Fonction G de Hofstadter et au-delà: un exemple curieux mêlant calculs et preuves sur ordinateur

Pierre Letouzey (IRIF, UPC, Inria)

FSMP, 2 décembre 2023

Avertissement : la préparation de cet exposé a subi. . .

La loi d'Hofstadter : il faut toujours plus longtemps que prévu, même en tenant compte de la loi d'Hofstadter.

Prélude : une première fonction récursive

Que dire de la fonction $d : \mathbb{N} \to \mathbb{N}$ définie ainsi ?

$$\begin{cases} d(0) = 0 \\ d(1) = 0 \\ d(n) = 1 + d(n-2) & \text{pour } n \geq 2 \end{cases}$$

Prélude : une première fonction récursive

Que dire de la fonction $d : \mathbb{N} \to \mathbb{N}$ définie ainsi ?

$$\begin{cases} d(0) = 0 & \\ d(1) = 0 & \\ d(n) = 1 + d(n-2) & \text{pour } n \geq 2 \end{cases}$$

n	0	1	2	3	4	5	6	...
$d(n)$	0	0	1	1	2	2	3	...

Prélude : une première fonction récursive

Que dire de la fonction $d : \mathbb{N} \to \mathbb{N}$ définie ainsi ?

$$\begin{cases} d(0) = 0 \\ d(1) = 0 \\ d(n) = 1 + d(n-2) & \text{pour } n \geq 2 \end{cases}$$

n	0	1	2	3	4	5	6	...
$d(n)$	0	0	1	1	2	2	3	...

En fait $d(n) = \lfloor n/2 \rfloor$ (arrondi à l'entier inférieur).

Prélude : une première fonction récursive

Que dire de la fonction $d : \mathbb{N} \to \mathbb{N}$ définie ainsi ?

$$\begin{cases} d(0) = 0 \\ d(1) = 0 \\ d(n) = 1 + d(n-2) & \text{pour } n \geq 2 \end{cases}$$

n	0	1	2	3	4	5	6	...
$d(n)$	0	0	1	1	2	2	3	...

En fait $d(n) = \lfloor n/2 \rfloor$ (arrondi à l'entier inférieur).

Justification : par récurrence !

Plus délicat : soustrayons !

Que dire de la fonction $f_0 : \mathbb{N} \to \mathbb{N}$ définie ainsi ?

$$\begin{cases} f_0(0) = 0 & \\ f_0(n) = n - f_0(n-1) & \text{pour } n \geq 1 \end{cases}$$

Plus délicat : soustrayons !

Que dire de la fonction $f_0 : \mathbb{N} \to \mathbb{N}$ définie ainsi ?

$$\begin{cases} f_0(0) = 0 & \\ f_0(n) = n - f_0(n-1) & \text{pour } n \geq 1 \end{cases}$$

n	0	1	2	3	4	5	6	...
$f_0(n)$	0	1	1	2	2	3	3	...

Plus délicat : soustrayons !

Que dire de la fonction $f_0 : \mathbb{N} \to \mathbb{N}$ définie ainsi ?

$$\begin{cases} f_0(0) = 0 & \\ f_0(n) = n - f_0(n-1) & \text{pour } n \geq 1 \end{cases}$$

n	0	1	2	3	4	5	6	...
$f_0(n)$	0	1	1	2	2	3	3	...

En fait $f_0(n) = d(n+1) = \lfloor (n+1)/2 \rfloor$.

Plus délicat : soustrayons !

Que dire de la fonction $f_0 : \mathbb{N} \to \mathbb{N}$ définie ainsi ?

$$\begin{cases} f_0(0) = 0 & \\ f_0(n) = n - f_0(n-1) & \text{pour } n \geq 1 \end{cases}$$

n	0	1	2	3	4	5	6	...
$f_0(n)$	0	1	1	2	2	3	3	...

En fait $f_0(n) = d(n+1) = \lfloor (n+1)/2 \rfloor$.

