

Fichiers en B-Arbres

C'est l'une des méthodes d'accès les plus efficaces connues à ce jour pour les fichiers volumineux et hautement dynamiques

Fichiers de type arbres m-aires de recherche complètement équilibrés

- Mêmes structures de blocs et mêmes déclarations
- Mêmes mécanismes de recherche (d'une valeur et par intervalle de valeurs)

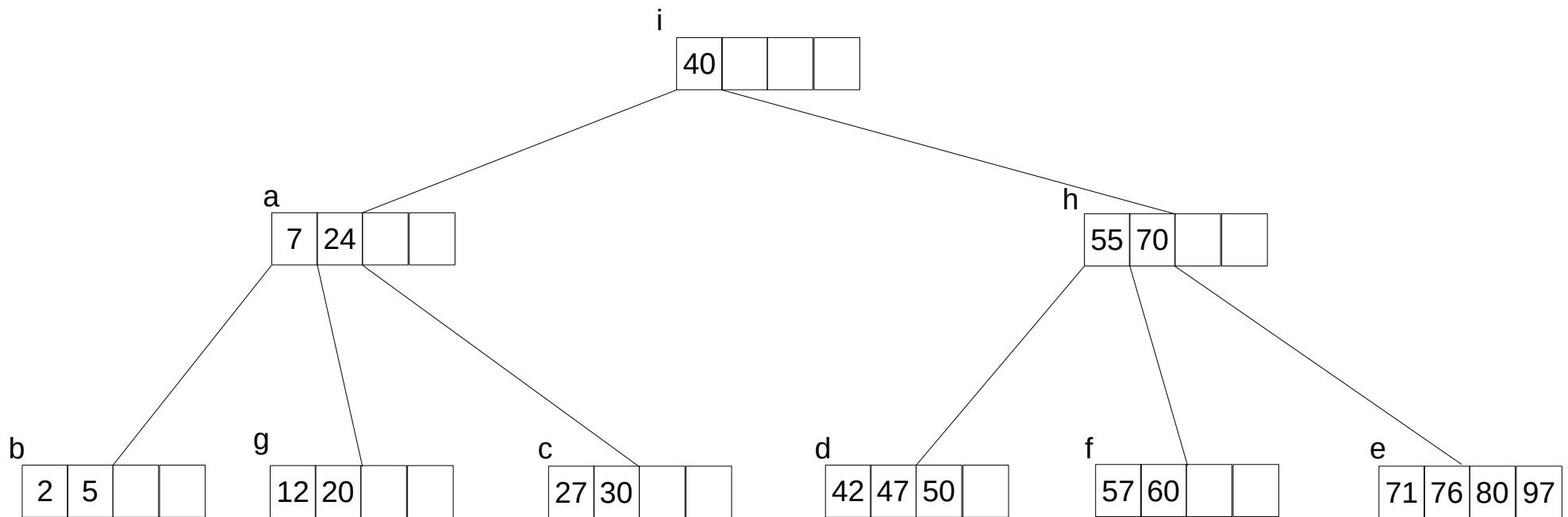
Propriétés

Un B-arbre d'ordre **N** est un arbre m-aire de recherche d'ordre **N** vérifiant les propriétés suivantes :

- Tous les nœuds (autres que la racine) doivent être remplis au minimum à 50 % de leur capacité ($\text{degré minimal} = \lceil N/2 \rceil$)
- Le nœud racine peut contenir au minimum une seule valeur et deux fils
- Toutes les feuilles sont dans le même niveau
- Dans un nœud interne, tous les fils ($Fils_1, Fils_2, \dots, Fils_{\text{degré}}$) sont différents de -1 et dans un nœud feuille, tous les fils ($Fils_1, Fils_2, \dots, Fils_{\text{degré}}$) sont à -1

Exemple de B-arbre d'ordre 5

- La capacité maximal d'un nœud est donc : 4 valeurs (et 5 fils)
- A part la racine, la capacité minimal d'un nœud est : 2 valeurs (donc 3 fils)



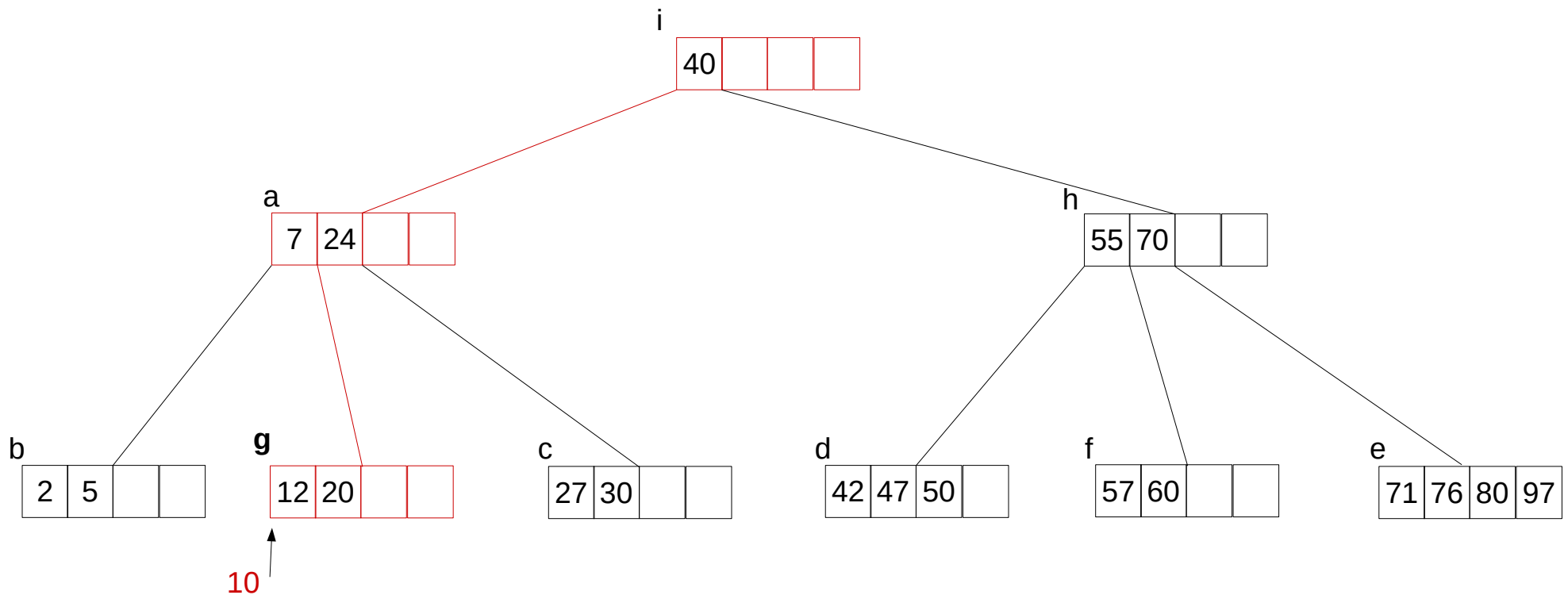
Mécanisme d'insertion

Insertion d'une nouvelle valeur x :

1. Rechercher x ,
Si x existe déjà, Aller à 4. (Fin)
Sinon, Soit P le dernier nœud visité (c'est une feuille) et Aller à 2.
2. Si (P n'est pas plein)
Insérer x dans P (par décalages internes) et Aller à 4. (Fin)
3. // Si P est plein, il faut éclater P en deux nœuds :
Allouer un nouveau nœud $\Rightarrow Q$
 P : contiendra la 1^{ère} moitié,
 Q : contiendra la dernière moitié
Soit m la valeur du milieu séparant les 2 moitiés
// il faut maintenant insérer m dans le père de P ...
 $x \leftarrow m$; $P \leftarrow \text{pere}(P)$; Si ($P == \text{nil}$), $P \leftarrow \underline{\text{AllouerNouvelleRacine}}$
Aller à 2.
4. // Fin

