011000101010110001010

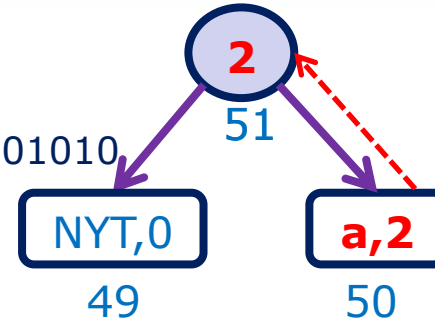
Symbole	Code
NYT	0
a	1
d	
k	
r	
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (3)

- Message décodé: a **a**
- Code: 00000**1**0100010000011000101010110001010

Symbole	Code
NYT	0
a	1
d	
k	
r	
v	
a	00000
d	00011
k	01010
r	10001
v	10101

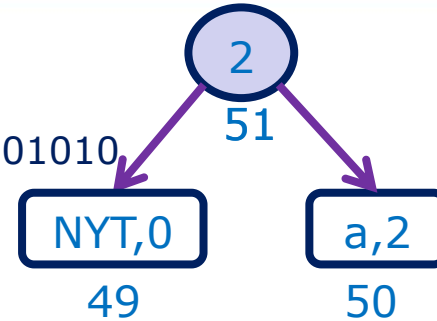


Décodage adaptatif de Huffman (4)

- Message décodé: a a
- Code: 0000010100010000011000101010110001010

Symbole	Code
NYT	0
a	1
d	
k	
r	
v	

a	00000
d	00011
k	01010
r	10001
v	10101

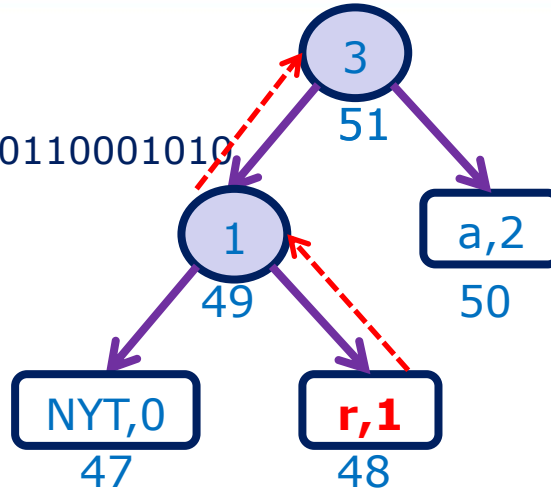


Décodage adaptatif de Huffman (5)

- Message décodé: a a **r**
- Code: 0000010**10001**0000011000101010110001010

Symbole	Code
NYT	0
a	1
d	
k	
r	
v	

a	00000
d	00011
k	01010
r	10001
v	10101

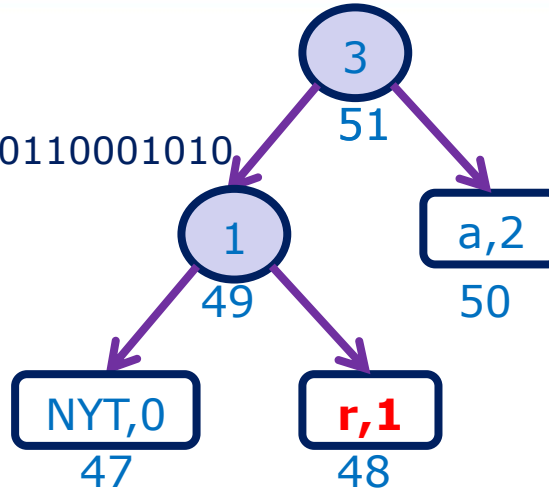


Décodage adaptatif de Huffman (6)

