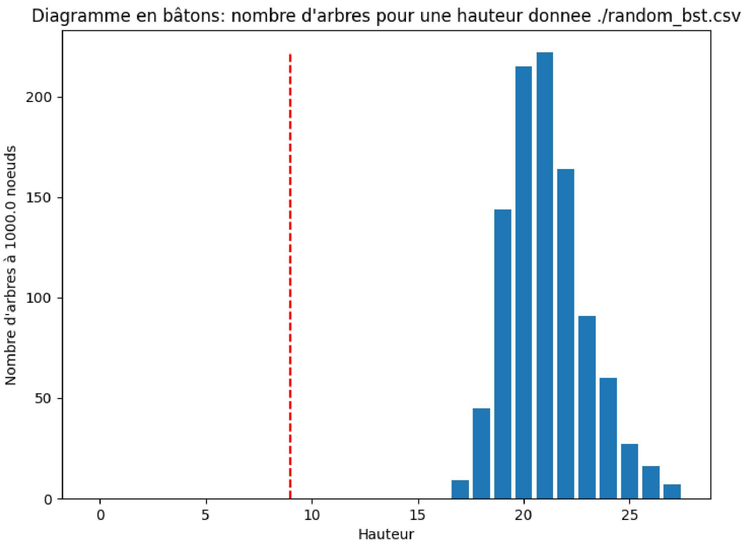
TP n°21 - Équilibrage des ABR - Corrigé

Exercice 1 (Analyse de l'équilibre d'ABR générés aléatoirement sans rééquilibrage).
Distribution des hauteurs pour 1000 arbres de 1000 nœuds générés aléatoirement sans rééquilibrage. La hauteur optimale (celle correspondant à un arbre complet) est indiqué en pointillés rouges :



Exercice 2 (Les arbres AVL sont équilibrés).

1. Soit $\mathcal{A}$ l'ensemble des arbres AVL. Montrer qu'il existe une constante $c > 1$ telle que, pour tout $t \in \mathcal{A}$, on a :

$$c^{h(t)+1} - 1 \leq n(t) \leq 2^{h(t)+1} - 1$$

Que vaut c ?

2. En déduire que les arbres AVL sont équilibrés au sens de la première définition.

3. Donner la plage de valeurs de hauteur pour des arbres AVL à 1000 nœuds puis à 100000 nœuds. Quelle aurait été la borne supérieure de cette plage de valeurs sans la propriété AVL ?

Corrigé de l'exercice 1. [Retour à l'énoncé]

1. La majoration a déjà été démontrée pour tous les arbres binaires, donc a fortiori pour les arbres AVL. On se concentre donc uniquement sur la preuve de la minoration. On montre cette propriété par induction structurelle : soit $c > 1$ une constante. Soit $t \in \mathcal{A}$.

$$\mathcal{P}(t) : c^{h(t)+1} - 1 \leq n(t)$$

Cas de base. Pour l'arbre vide, on a $h(E) = -1$ et $n(E) = 0$. L'inégalité ci-dessus donne

$$0 = c^0 - 1 \leq 0$$

qui est vraie. Donc $\mathcal{P}(E)$ est vraie.

Induction. On fixe $t \in \mathcal{A} \setminus \{E\}$, on considère l'ordre structurel $\preceq$ sur $\mathcal{A}$ et on suppose que, $\forall t' \prec t$, $\mathcal{P}(t')$ est vraie (hypothèse d'induction).
Comme $t \neq E$, t a été construit avec le constructeur N et on a $t = N(\ell, x, r)$. Par définition de l'ordre structurel, $\ell \prec t$ et $r \prec t$: on peut donc appliquer l'hypothèse d'induction à ces deux sous-arbres :

$$\ell \prec t \to c^{h(\ell)+1} - 1 \leq n(\ell)$$
$$r \prec t \to c^{h(r)+1} - 1 \leq n(r)$$

Or, par définition $n(t) = 1 + n(\ell) + n(r)$, on a donc :

$$n(t) = 1 + n(\ell) + n(r) \geq 1 + (c^{h(\ell)+1} - 1) + (c^{h(r)+1} - 1)$$

$$\Leftrightarrow n(t) \geq c^{h(\ell)+1} + c^{h(r)+1} - 1$$

La hauteur de t est définie inductivement par :

$$h(t) = 1 + \max\left(h(\ell),\, h(r)\right)$$

Procédons par disjonction des cas

Supposons que $h(l) \geq h(r)$**.** On a donc $h(t) = 1 + h(l)$ et $0 \leq h(l) - h(r) \leq 1$ car $t \in \mathcal{A}$ est supposé AVL.