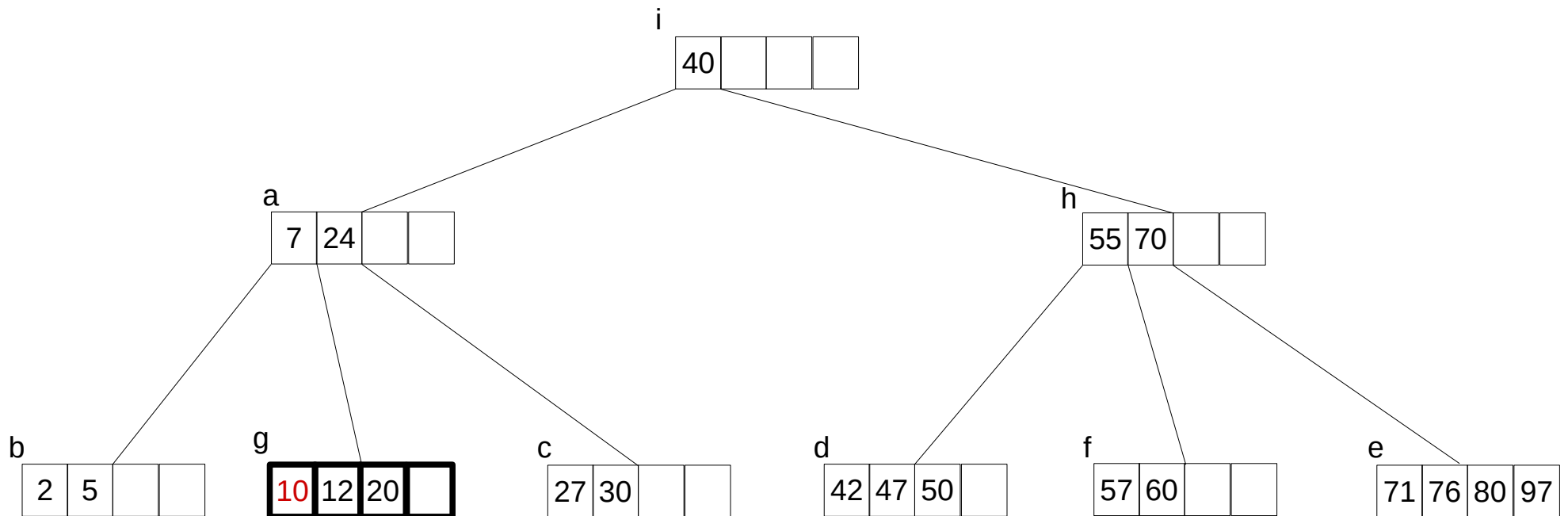
Insertion de la valeur 10

- La recherche de 10 mène vers le nœud feuille **g** (pos = **1** dans **g**)
- Comme **g** n'est pas plein, l'insertion se fait par décalages internes



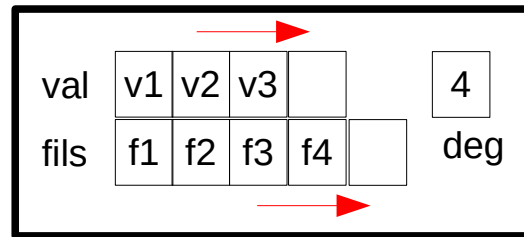
Résultat de l'insertion de la valeur 10

- La recherche de 10 mène vers le nœud feuille **g** (pos = 1 dans **g**)
- Comme **g** n'est pas plein, l'insertion se fait par décalages internes

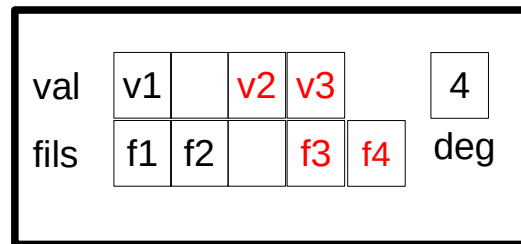


Insertion de **x** à la position 2 (et son fils droit **fd**), par **décalages internes**, dans un nœud **P** non plein :

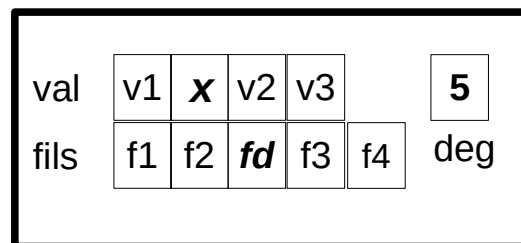
Avant :



Après décalages internes (valeurs et fils) :

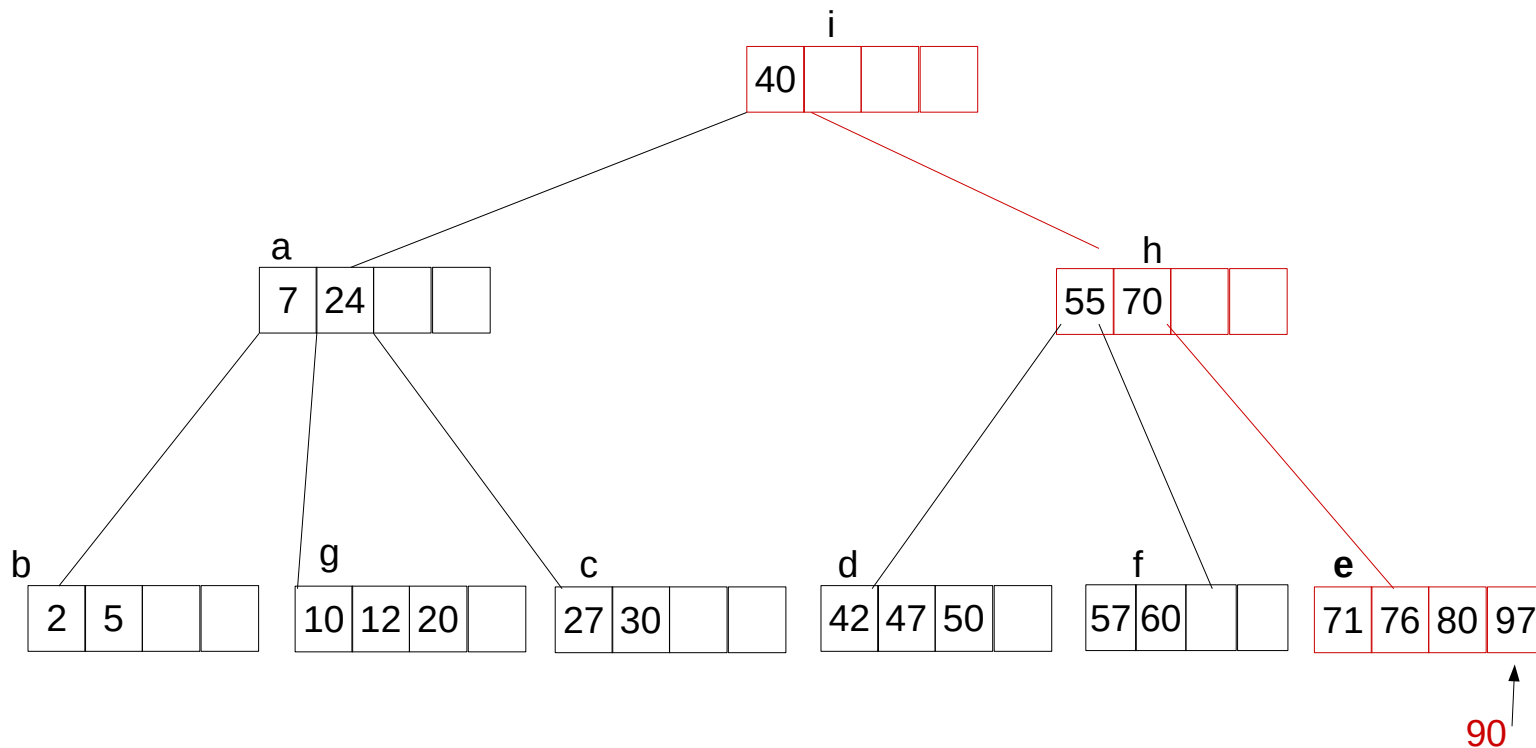


Après insertion de **x** (et son **fd**) et m-a-j du *degré* :