- Message décodé: a a **r**
- Code: 0000010100010000011000101010110001010

Symbole	Code
NYT	00
a	1
d	
k	
r	01
v	

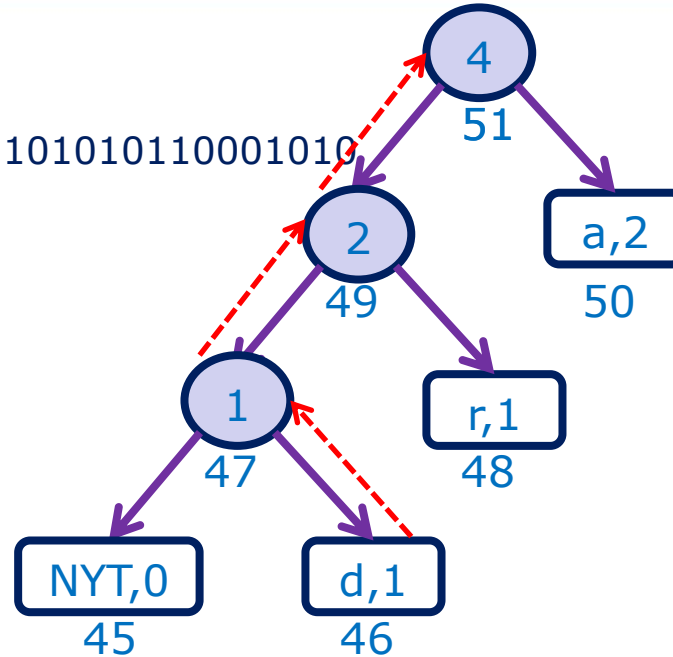
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (7)

- Message décodé: a a r **d**
- Code: 0000010100010000011000101010110001010

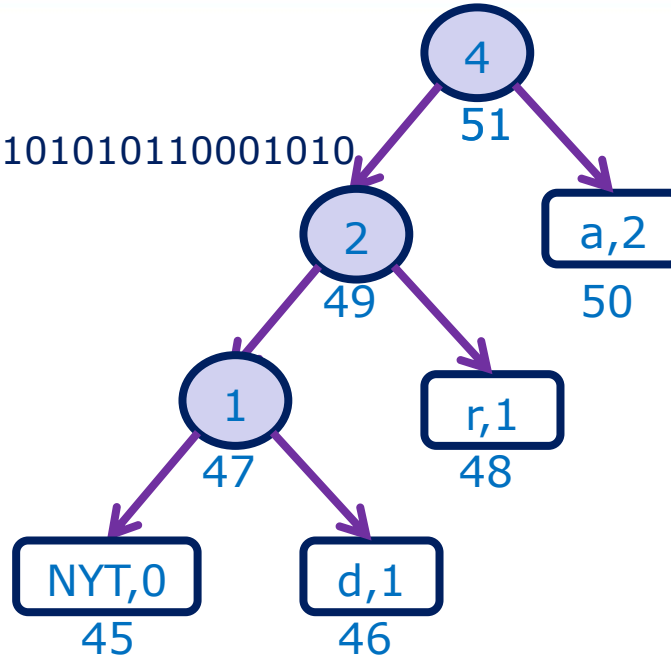
Symbole	Code
NYT	00
a	1
d	
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (7)

