Skip lists - Listes à sauts

François Schwarzentruber

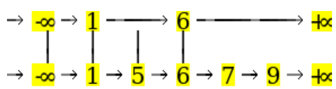
8 mai 2023

1 Définition

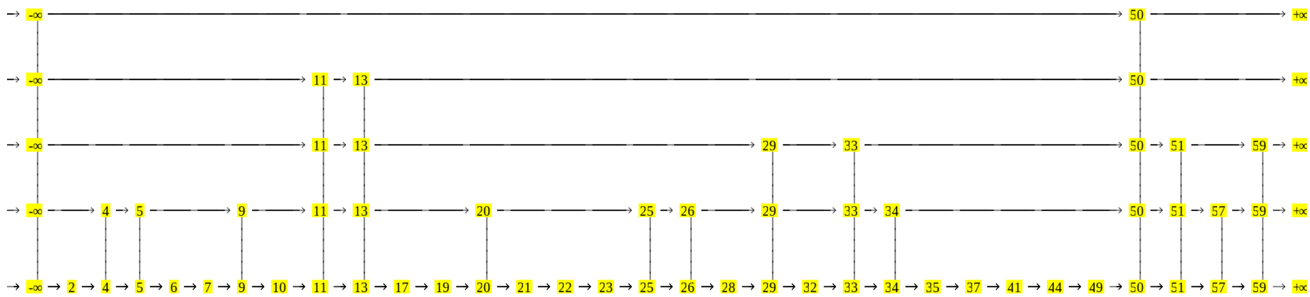
Vidéo 1 `https://www.youtube.com/watch?v=HIrh1NJmQxM`

Définition 2 (informelle) Une *skip list* est une liste simplement chaînée triée, avec des voies rapides (niveaux) en plus, pour prendre des raccourcis. Chaque niveau raffine le précédent. Le dernier niveau est la liste chaînée. Chaque nœud possède plusieurs pointeurs, un par voie. Le premier élément est une *sentinelle* et vaut $-\infty$.

Exemple 3 Voici une skip list avec 2 niveaux. Le niveau du bas est la liste chaînée habituelle, et le niveau du haut est la voie rapide.



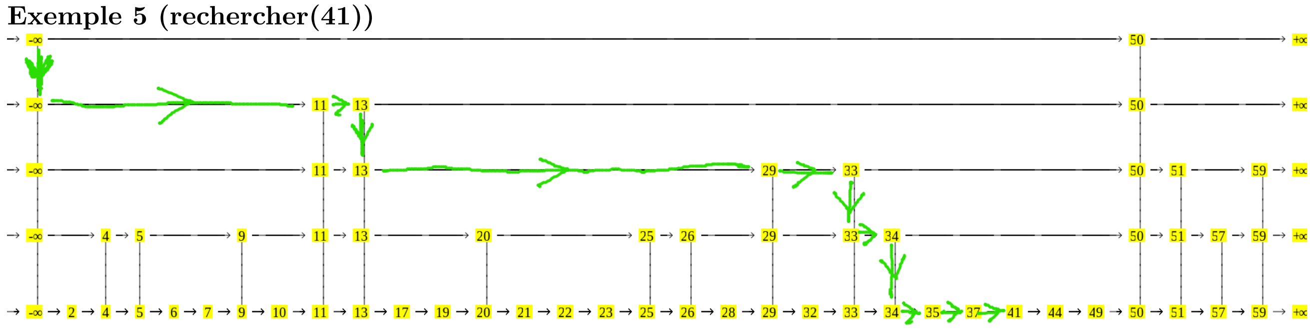
Exemple 4 Voici une skip list avec 5 niveaux.



2 Recherche

fonction rechercher(x)
  skippy $:= -\infty$ sur le niveau le plus rapide (en haut à gauche)
  tant que *skippy* n'est pas sorti
    avancer *skippy* vers la droite le plus possible sans dépasser v
    si x trouvé **alors**
      renvoyer oui
    sinon
      descendre *skippy* d'un niveau
  renvoyer non

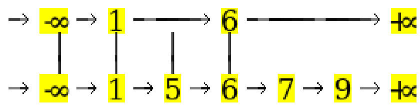
Exemple 5 (rechercher(41))