Autre écriture possible $f_0(n) = \lceil n/2 \rceil$ (arrondi à l'entier supérieur).

Plus délicat : soustrayons !

Que dire de la fonction $f_0 : \mathbb{N} \to \mathbb{N}$ définie ainsi ?

$$\begin{cases} f_0(0) = 0 & \\ f_0(n) = n - f_0(n-1) & \text{pour } n \geq 1 \end{cases}$$

n	0	1	2	3	4	5	6	...
$f_0(n)$	0	1	1	2	2	3	3	...

En fait $f_0(n) = d(n+1) = \lfloor (n+1)/2 \rfloor$.

Autre écriture possible $f_0(n) = \lceil n/2 \rceil$ (arrondi à l'entier supérieur).

Justification : si $n - 1 \geq 1$, on "re-expanse" $f_0(n-1)$:

$$f_0(n) = n - ((n-1) - f_0(n-2)) = 1 + f_0(n-2)$$

La fonction G d'Hofstadter

Douglas Hofstadter définit au chap. 5 de “Gödel,Escher,Bach” :

$$\begin{cases} G(0) = 0 & \\ G(n) = n - G(G(n-1)) & \text{pour } n \geq 1 \end{cases}$$

La fonction G d'Hofstadter

Douglas Hofstadter définit au chap. 5 de “Gödel,Escher,Bach” :

$$\begin{cases} G(0) = 0 & \\ G(n) = n - G(G(n-1)) & \text{pour } n \geq 1 \end{cases}$$

n	0	1	2	3	4	5	6	7	8	9	...	13	...
$G(n)$	0	1	1	2	3	3	4	4	5	6	...	8	...

La fonction G d'Hofstadter

Douglas Hofstadter définit au chap. 5 de "Gödel,Escher,Bach" :

$$\begin{cases} G(0) = 0 & \\ G(n) = n - G(G(n-1)) & \text{pour } n \geq 1 \end{cases}$$

n	0	1	2	3	4	5	6	7	8	9	...	13	...
$G(n)$	0	1	1	2	3	3	4	4	5	6	...	8	...

Remarque: G semble transformer tout nombre de Fibonacci $F_i > 1$ en le précédent (on y reviendra):

$$G(F_i) = F_{i-1}$$

Quelques premières propriétés de G

- Existence et premier encadrement: $0 \leq G(n) \leq n$.
- $G(0) = 0$, $G(1) = 1$ puis $1 \leq G(n) < n$.
- A chaque étape, G "monte" (de $+1$) ou "stagne" ($+0$).
- Jamais deux $+0$ de suite.
- Jamais trois $+1$ de suite.

Quelques premières propriétés de G

- Existence et premier encadrement: $0 \leq G(n) \leq n$.
- $G(0) = 0$, $G(1) = 1$ puis $1 \leq G(n) < n$.
- A chaque étape, G "monte" (de $+1$) ou "stagne" ($+0$).
- Jamais deux $+0$ de suite.
- Jamais trois $+1$ de suite.

Ici en fait : $G(n) = \lfloor (n+1)/\varphi \rfloor$

où $\varphi = (1+\sqrt{5})/2 \approx 1.618$ est le nombre d'or

et $1/\varphi = \varphi - 1 \approx 0.618$.

Quelques premières propriétés de G

- Existence et premier encadrement: $0 \leq G(n) \leq n$.
- $G(0) = 0$, $G(1) = 1$ puis $1 \leq G(n) < n$.
- A chaque étape, G "monte" (de $+1$) ou "stagne" ($+0$).
- Jamais deux $+0$ de suite.
- Jamais trois $+1$ de suite.