Résultat de l'insertion de la valeur 90

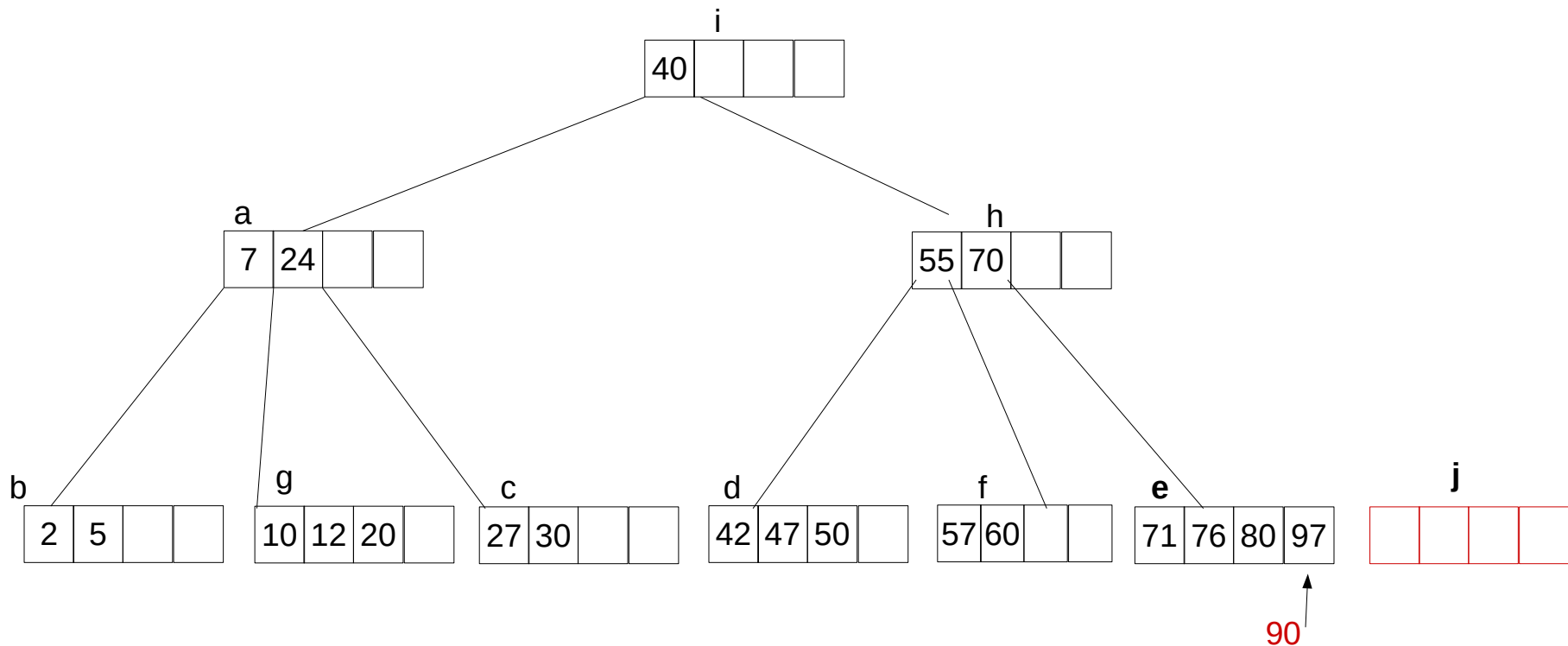
- La recherche de 90 mène vers le nœud feuille **e** (pos = 4 dans **e**)
- Comme **e** est déjà plein, il y aura éclatement du nœud :



Résultat de l'insertion de la valeur 90

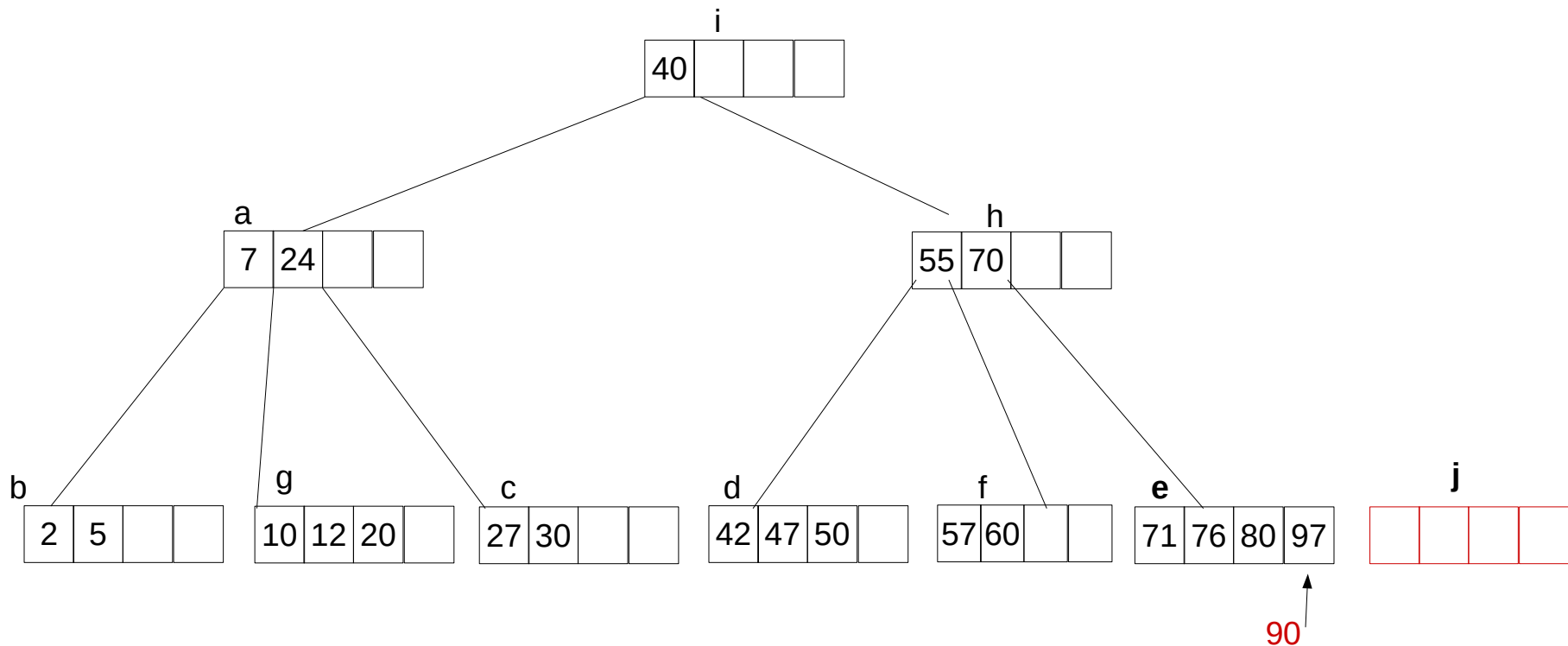
- La recherche de 90 mène vers le nœud feuille **e** (pos = 4 dans **e**)
- Comme **e** est déjà plein, il y aura éclatement du nœud :

1) Allocation d'un nouveau nœud (**j**)