- Message décodé: a a r **d**
- Code: 0000010100010000011000101010110001010

Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101

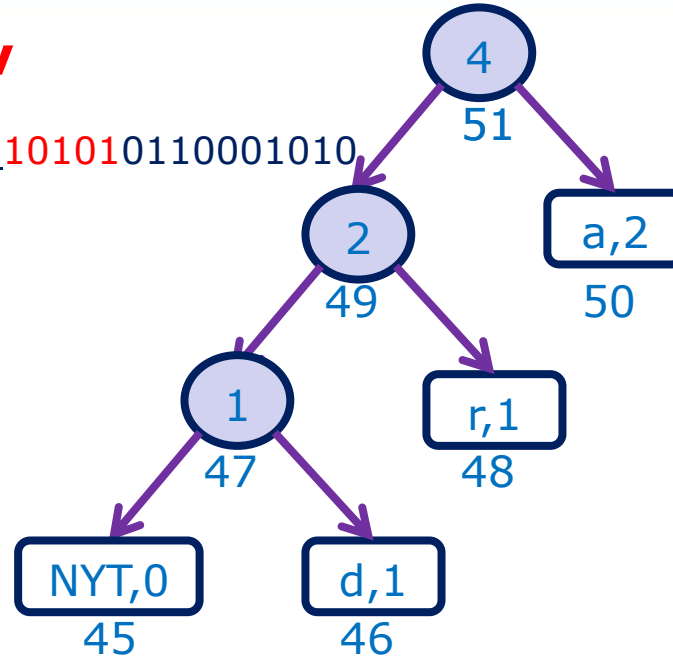


Décodage adaptatif de Huffman (8)

- Message décodé: a a r d **v**
- Code: 0000010100010000011000101010110001010

Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	

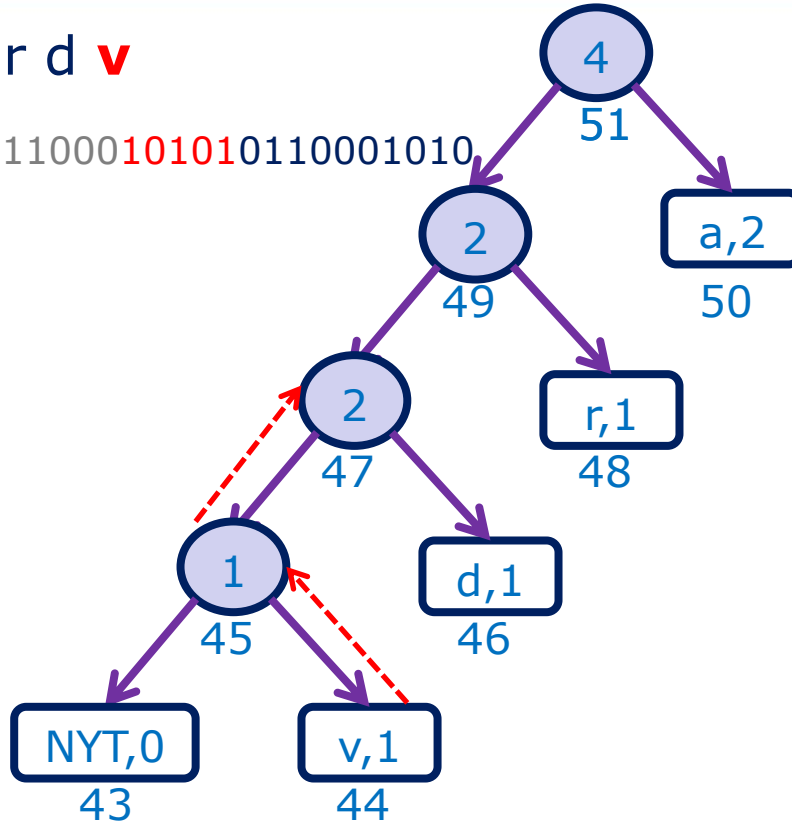
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (9)