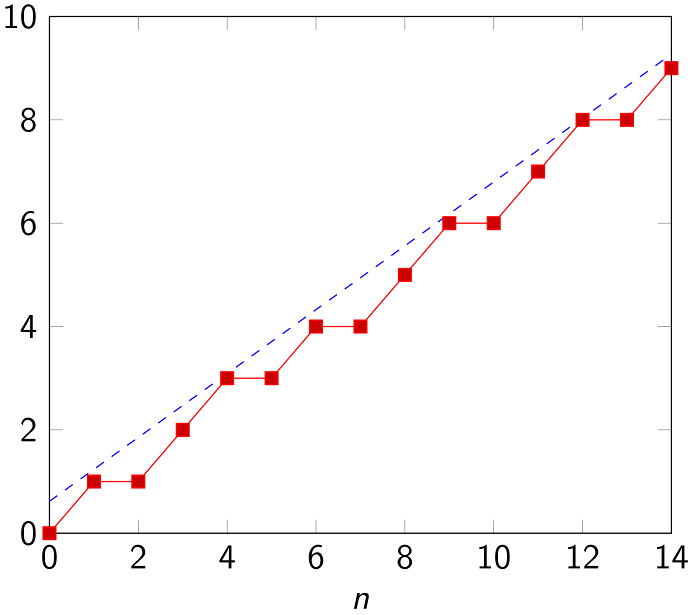
Ici en fait : $G(n) = \lfloor (n+1)/\varphi \rfloor$

où $\varphi = (1 + \sqrt{5})/2 \approx 1.618$ est le nombre d'or

et $1/\varphi = \varphi - 1 \approx 0.618$.

Arrondi de droite à pente irrationnelle : voir les *mots sturmiens*.

Tracés de $G(n)$ et $(n+1)/\varphi$:



Généralisons : la fonction H

Comme Hofstadter, varions le nombre d'appels imbriqués:

$$\begin{cases} H(0) = 0 & \\ H(n) = n - H(H(H(n-1)) & \text{pour } n \geq 1 \end{cases}$$

Généralisons : la fonction H

Comme Hofstadter, varions le nombre d'appels imbriqués:

$$\begin{cases} H(0) = 0 & \\ H(n) = n - H(H(H(n-1))) & \text{pour } n \geq 1 \end{cases}$$

Mêmes propriétés de base que G, sauf que:

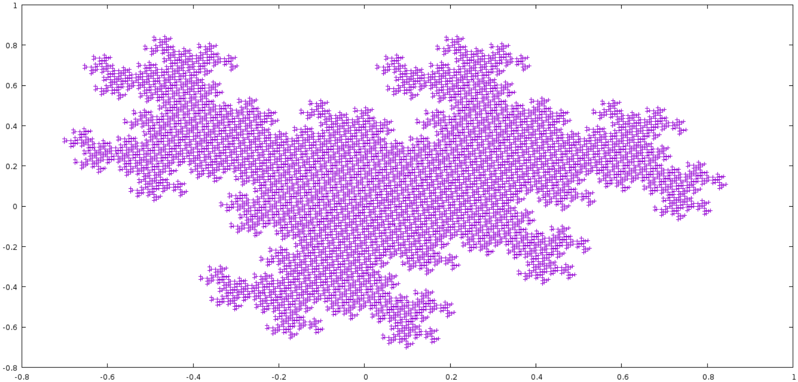
- Au plus trois $+1$ successifs
- Pas d'équation simple et exacte à base de $\lfloor\ \rfloor$
- Par contre: $H(n) = \lfloor \tau n \rfloor + 0$ ou 1

  avec $\tau \approx 0.6823$ racine réelle de $X^3 + X - 1$

Surprise !

Posons $\delta(n) = H(n) - \tau.n$

Affichage des points $(\delta(i), \delta(H(i))$ avec i=0..10000



Fractale de Rauzy et variantes

- Fractale précédente déjà connue, mais obtenue différemment.
- Variante de la fractale de Rauzy, via une autre variante de Fibonacci et un autre nombre de Pisot-Vijayaraghavan, mais l'étude est similaire.