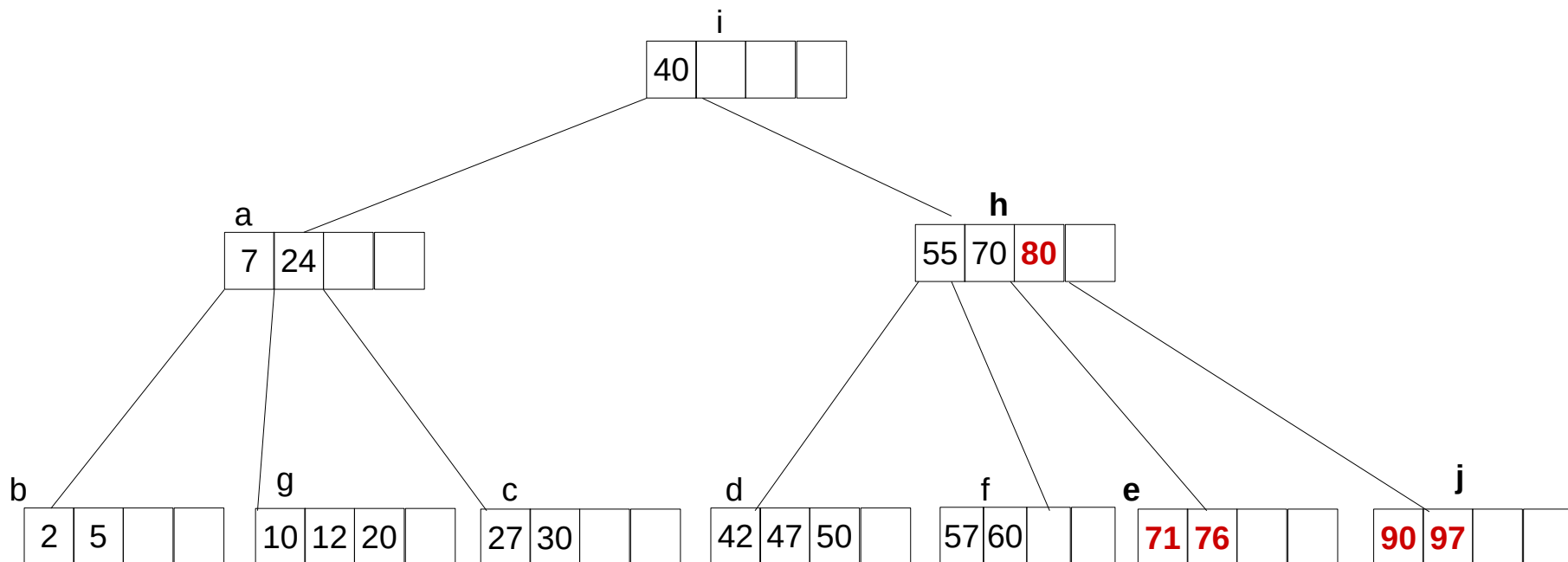
Résultat de l'insertion de la valeur 90

- La recherche de 90 mène vers le nœud feuille **e** (pos = 4 dans **e**)
- Comme **e** est déjà plein, il y aura éclatement du nœud :
 - 1) Allocation d'un nouveau nœud (**j**)
 - 2) formation de la séquence ordonnée : 71 76 80 90 97



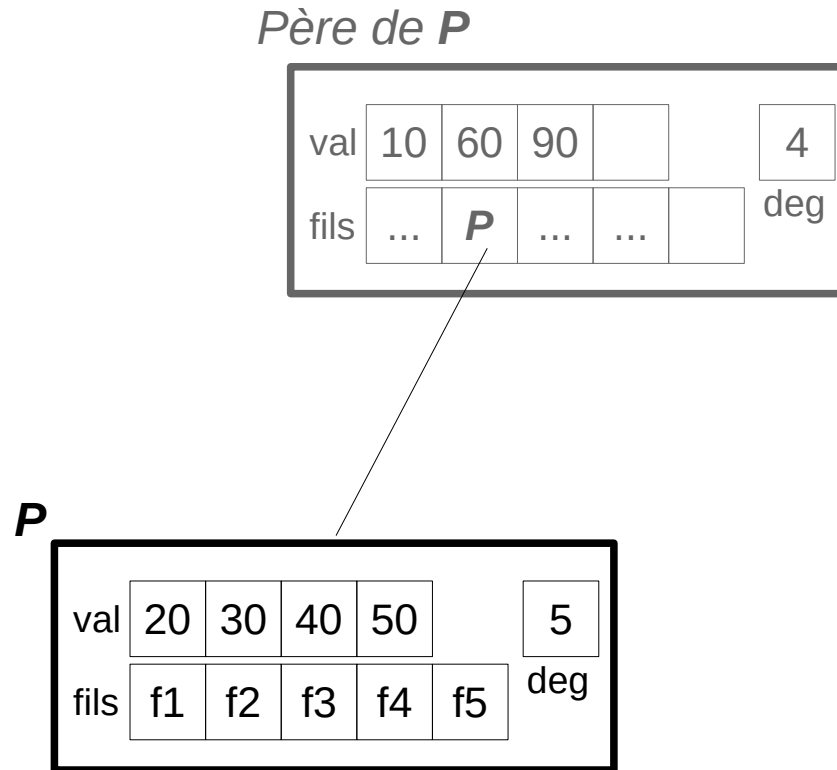
Résultat de l'insertion de la valeur 90

- La recherche de 90 mène vers le nœud feuille **e** (pos = 4 dans **e**)
- Comme **e** est déjà plein, il y aura éclatement du nœud :
 - 1) Allocation d'un nouveau nœud (**j**)
 - 2) formation de la séquence ordonnée : $[71\ 76]\ 80\ [90\ 97]$
 - 3) **Partage de la séquence** : $--(e)--(h)--(j)--$
Avec insertion de la valeur du milieu (80) dans le nœud parent h



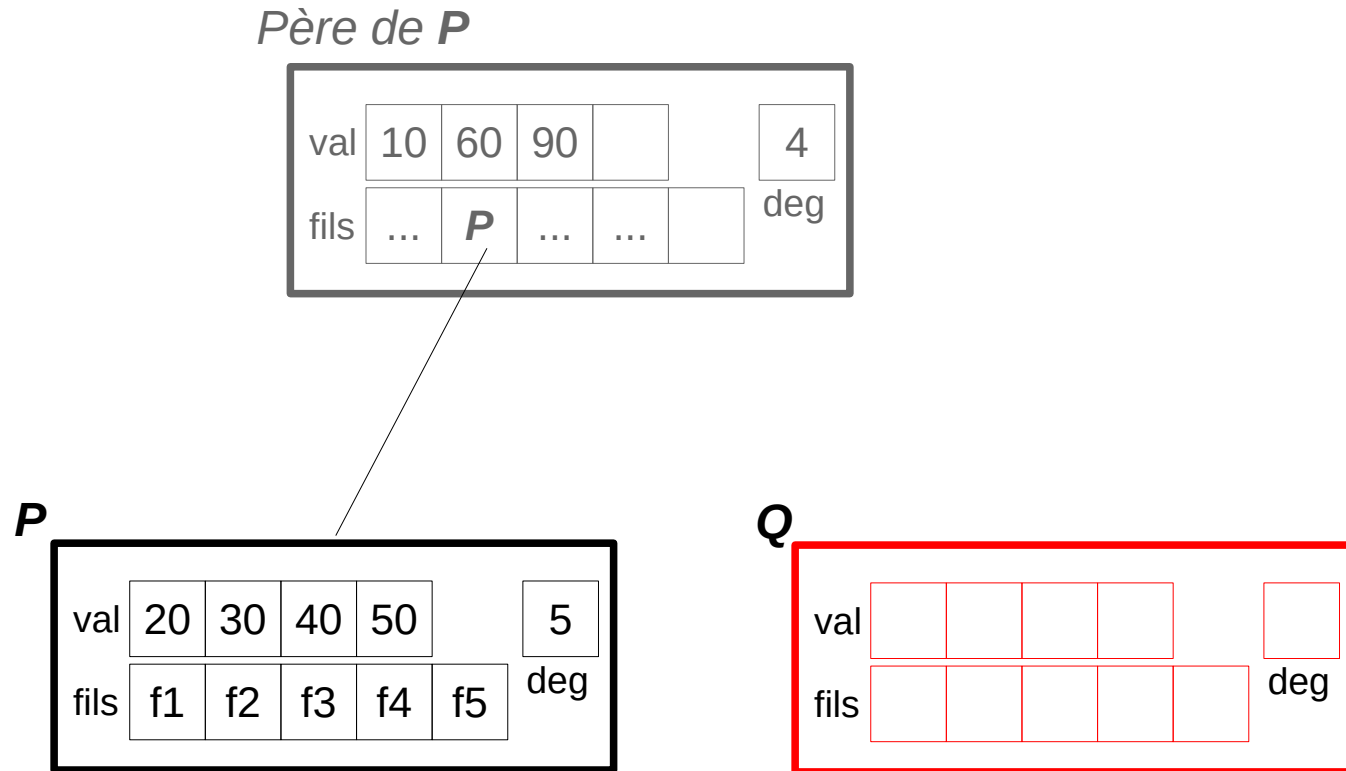
Insertion de $x = 25$ (et son fils droit *fd*), par **éclatement**, dans un nœud ***P*** plein :

Avant :



Insertion de $x = 25$ (et son fils droit **fd**), par **éclatement**, dans un nœud **P** plein :