- Message décodé: a a r d **v**
- Code: 0000010100010000011000101010110001010

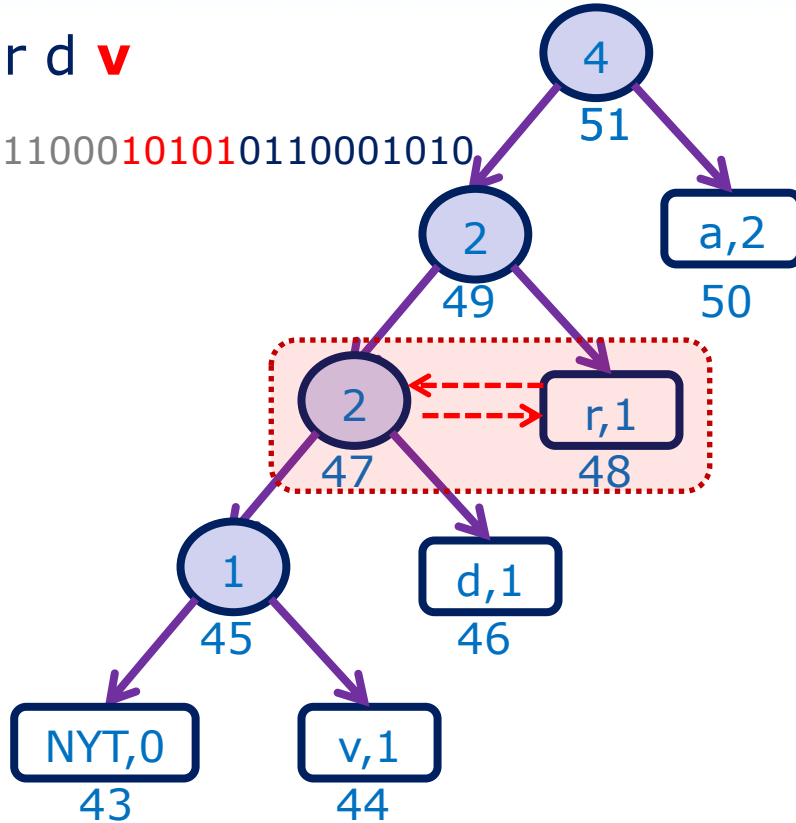
Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (10)

- Message décodé: a a r d **v**
- Code: 0000010100010000011000101010110001010

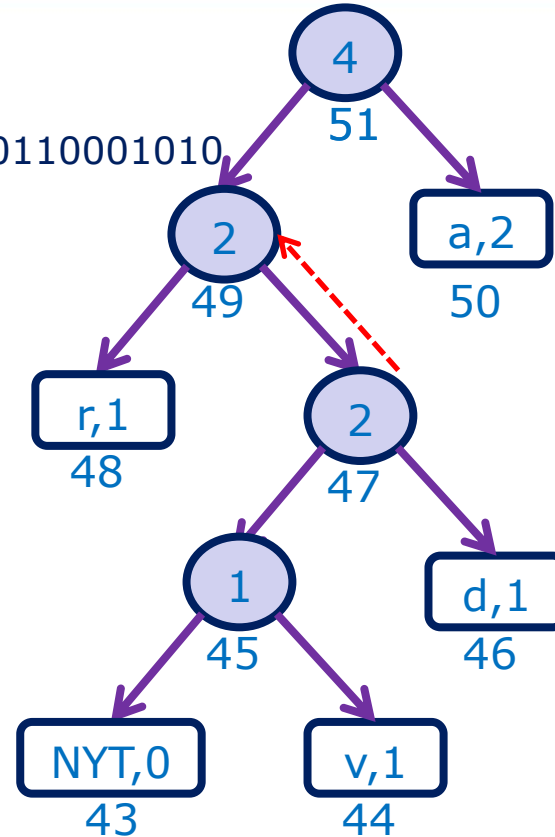
Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (11)

- Message décodé: a a r d v
- Code: 0000010100010000011000101010110001010

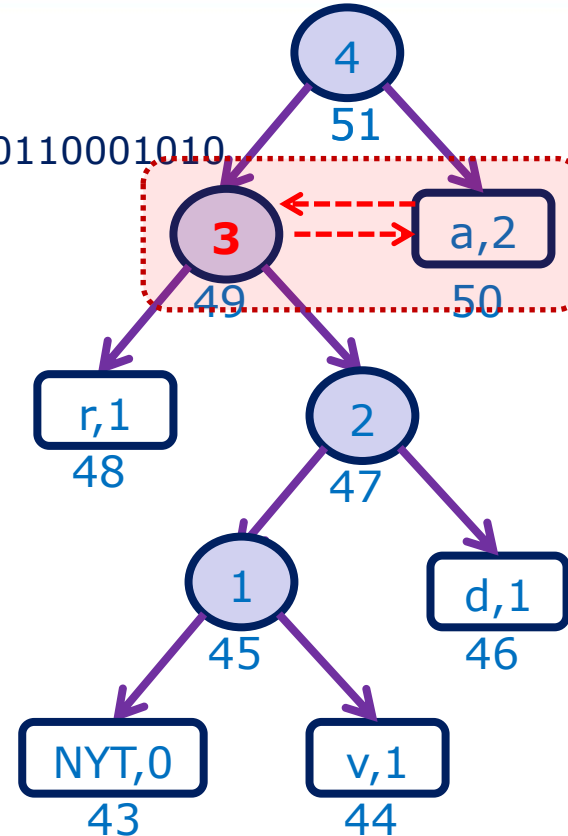
Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (12)