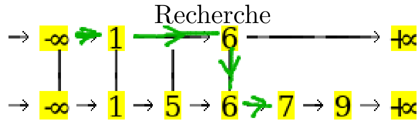
3 Ajout

```
fonction ajouter(x)
    aller à la position dans le niveau du bas où x doit être inséré (voir rechercher)
    y insérer x
    boucle
        lancer une pièce de monnaie
        si pile alors
            promouvoir x
        sinon
            s'arrêter
```

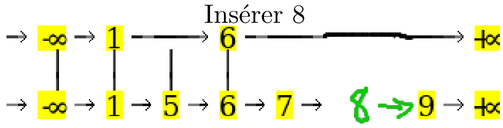
Exemple 6 (ajouter(8))



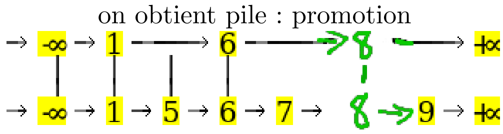
Recherche



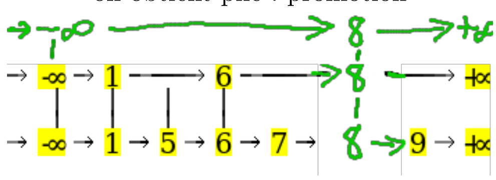
Insérer 8



on obtient pile : promotion



on obtient pile : promotion



on obtient face : on s'arrête

4 Meilleur équilibrage

Proposition 7 Pour le meilleur équilibrage d'une skiplist à k niveaux et n éléments, la complexité pire cas de la recherche est $O(kn^{1/k})$.

Proposition 8 Pour le meilleur équilibrage d'une skiplist à $\log n$ niveaux et n éléments, la complexité pire cas de la recherche est $O(\log n)$.

5 Rappels de probabilité

Proposition 9 Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On a : $\mathbb{E}(X) = \sum_{k=1}^{\infty} \mathbb{P}(X \geq k)$.

Proposition 10 (loi géométrique de paramètre $\frac{1}{2}$) Soit X le nombre de piles consécutifs en lançant une pièce de monnaie équilibrée. On a : $\mathbb{E}(X) = 1$.

Définition 11 (espérance conditionnelle) Soit X une variable aléatoire discrète et E un événement de probabilité non nulle.

$$\mathbb{E}(X \mid E) = \frac{1}{\mathbb{P}(E)} \sum_{k=1}^{\infty} k\mathbb{P}(X{=}k \ \cap \ E).$$

Proposition 12 Si $A \sqcup B = \Omega$ (partition) alors $\mathbb{E}(X) = \mathbb{E}(X \mid A)\mathbb{P}(A) + \mathbb{E}(X \mid B)\mathbb{P}(B)$.

6 Complexité

Proposition 13 Le nombre de niveau d'une skip list avec n éléments est $O(\log n)$ avec une grande probabilité. Plus précisément, pour tout $c \geq 1$, $\mathbb{P}(\text{nb de niveaux} > c\log n) \leq \frac{1}{n^{c-1}}$.

Corollaire 14 $\mathbb{E}(\text{nb de niveaux}) = O(\log n)$.

Proposition 15 L'espérance de la mémoire utilisée est $O(n)$.

Proposition 16 Une recherche dans une skip list coûte $O(\log n)$ en temps espéré.

7 Encore des rappels de probabilité

Théorème 17 (inégalité de Chernoff) Soit Y le nombre total de faces dans une suite de m lancées de pièces. Pour tout $r > 0$,

$$\mathbb{P}(Y \geq \frac{m}{2} + r) \leq e^{-2r^2/m}.$$

8 Complexité avec grande probabilité

Lemme 18 Pour tout constante c, il existe d tel que, avec grande probabilité, le nombre de piles en lançant $d\log n$ pièces est au moins $c\log n$. Plus formellement :

$$\mathbb{P}(\text{nb de piles obtenus durant au moins } d\log n \text{ lancés} \geq c\log n) \quad \geq \quad 1 - \frac{1}{n^c}.$$

Proposition 19 Une recherche dans une skip list coûte $O(\log n)$ avec une grande probabilité.

9 Implémentation en Python

`https://github.com/francoisschwarzentruber/skiplist/blob/main/skiplist.py`

10 Notes bibliographiques

Les skip lists ont été inventées par William Pugh [**?**]. Elles ont inspiré les *skip graphs* utilisés en réseau pair-à-pair [**?**]. Vous trouvez des informations pour améliorer l'implémentation ici : [**?**].

Références