Allocation d'un nouveau bloc : **Q**



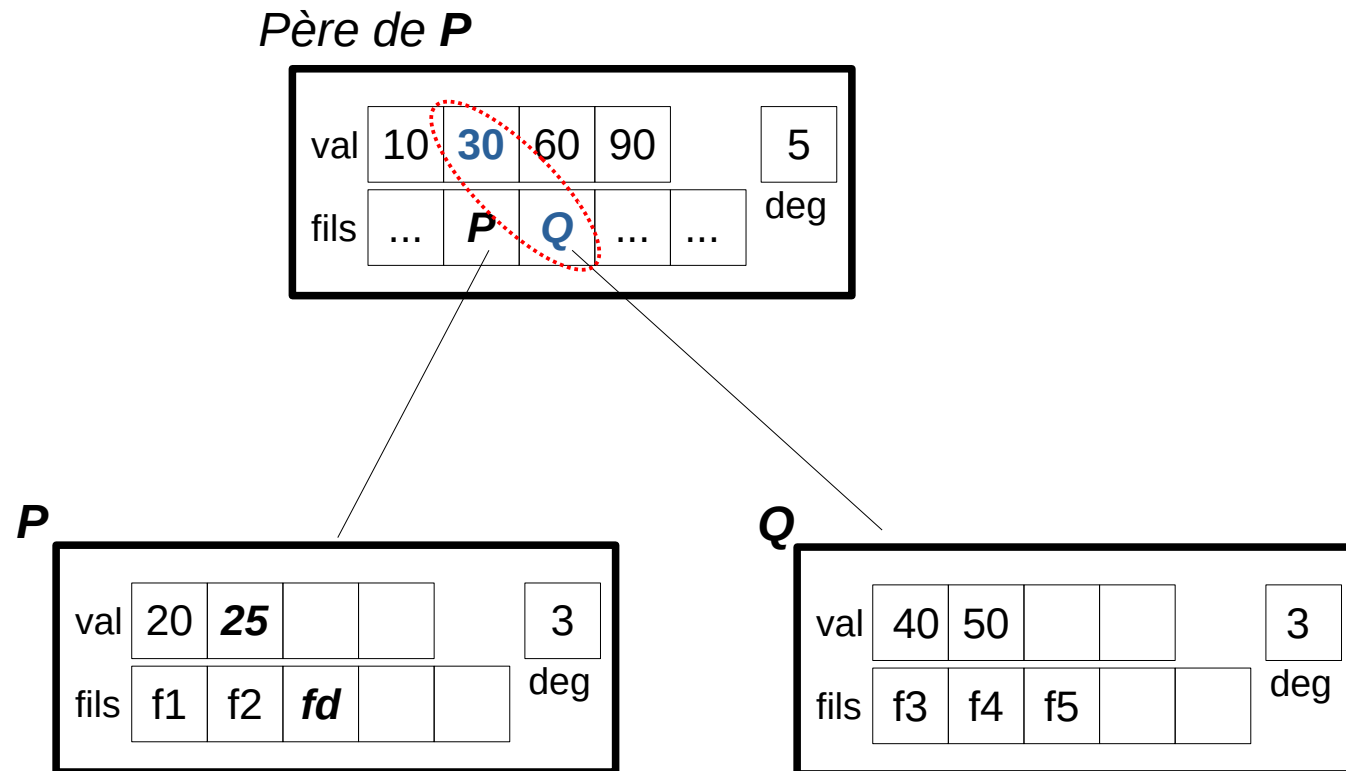
Insertion de $x = 25$ (et son fils droit fd), par **éclatement**, dans un nœud P plein :

Partage de la séquence ordonnée entre : P (1^{ère} moitié) et Q (dernière moitié)

[20 , 25 , **30** , 40 , 50]

---- P --- --- Q ---

et Insertion de la valeur du milieu 30 (avec son fils droit Q) dans le père de P



Insertion détaillée de x

Posons $N = 2d + 1$ (pour simplifier la présentation)
-un nœud peut donc contenir **entre d et $2d$** valeurs-

Rechercher x (avec pile) $\rightarrow P$ la feuille cible.

Initialement posons $t = -1$ (le fil à droite de x)

TQ (P contient déjà $2d$ valeurs)

// éclatement de P ...

$Q \leftarrow \text{AllocBloc}(\dots)$

Former la séquence ordonnée composée du contenu de $P + (x, t)$:

Val : $v_1 \ v_2 \ \dots \ x \ \dots \ v_{2d+1}$

// x est l'un des v_i (avec $i \in [1, 2d+1]$)

Fils : $f_1 \ f_2 \ f_3 \ \dots \ t \ \dots \ f_{2d+1} \ f_{2d+2}$

// t est l'un des f_{i+1}

La Séquence ordonnée contient maintenant **$2d+1$ valeurs** et **$2d+2$ fils**

= [$f_1 \ v_1 \ f_2 \ v_2 \ \dots \ f_d \ v_d \ f_{d+1} \ v_{d+1} \ f_{d+2} \ v_{d+2} \ \dots \ f_{2d+1} \ v_{2d+1} \ f_{2d+2}$]

Partager la séquence ordonnée entre P et Q en isolant la valeur du milieu (v_{d+1}) :

Les d premières valeurs et $d+1$ premiers fils [$f_1 \ v_1 \ f_2 \ v_2 \ f_3 \ \dots \ v_d \ f_{d+1}$] $\rightarrow P$

Les d dernières valeurs et $d+1$ derniers fils [$f_{d+2} \ v_{d+2} \ f_{d+3} \ v_{d+3} \ \dots \ v_{2d+1} \ f_{2d+2}$] $\rightarrow Q$

// la **valeur du milieu** avec Q comme fils droit : (v_{d+1}, Q) à insérer dans le père de P ...

$x \leftarrow v_{d+1}$; $t \leftarrow Q$; **SI** (Non Pilevide) Depiler(P) **SINON** $P \leftarrow \text{AllocBloc}(\dots)$ /* nouvelle racine */ **FSI**

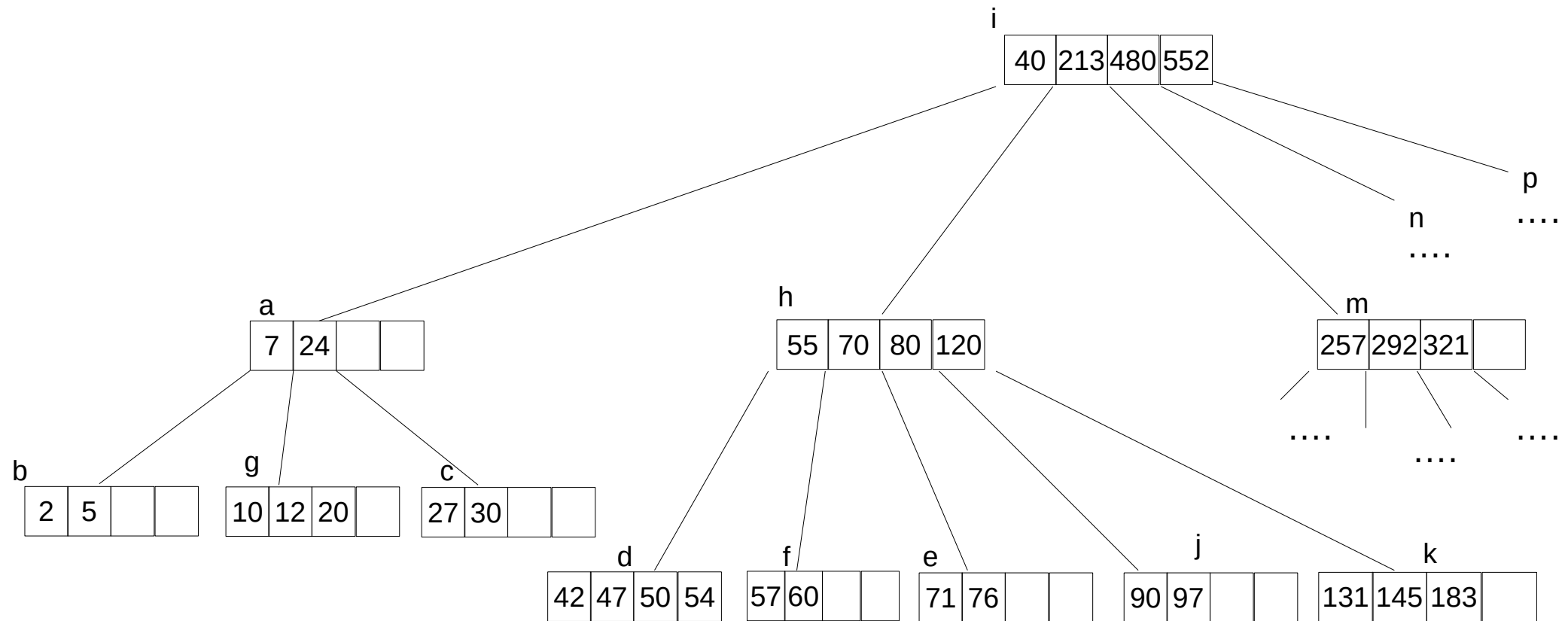
FTQ

// P contient donc moins de $2d$ valeurs

Insérer (x, t) dans P (par décalages internes)