Références:

- G. Rauzy, *Nombres algébriques et substitutions*, 1982.
- N. Pytheas Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics*, 2002

Généralisons encore : une famille de fonctions f_k

Notons $k+1$ le nombre d'appels récursifs souhaités. On définit:

$$\begin{cases} f_k(0) = 0 & \\ f_k(n) = n - \underbrace{f(\cdots f}_{k+1}(n-1)\cdots) & \text{pour } n \geq 1 \end{cases}$$

Généralisons encore : une famille de fonctions f_k

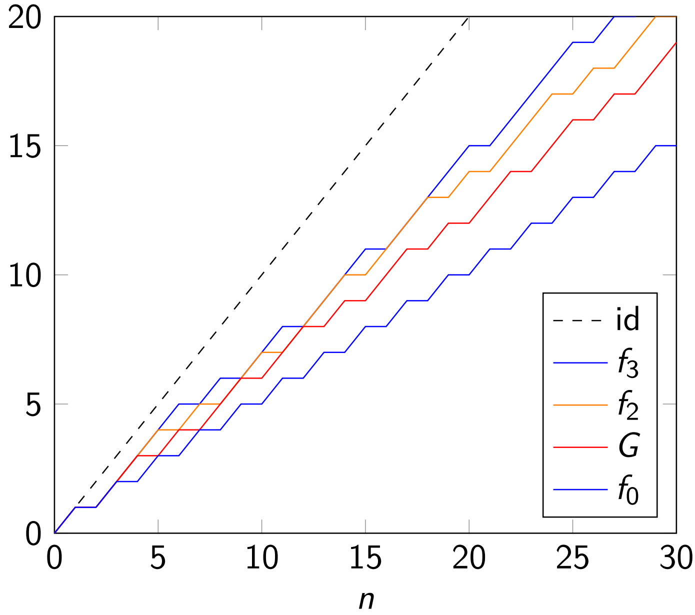
Notons $k+1$ le nombre d'appels récursifs souhaités. On définit:

$$\begin{cases} f_k(0) = 0 & \\ f_k(n) = n - \underbrace{f(\cdots f}_{k+1}(n-1)\cdots) & \text{pour } n \geq 1 \end{cases}$$

On retrouve les cas particuliers précédents :

- $f_0 = \lfloor (n+1)/2 \rfloor$ avec un seul appel récursif
- $f_1 = G$ avec deux appels récursifs
- $f_2 = H$ avec trois appels récursifs

Tracés



Premières propriétés de f_k

- Existence et premier encadrement: $0 \leq f_k(n) \leq n$
- $f_k(0) = 0$, $f_k(1) = 1$ puis $1 \leq f_k(n) < n$
- A chaque étape, f_k “monte” (de $+1$) ou “stagne” ($+0$).
- Jamais deux $+0$ de suite
- Au plus $k+1$ montées $+1$ de suite

Nota: pour $k \geq 2$, $f_k(n)$ n’a pas d’expression exacte via $\lfloor\ \rfloor$.

De la science toute récente !

$\boxed{\text{Théorème: pour tous } k \text{ et } n\text{, on a } f_k(n) \leq f_{k+1}(n)}$

Bref, f_0 est partout en dessous de G, qui est en dessous de H, etc

De la science toute récente !

$\boxed{\text{Théorème: pour tous } k \text{ et } n\text{, on a } f_k(n) \leq f_{k+1}(n)}$

Bref, f_0 est partout en dessous de G, qui est en dessous de H, etc

- Conjecturé en 2018.
- Coeur de la preuve par Shuo Li il y a 15 jours !
- Version certifiée sur machine cette semaine (preuve Coq) !
- Preuve par reformulation en un problème de mots infinis.