$$\begin{aligned} n(t) &\geq c^{h(\ell)+1} + c^{h(r)+1} - 1 \\ &= c^{h(\ell)+2}\left(\frac{1}{c} + c^{h(r)+1-h(l)-2}\right) - 1 \\ &= c^{(h(\ell)+1)+1}\left(\frac{1}{c} + \frac{1}{c^{h(l)-h(r)+1}}\right) - 1 \\ &= c^{h(t)+1}\left(\frac{1}{c} + \frac{1}{c^{h(l)-h(r)+1}}\right) - 1 \\ &\geq c^{h(t)+1}\left(\frac{1}{c} + \frac{1}{c^2}\right) - 1 \text{ car } c > 1 \text{ et } h(l) - h(r) \leq 1 \end{aligned}$$

Si on arrive à trouver une constante $c > 1$ telle que :

$$\frac{1}{c} + \frac{1}{c^2} \geq 1$$

c'est gagné. Et tant qu'à faire, on cherche le plus grand $c > 1$ possible vérifiant cette condition : dans ce cas, on aura démontré notre minoration

$$n(t) \geq c^{h(t)+1} - 1$$

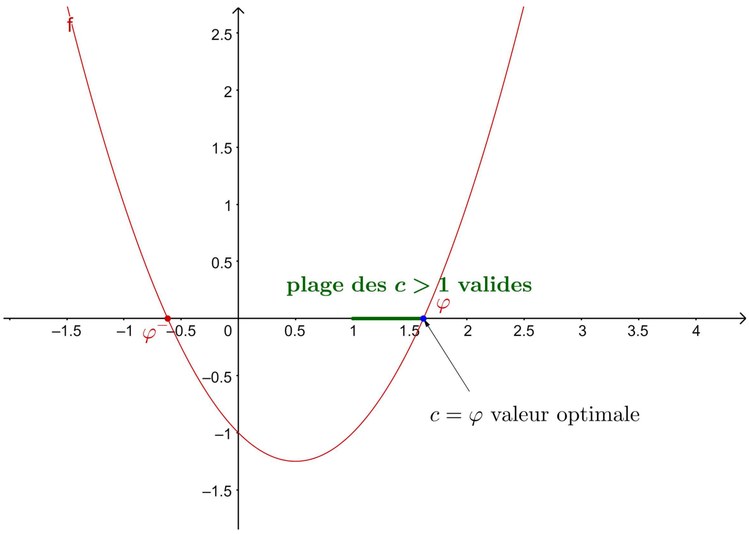
avec $c > 1$ optimale : c'est la minoration la plus fine que l'on puisse trouver.
Cherchons donc tous les $c > 1$ vérifiant :

$$\frac{1}{c} + \frac{1}{c^2} \geq 1$$

De telles valeurs de c vérifient donc (on multiplie par $c^2 > 0$ de deux côtés et on passe tout à droite par exemple) :

$$P(c) \leq 0$$

où $P(X) = X^2 + X - 1$. Ce polynôme du second degré a deux racines réelles : l'une positive, $\varphi = \dfrac{1+\sqrt{5}}{2} \approx 1.6$ (nombre d'or) et une autre négative $\varphi^- = \dfrac{1-\sqrt{5}}{2}$.



L'inégalité est donc vraie pour tous les $c \in]1; \varphi]$, mais elle est optimale (car il y a égalité et non minoration lors de la dernière étape) pour :

$$c = \varphi = \frac{1+\sqrt{5}}{2} \approx 1.6$$

On a donc montré que, pour cette valeur de $c = \varphi > 1$, $\mathcal{P}(t)$ est vraie (et optimale !) sous hypothèse d'induction.

Supposons que $h(r) \geq h(l)$**.** Ce cas se traite de manière totalement symétrique.

On a donc montré que dans les deux cas de la disjonction, $\mathcal{P}(t)$ est vraie (et optimale !) sous hypothèse d'induction.

Conclusion : Par principe d'induction, la propriété est vraie pour tous les arbres binaires AVL avec $c = \varphi = \dfrac{1+\sqrt{5}}{2} > 1$.

2. On a déjà, en prenant le logarithme base 2 de la deuxième inégalité que :

$$\lceil \log_2(n(t)+1) \rceil - 1 \leq h(t)$$

On prend le logarithme base $c > 1$ de l'inégalité de gauche pour trouver la majoration :

$$h(t) \leq \left\lfloor \log_\varphi(n(t)+1) \right\rfloor - 1$$

On donc écrire :

$$\lceil \log_2(n(t)+1) \rceil - 1 \leq h(t) \leq \left\lfloor \log_\varphi(n(t)+1) \right\rfloor - 1$$

ce qui se réécrit :

$$\lceil \log_2(n(t)+1) \rceil - 1 \leq h(t) \leq \underbrace{\frac{\ln(2)}{\ln(\varphi)}}_{\approx 1.44} \lfloor \log_2(n(t)+1) \rfloor - 1$$

3.
- Pour $n(t) = 1000$, on obtient $h(t) \in [\![9; 13]\!]$
- Pour $n(t) = 100000$, on obtient $h(t) \in [\![16; 24]\!]$