- Message décodé: a a r d **v**
- Code: 0000010100010000011000101010110001010

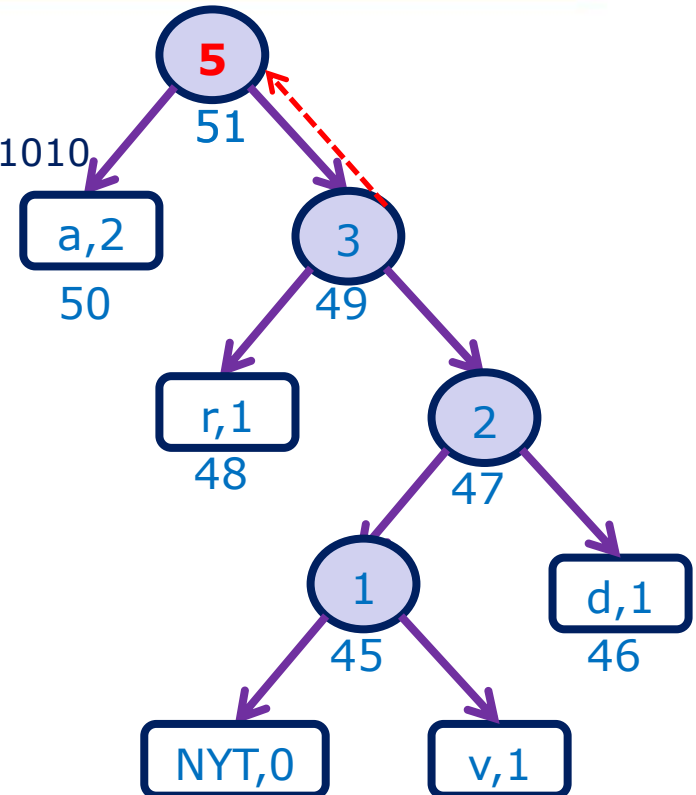
Symbole	Code
NYT	000
a	1
d	001
k	
r	01
v	
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (13)

- Message décodé: a a r d **v**
- Code: 0000010100010000011000101010110001010

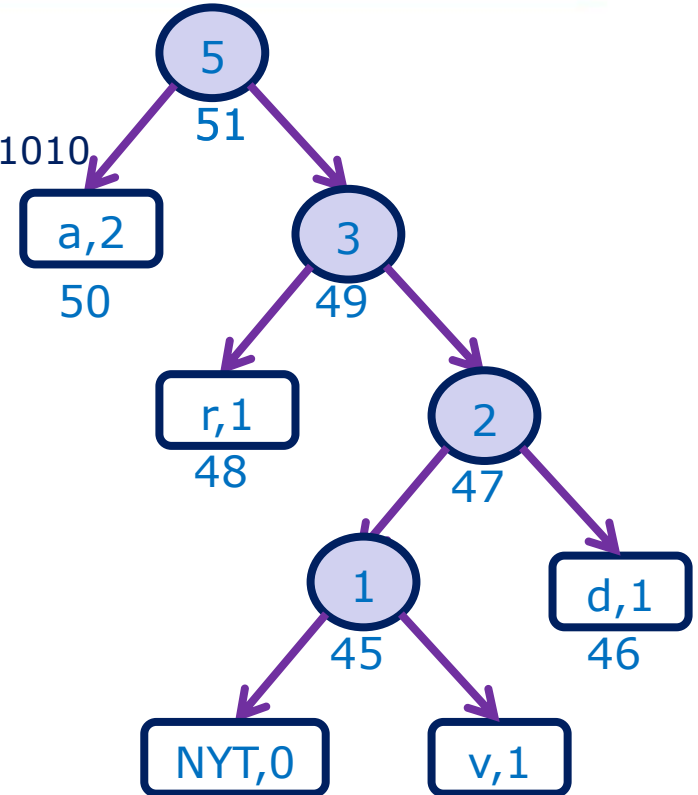
Symbole	Code
NYT	1100
a	0
d	111
k	
r	10
v	1101
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (14)

