

Cours de bases de données,
aspects systèmes,
<http://sys.bdpedia.fr>

Hachage linéaire

Le hachage linéaire

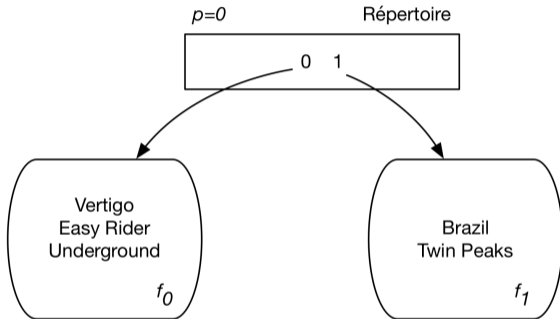
Hachage linéaire : un hachage qui s'adapte dynamiquement aux insertions et suppressions, avec une **évolution minimale** du répertoire et des fragments.

titre	$h(\text{titre})$
Vertigo	14
Brazil	43
Twin Peaks	25
Underground	24
Easy Rider	8
Psychose	33
Greystoke	17
Shining	16
Citizen Kane	48

Une fonction $h(c)$ immuable s'applique à la clé et renvoie un entier.

La structure

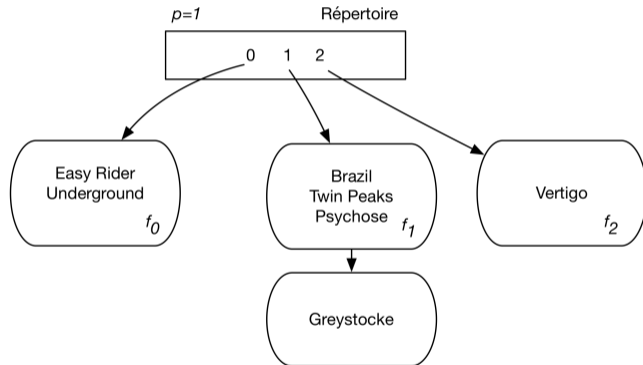
Comme les autres hachages, plus un paramètre : **l'indice de partitionnement**.



On considère la famille de fonctions $h_i(c) = h(c) \bmod 2^i$. Ici on applique h_1 (modulo 2).

Insertions

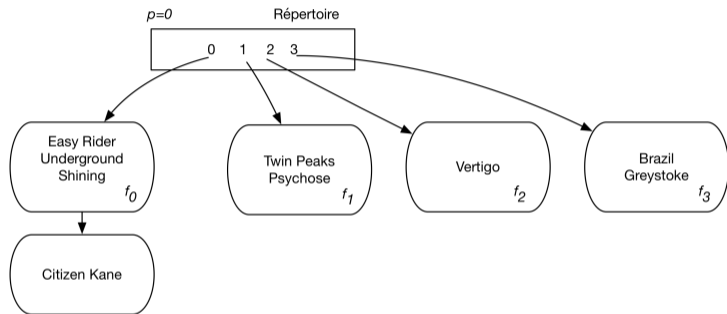
On insère Psychose, puis Greystoke qui vont tous deux dans f_1 . C'est f_1 qui déborde, **mais c'est f_0 qui éclate en (f_0, f_2)** .



Maintenant on a deux fonctions : h_1 s'applique pour f_1 , h_2 pour les autres. Et p vaut 1.

On insère Shing puis Citizen Kane

Tous les deux vont dans f_0 qui déborde. Mais maintenant c'est f_1 qui éclate.



On a **découplé** les éclatements et les débordements. On sait que les seconds seront rectifiés par les premiers, **à terme**.

p est revenu à 0 ; la fonction est h_2 : c'est reparti pour un tour.

Résumé du hachage linéaire

À chaque moment, on a deux fonctions (h_n, h_{n+1}) et un indice de partitionnement p .

- La fonction h_{n+1} s'applique à tous les fragments **avant** p car ils ont déjà éclaté.
- La fonction h_n à tous les autres fragments qui n'ont pas encore éclaté.

Quand le dernier fragment éclate ($p = n$) la fonction h_n ne sert plus et la paire de fonction de hachage devient (h_{n+1}, h_{n+2})

Brillant !