C'est beaucoup mieux que la majoration sur les arbres binaires générale qui est :

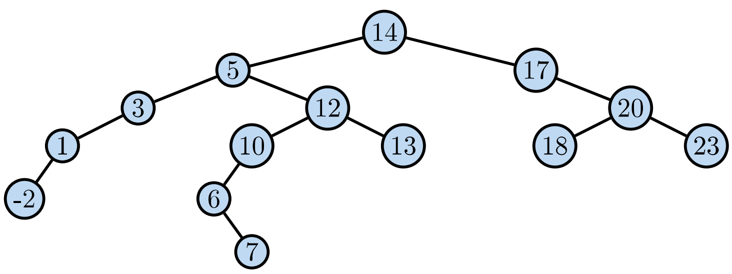
$$h(t) \leq n(t) - 1$$

- Pour $n(t) = 1000$, on obtient $h(t) \in [\![9; 999]\!]$
- Pour $n(t) = 100000$, on obtient $h(t) \in [\![16; 99999]\!]$

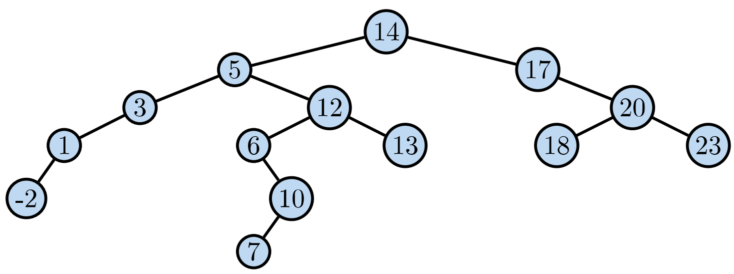
Exercice 3 (Opérations de rotation (gauche ou droite) autour d'un nœud).

1. Au considère l'ABR de la figure ci-dessous (c'est le même arbre que celui du premier TP sur les ABR !). Dessiner l'arbre obtenu après une rotation droite autour du nœud d'étiquette 10.
2. A partir de l'arbre modifié, faire à présent une rotation gauche autour du nœud d'étiquette 17 de l'arbre donné ci-dessus.
3. Faire à présent une rotation droite autour du nœud d'étiquette 14.
4. Montrer que les opérations de rotation gauche et droite conservent la propriété d'ABR (et vérifiez le sur vos manipulations précédentes !)

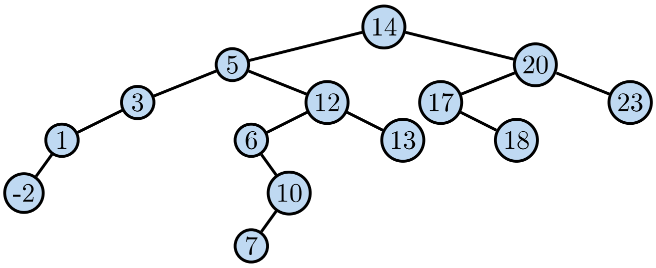
Corrigé de l'exercice 2. [Retour à l'énoncé]



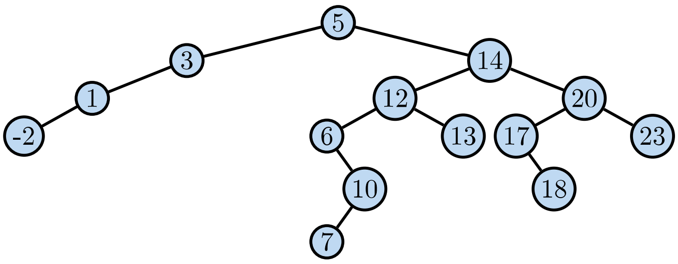
1. Rotation droite autour du nœud d'étiquette 10 :



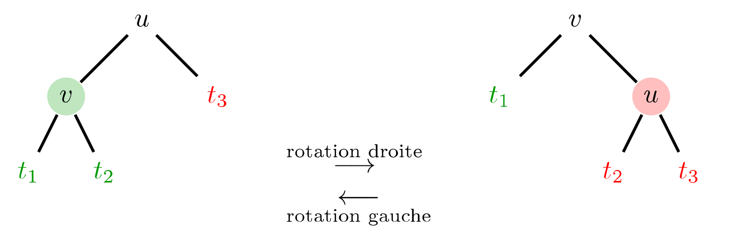
2. Rotation gauche autour du nœud d'étiquette 17 :



3. Rotation droite autour du nœud d'étiquette 14 :



4. Repartons du schéma décrivant les opérations de rotation gauche et droite :



$t_1 < v < t_2 < u < t_3$ $\qquad$ $t_1 < v < t_2 < u < t_3$

Si l'ordre $t_1 < v < t_2 < u < t_3$ est vrai à gauche, il reste vrai à droite après une rotation droite et réciproquement, s'il est vrai à droite, il reste vrai après une rotation gauche.