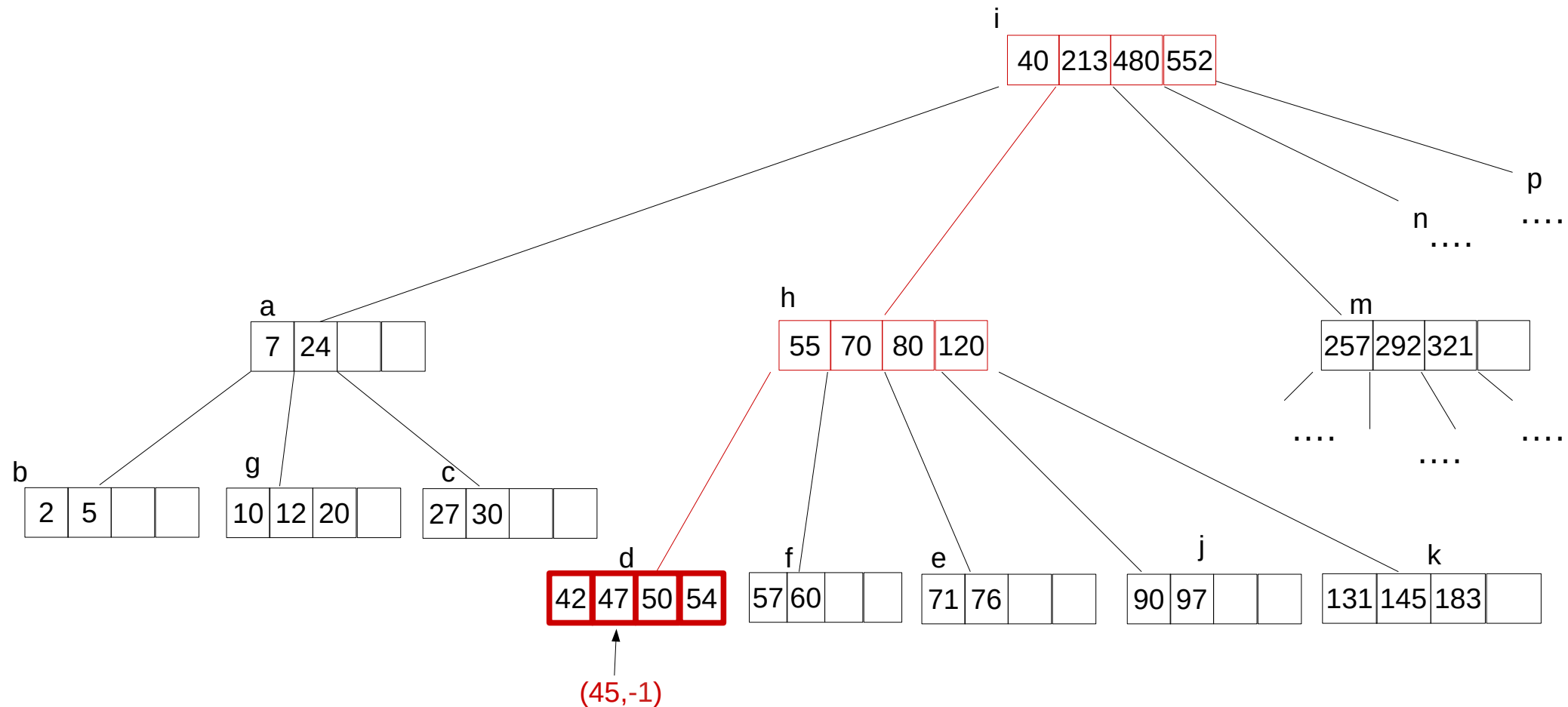
SI (P est une nouvelle racine) mettre l'ancienne racine comme *Fils1* de P **FSI**

Insertion de 45



Insertion de 45

Phase de descente (recherche de 45) : **lecture des blocs i , h et d**



Insertion de 45

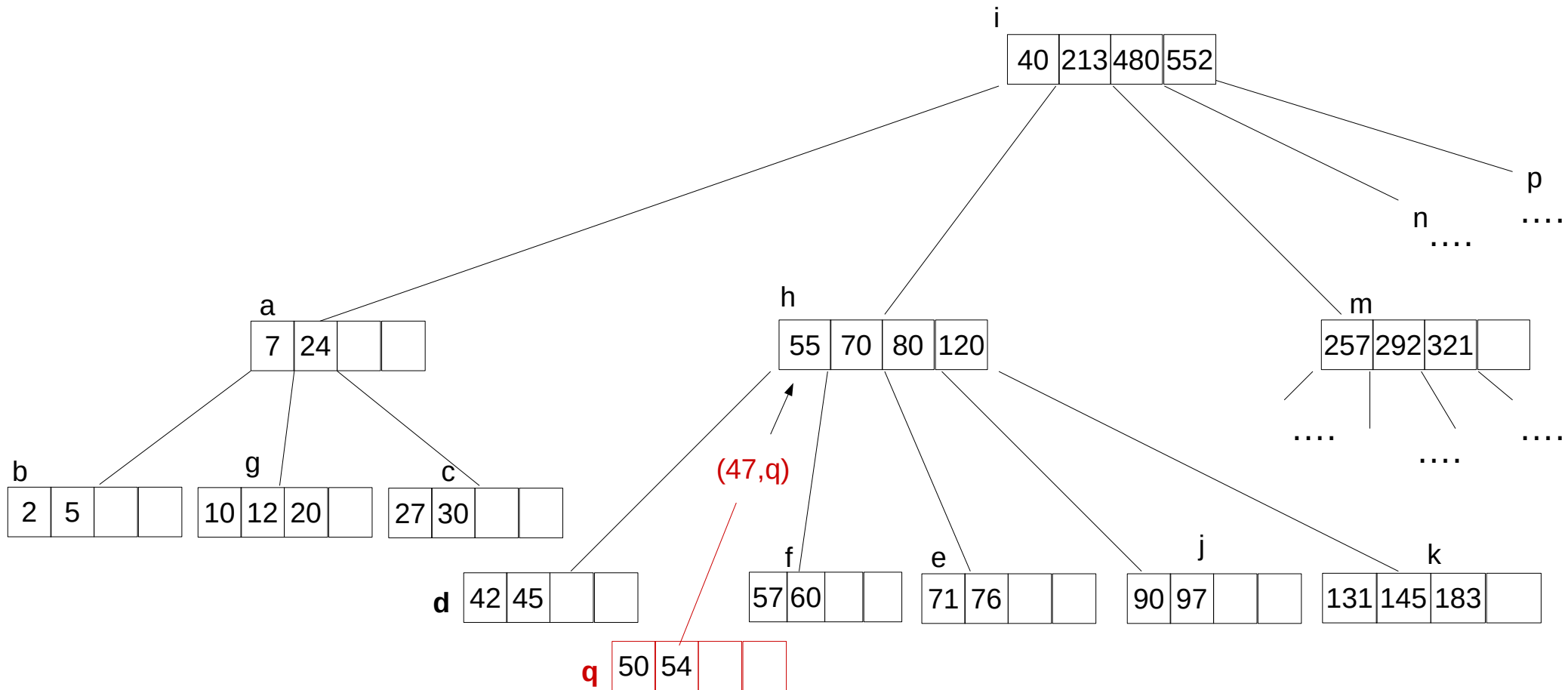
Eclatement du bloc d : Allocation du bloc q et partage de la séquence ordonnée

[42 45] 47 [50 54]