Partie 2 : Fibonacci généralisé et numération

Les nombres de Fibonacci

$$\begin{cases} F_0 & = 1 \\ F_1 & = 2 \\ F_{n+2} & = F_n + F_{n+1} \end{cases}$$

Les nombres de Fibonacci

$$\begin{cases} F_0 & = 1 \\ F_1 & = 2 \\ F_{n+2} & = F_n + F_{n+1} \end{cases}$$

(F_i) : 1 2 3 5 8 13 21 34 55 89 ...

NB: définition inhabituelle, pas de 0, un seul 1.

Théorème de Zeckendorf

Théorème (Zeckendorf): tout nombre entier peut s'écrire comme somme de nombres de Fibonacci tous différents et sans voisins. Cette décomposition est unique.

Par exemple: $17 = 13 + 3 + 1 = F_5 + F_2 + F_0$

On écrit alors parfois $17 = 100101_F$

1	=	1_F	7	=	1010_F
2	=	10_F	8	=	10000_F
3	=	100_F	9	=	10001_F
4	=	101_F	10	=	10010_F
5	=	1000_F	11	=	10100_F
6	=	1001_F	12	=	10101_F

G et Fibonacci

- $G(F_i) = F_{i-1}$ (avec la convention $F_{0-1} = F_0 = 1$)

G et Fibonacci

- $G(F_i) = F_{i-1}$ (avec la convention $F_{0-1} = F_0 = 1$)
- Et même, G décale les décompositions: $G(\Sigma F_i) = \Sigma F_{i-1}$

G et Fibonacci

- $G(F_i) = F_{i-1}$ (avec la convention $F_{0-1} = F_0 = 1$)
- Et même, G décale les décompositions: $G(\Sigma F_i) = \Sigma F_{i-1}$
- Propriété cruciale, preuve délicate
- Exemple: $G(17) = G(100101_F) = 10011_F = 11$
- Voisins possibles dans la décomposition obtenue, on peut la *renormaliser* ensuite, p.ex. $10011_F = 10100_F = 11$

Fibonacci généralisé

Soit k un entier naturel. On définit:

$$\begin{cases} A_n^k & = n+1 \qquad\qquad\qquad \text{pour } n \leq k \\ A_{n+1}^k & = A_n^k + A_{n-k}^k \qquad\qquad \text{pour } n \geq k \end{cases}$$

Fibonacci généralisé

Soit k un entier naturel. On définit:

$$\begin{cases} A_n^k & = n+1 \qquad\qquad\qquad \text{pour } n \leq k \\ A_{n+1}^k & = A_n^k + A_{n-k}^k \qquad\qquad \text{pour } n \geq k \end{cases}$$

- A^0 : 1 2 4 8 16 32 64 128 256 512 ... (Puissances de 2)
- A^1 : 1 2 3 5 8 13 21 34 55 89 ... (Fibonacci F_i)
- A^2 : 1 2 3 4 6 9 13 19 28 41 ... (Narayana's Cows)
- A^3 : 1 2 3 4 5 7 10 14 19 26 ...

Zeckendorf généralisé

Soit k un entier naturel.

Théorème (Zeckendorf): tout nombre entier peut s'écrire comme somme de nombres A_i^k dont les indices diffèrent tous d'au moins $k+1$. Cette décomposition est unique.

Théorème: f_k décale cette décomposition : $f_k(\Sigma A_i^k) = \Sigma A_{i-1}^k$ (toujours avec la convention $A_{0-1}^k = A_0^k = 1$)

Là encore, f_k dénormalise au passage certaines décompositions.

Important: f_k "stagne" en n lorsque n contient $A_0^k = 1$ dans sa décomposition.

Partie 3 : Lien avec des mots infinis

Une substitution de lettres

Soit k un entier naturel. On utilise $\mathcal{A} = [0..k]$ comme alphabet et on définit une *substitution* $\sigma_k : \mathcal{A} \to \mathcal{A}^*$ ainsi:

$$\begin{cases} \sigma_k(n) = (n+1) & \text{pour } n < k \\ \sigma_k(k) = k.0 & \end{cases}$$

Ceci engendre un mot infini m_k à partir de la lettre k (on parle de mot *morphique*)

Par exemple:

- $m_1 = 1011010110110...$ (dual du mot de Fibonacci)
- $m_2 = 2012202012012201202...$

Vision par blocs de lettres

Si l'on découpe m_k à chaque lettre k et que l'on marque la taille des blocs entre ces k, on réobtient m_i.