- Message décodé: a a r d v **a**
- Code: 000001010001000001100010101**0**110001010

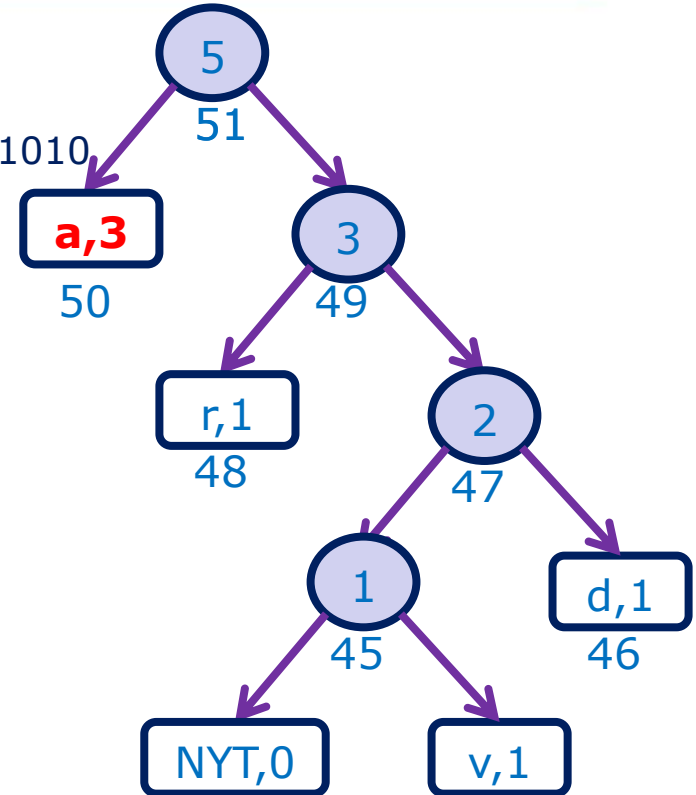
Symbole	Code
NYT	1100
a	0
d	111
k	
r	10
v	1101
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (15)

- Message décodé: a a r d v a
- Code: 0000010100010000011000101010110001010

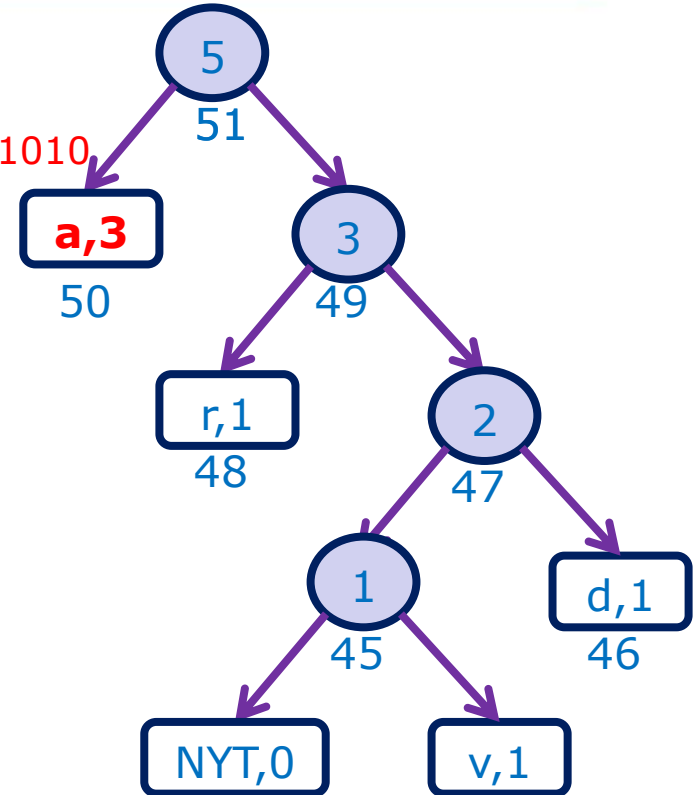
Symbole	Code
NYT	1100
a	0
d	111
k	
r	10
v	1101
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (16)