--- d --- (h) --- q ---

||

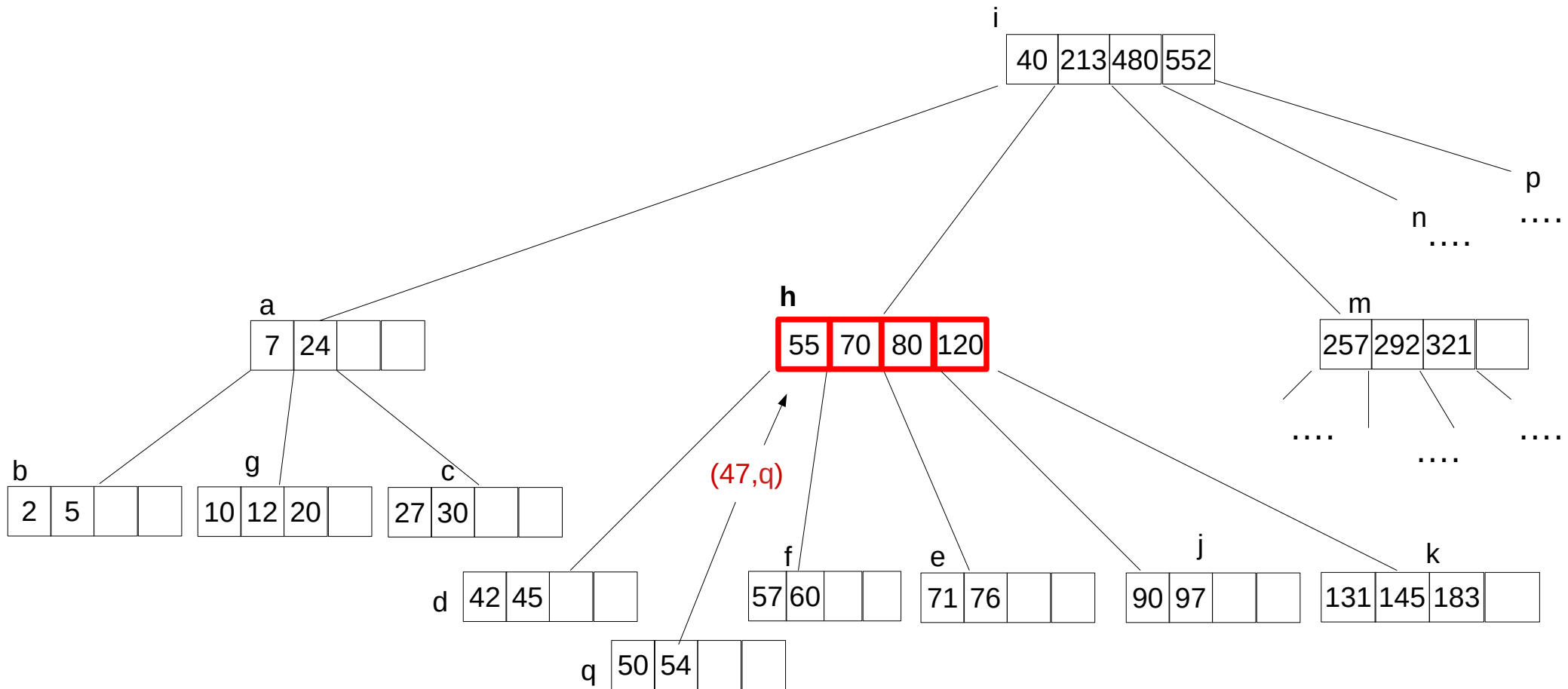
==> insertion de (47,q) dans le bloc h



Insertion de 45

h est plein \Rightarrow Eclatement du bloc **h** : **séquence ordonnée**

[_d **47** _q 55 _f 70 _e 80 _j 120 _k]

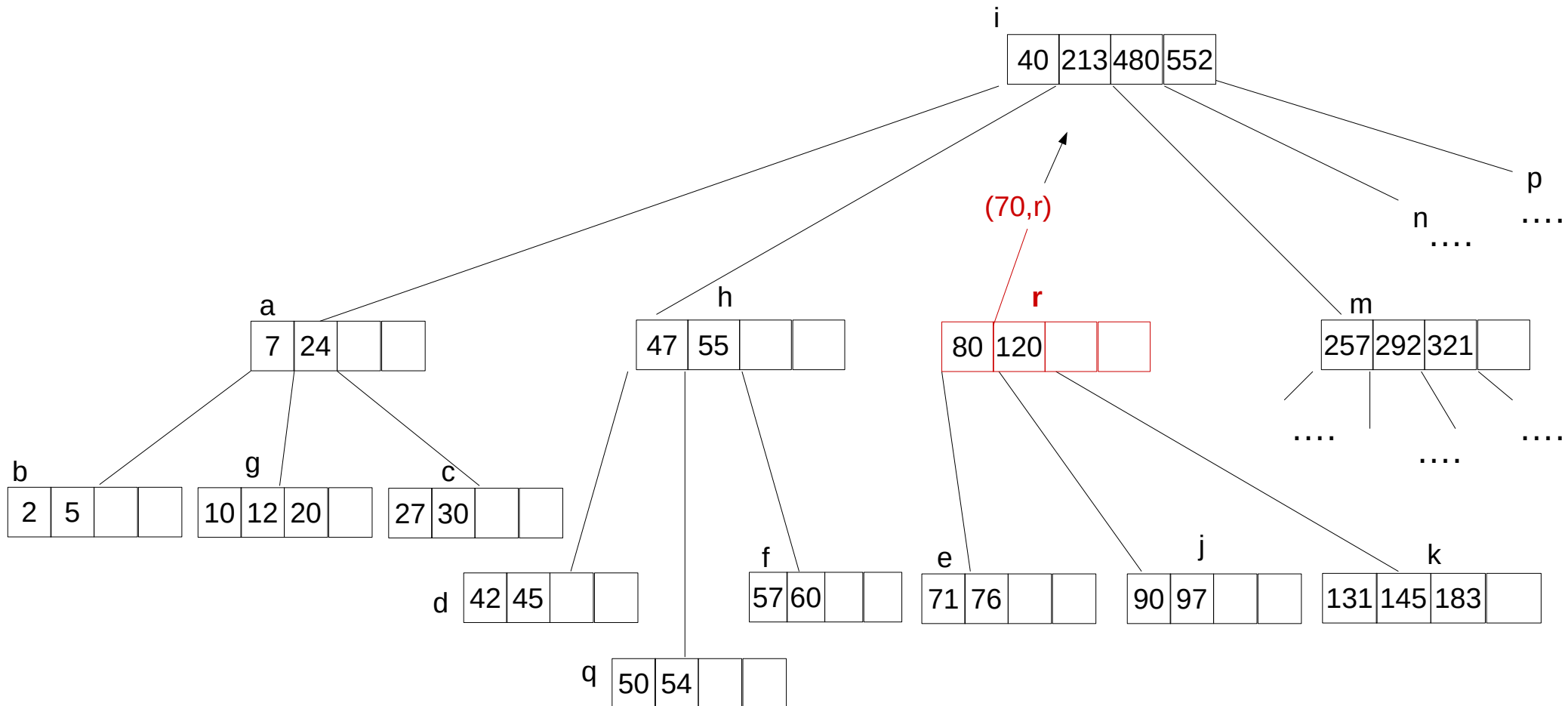


Insertion de 45

Eclatement du bloc h : Allocation du bloc r et partage de la séquence ordonnée

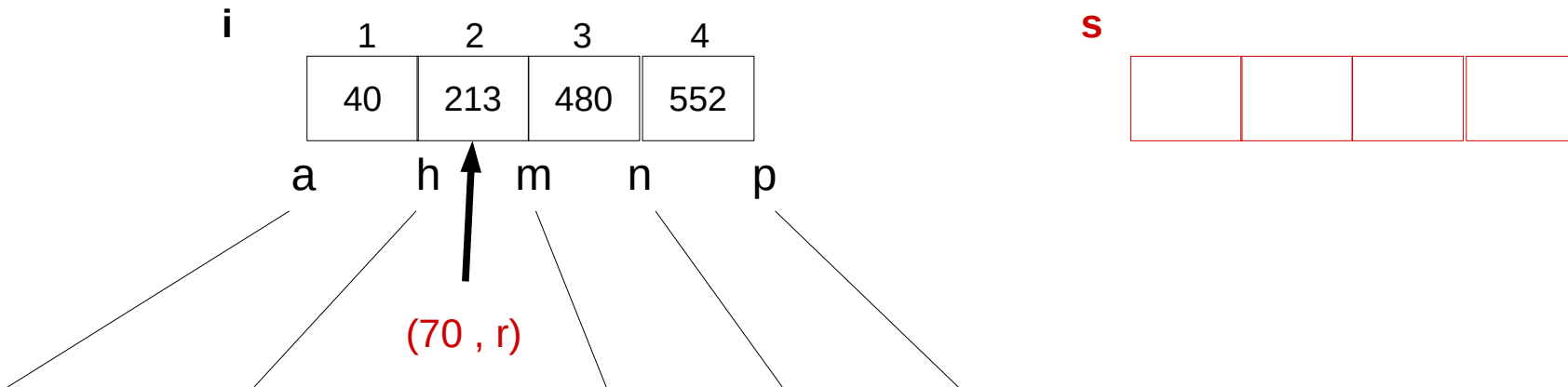
[d 47 q 55 f] 70 [e 80 j 120 k]
|----- h -----| |----- r -----|