Exemple:

$$\begin{array}{lcccccc} m_2 & = & 201 & 2 & 20 & 201 & 201\ldots \\ & = & 2 & 0 & 1 & 2 & 2\ldots \end{array}$$

Vision par blocs de lettres

Si l'on découpe m_k à chaque lettre k et que l'on marque la taille des blocs entre ces k, on réobtient m_i.

Exemple:

$$\begin{array}{ccccccc} m_2 & = & 201 & 2 & 20 & 201 & 201\ldots \\ & = & 2 & 0 & 1 & 2 & 2\ldots \end{array}$$

A contrario, ceci donne une méthode d'expansion de m_k (c'est en fait la substitution $(\sigma_k)^k$).

Equation récursive alternative

m_k est la limite de $\sigma_k^n(k)$ quand $n \to \infty$

Mais aussi la limite de préfixes finis $M_{k,n}$ définis ainsi:

- $M_{k,n} = k.0...(n-1)$ pour $n \leq k$
- $M_{k,n+1} = M_{k,n}.M_{k,n-k}$ pour $k \leq n$

Equation récursive alternative

m_k est la limite de $\sigma_k^n(k)$ quand $n \to \infty$

Mais aussi la limite de préfixes finis $M_{k,n}$ définis ainsi:

- $M_{k,n} = k.0...(n-1)$ pour $n \leq k$
- $M_{k,n+1} = M_{k,n}.M_{k,n-k}$ pour $k \leq n$

Remarque : $|M_{k,n}| = A_n^k$

Lien avec f_k

Theorème : si l'on projette vers 1 chaque lettre non-nulle de m_k, on obtient le mot infini des montées et des plats de f_k

Autrement dit, f_k stagne là il y a des 0 dans m_k.

Et en cumulant: le nombre de 0 dans m_k parmi ses n premières lettres donne $n - f_k(n)$.

Partie 4 : Formalisation sur machine

Comment définir f_k en Coq ?

Voir https://github.com/letouzey/hofstadter_g

```
Fixpoint recf k p n :=
 match p, n with
 | S p, S n => S n - Nat.iter (S k) (recf k p) n
 | _, _ => 0
 end.

Definition f k n := recf k n n.

Lemma f_k_0 : forall k, f k 0 = 0.
Proof.
 intro. compute. reflexivity.
Qed.

Lemma f_pred : forall k n, f k n = n - Nat.iter (S k) (f k) (n-1).
Proof.
 ...
Qed.
```

Quelques énoncés prouvés en Coq

Voir https://github.com/letouzey/hofstadter_g

```
(* Action de f_k sur la décomposition de n en sommes de k-bonacci *)
Lemma f_decomp : forall k n, f k n = sumA k (map pred (decomp k n)).
Proof.
 ...
Qed.

(* Proximité entre H = f 2 et son approximation linéaire *)
Lemma h_natpart_or :
 forall n, h n = nat_part (tau*n) \/ h n = S (nat_part (tau*n)).
Proof.
 ...
Qed.

(* Monotonie de la famille de fonctions f_k *)
Theorem f_grows : forall k n, f k n <= f (S k) n.
Proof.
 ...
Qed.
```

Bilan

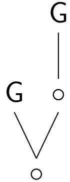
- Au tout début: une plage, un livre, du papier !
- Il reste encore des territoires à défricher, même d'approche élémentaire.
- Le calcul sur machine est essentiel pour affiner ses intuitions
- Des preuves rapidement trop longues/complexes pour être totalement fiabilisées par relecture humaine.
- Preuve sur machines : ardu, coûteux, mais faisable et gratifiant

Questions ?