- Message décodé: a a r d v a **k**
- Code: 0000010100010000011000101010110001010

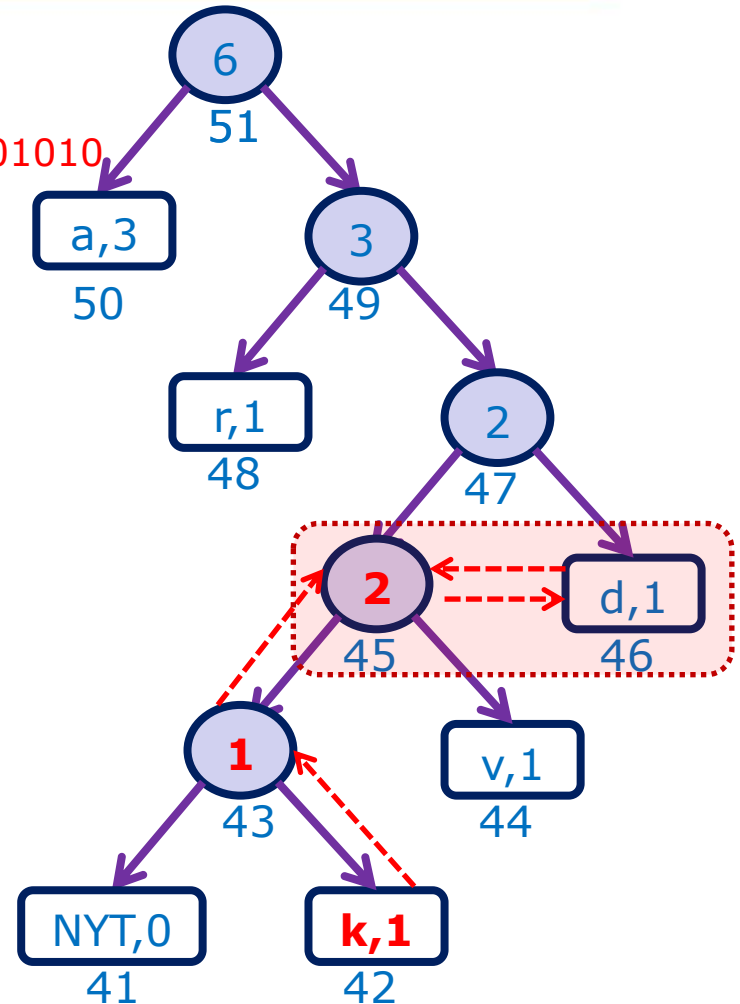
Symbole	Code
NYT	1100
a	0
d	111
k	
r	10
v	1101
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (17)

- Message décodé: a a r d v a **k**
- Code: 0000010100010000011000101010110001010

Symbole	Code
NYT	1100
a	0
d	111
k	
r	10
v	1101
a	00000
d	00011
k	01010
r	10001
v	10101



Décodage adaptatif de Huffman (18)

- Message décodé: a a r d v a **k**
- Code: 0000010100010000011000101010110001010

Symbole	Code
NYT	11100
a	0
d	110
k	11101
r	10
v	1111
a	00000
d	00011
k	01010
r	10001
v	10101