insertion de (70,r) dans le bloc i



Insertion de 45

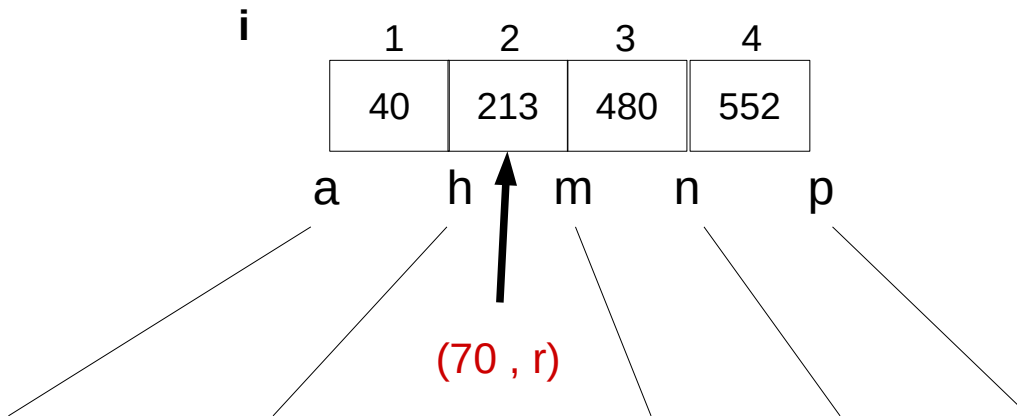
i est plein aussi \Rightarrow Eclatement du bloc i : Allocation du bloc s



Insertion de 45

Eclatement du bloc i : **formation** de la séquence ordonnée

[a 40 h **70** r 213 m 480 n 552 p]



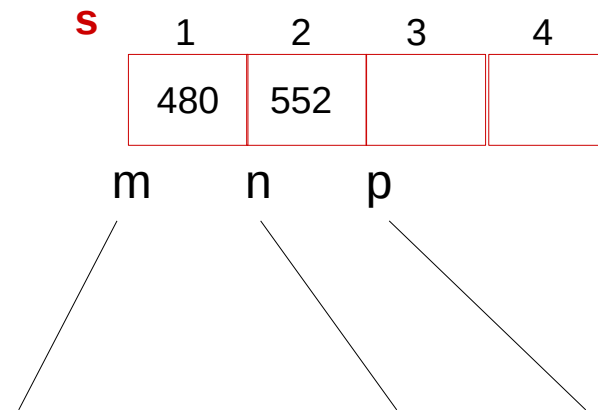
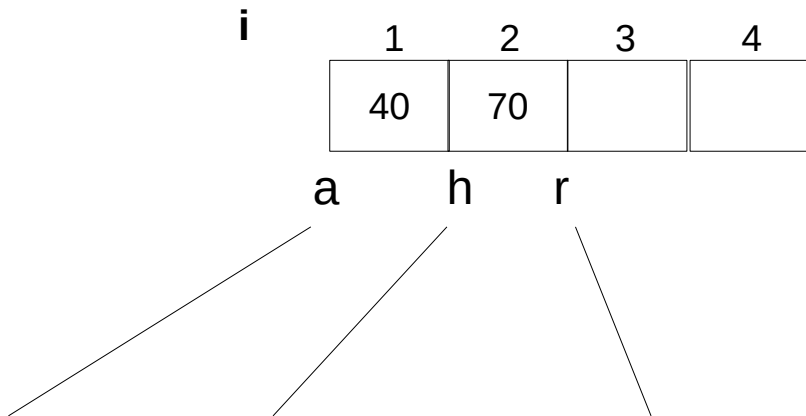
Insertion de 45

Eclatement du bloc i : **partage** de la séquence ordonnée

vers le
père de i

[a 40 h 70 r] 213 [m 480 n 552 p]
|----- i -----| |----- s -----|

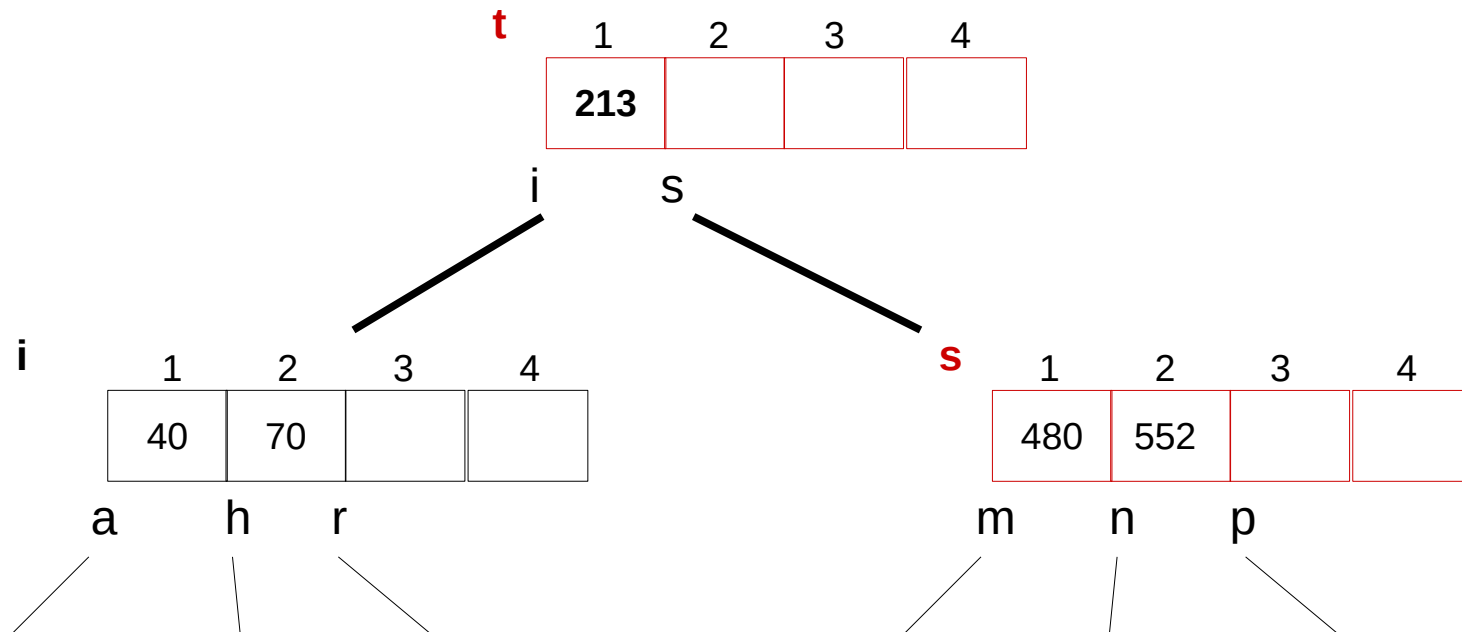
... 213 ...



Insertion de 45

Eclatement du bloc i :

Une nouvelle racine (**t**) est allouée pour recevoir la valeur **213** avec comme fils1 le nœud **i** et comme fils2 le nœud **s**



Insertion de 45

- Eclatement du bloc **d**: Allocation du bloc **q**
- Eclatement du bloc **h**: Allocation du bloc **r**
- Eclatement du bloc **i** : Allocation du bloc **s**
- Nouvelle racine : Allocation du bloc **t**