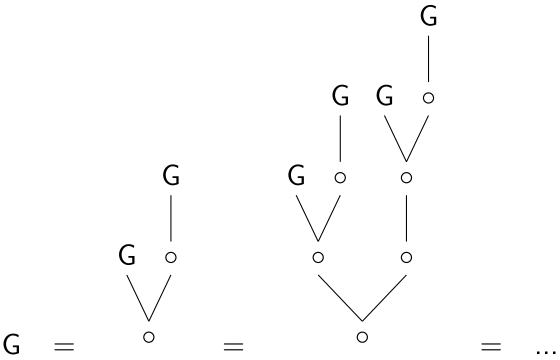
Partie 5 : G vu comme arbre infini

Un arbre infini rationnel

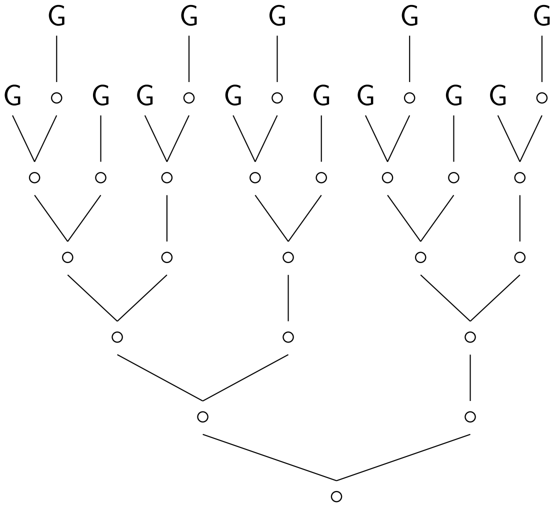
G =



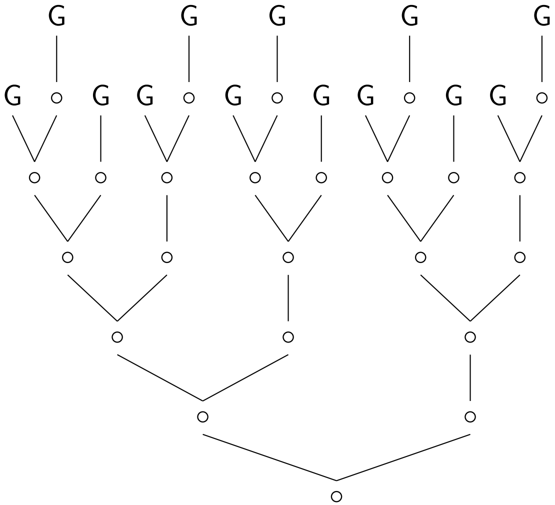
Un arbre infini rationnel



Un arbre infini rationnel



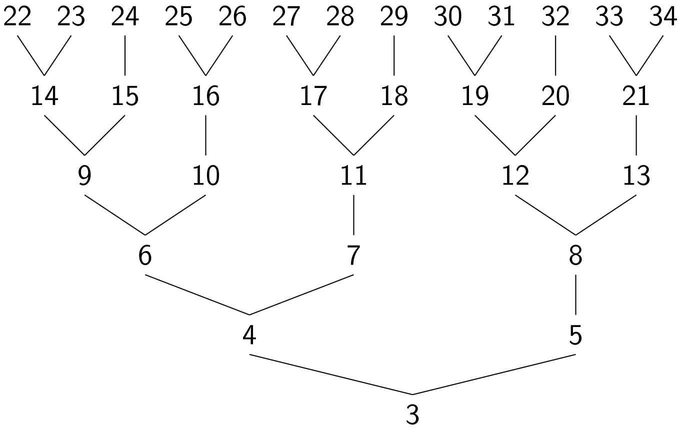
Un arbre infini rationnel



Combien de noeuds par niveau ?

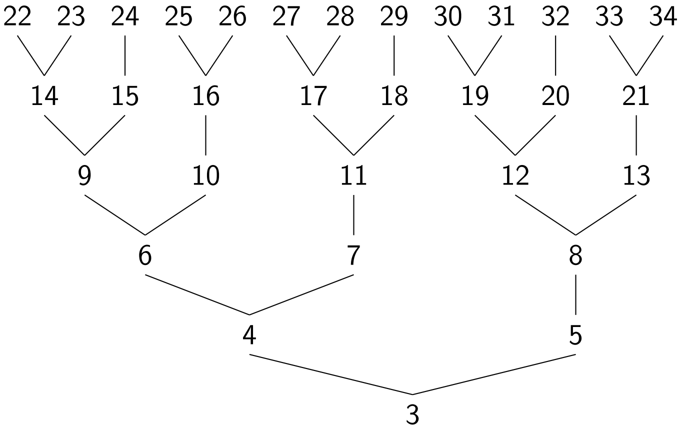
Numérotons les noeuds !

Parcours en largeur, de gauche à droite



Numérotons les noeuds !

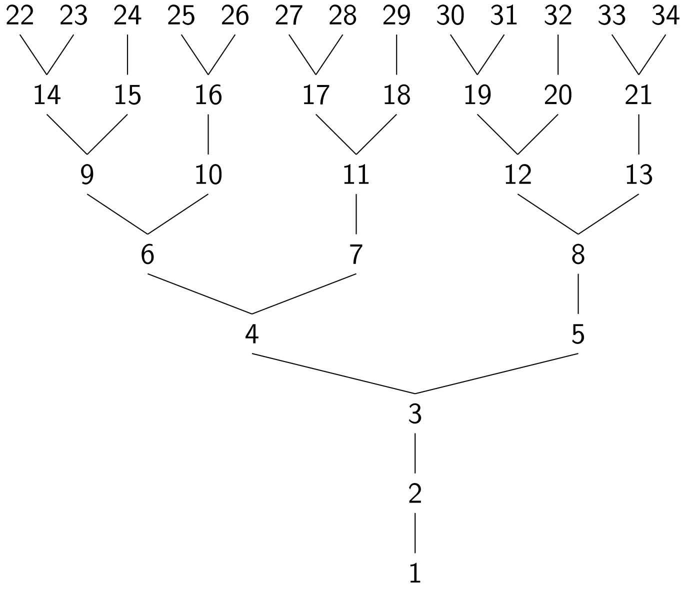
Parcours en largeur, de gauche à droite



Départ à 3 ? Pour faire apparaître les nombres de Fibonacci. . .

Théorème: le noeud n a $G(n)$ comme ancêtre.

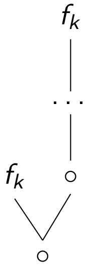
Ajout d'un tronc : l'arbre de G



Arbre généralisé

On allonge la branche de droite ($k + 1$ segments)

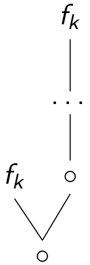
$f_k \quad =$



Arbre généralisé

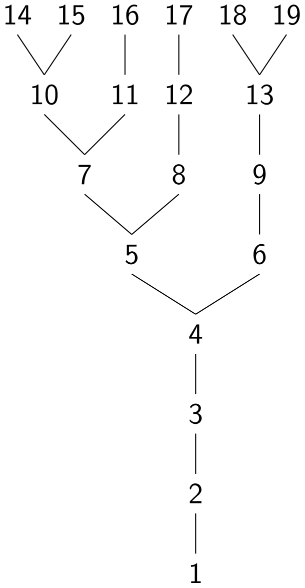
On allonge la branche de droite ($k + 1$ segments)

$f_k \quad =$



Et encore un tronc sur mesure (1 puis $k + 1$ segments)

Arbre pour f_2 (H de Hofstadter)



Arbre pour f_0