GENEALOGIE MATHEMATIQUE

Hervé Stève,
herve.steve@hotmail.fr

Kafemath du 8 novembre 2018

À la Coulée Douce, Paris 12ème

Introduction

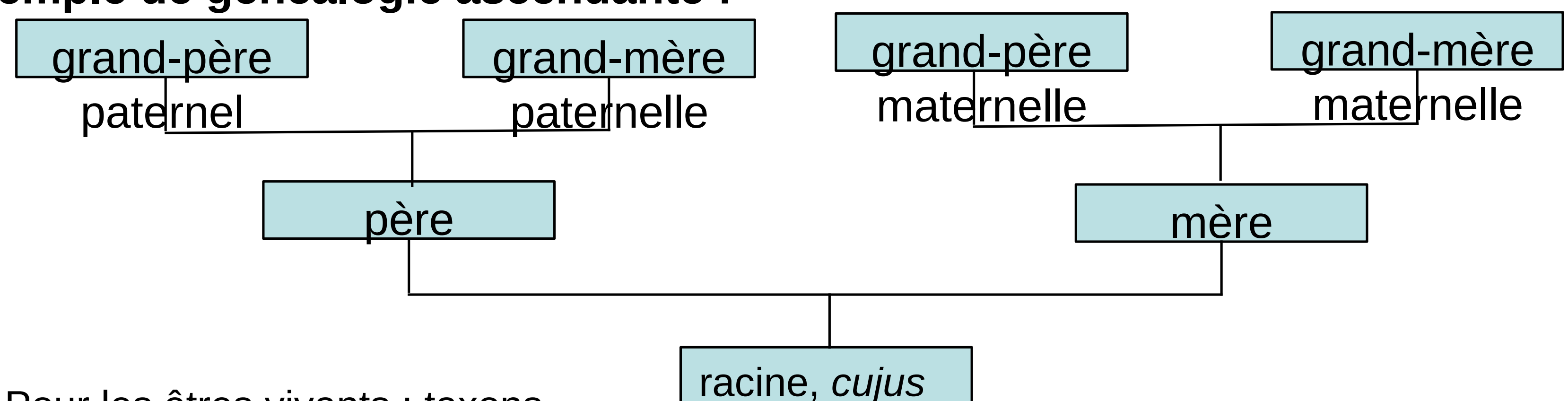
- **Arbres généalogiques**
- **Numérotation SOSA, implexe**
- **Arbres de Galton-Watson**
- **Généalogie des mathématiciens**
- **Généalogie des abeilles, des lapins**
- **Généalogie génétique**

Généalogie

Étude de générations d'individus : liste des membres d'une famille ayant une lien de filiation (ou parenté) … science apparentée à l'histoire

Arbre généalogique : individu « racine » vers des individus parents : « branches » et « feuilles » , généalogie ascendante (racine vers parents, grands-parents, …) ou bien descendante (racine vers les enfants, petits enfants, …)

Exemple de généalogie ascendante :



Pour les êtres vivants : taxons
Pour les animaux : pedigree
XIIème : Arbre biblique de Jessé, descendants jusqu'à Jésus …
XIVème : Arbre des rois francs de Bernard Gui …
Arbres en ligne sur la toile : MyHeritage, geneanet, roglo, …

Généalogie ascendante

Taille de l'arbre noté T(N) : avec N générations ou N-1 degré

Racine : degré 0 ou 1ère génération

Père, mère : degré 1 ou 2nde génération

Grands-parents : degré 2 ou 3ème génération …

Soit 2^{N-1} individus par génération

donc $T(N)=1+2+2^2+2^3+\ldots+2^{N-1}=2^N-1$ individus total

(formule $1+a+a^2+\ldots a^{n-1}=(a^n-1)/(a-1)$)

Si 30 ans par génération soit 17 générations jusqu'en 1539 (ordonnance de Villers-Cotterêts édicté par François 1er) et $T(17)=2^{16}-1=65\,535$ ancêtres théoriques !

En l'an 0, on aurait 60 générations soit T environ 1 milliard de milliards d'individus ! mais il y a les mariages consanguins de degrés + ou – élevés (endogamie) : Exemple) Charlemagne est 500 fois dans l'arbre de Saint-Louis !

Implexe : même ancêtre dans l'arbre

Indice d'implexe (N) : ancêtres réels / ancêtres théoriques = 2^{N-1}

Taux d'implexe (N) : % de doublons à la génération N

Numération de Sosa-Stradonitz

Généalogie ascendante
Jérôme de Sosa, généalogiste franciscain (1676)
Stéphane Kékulé von Stradonitz (1879) :
chimiste, ascendance des souverains et conjoints

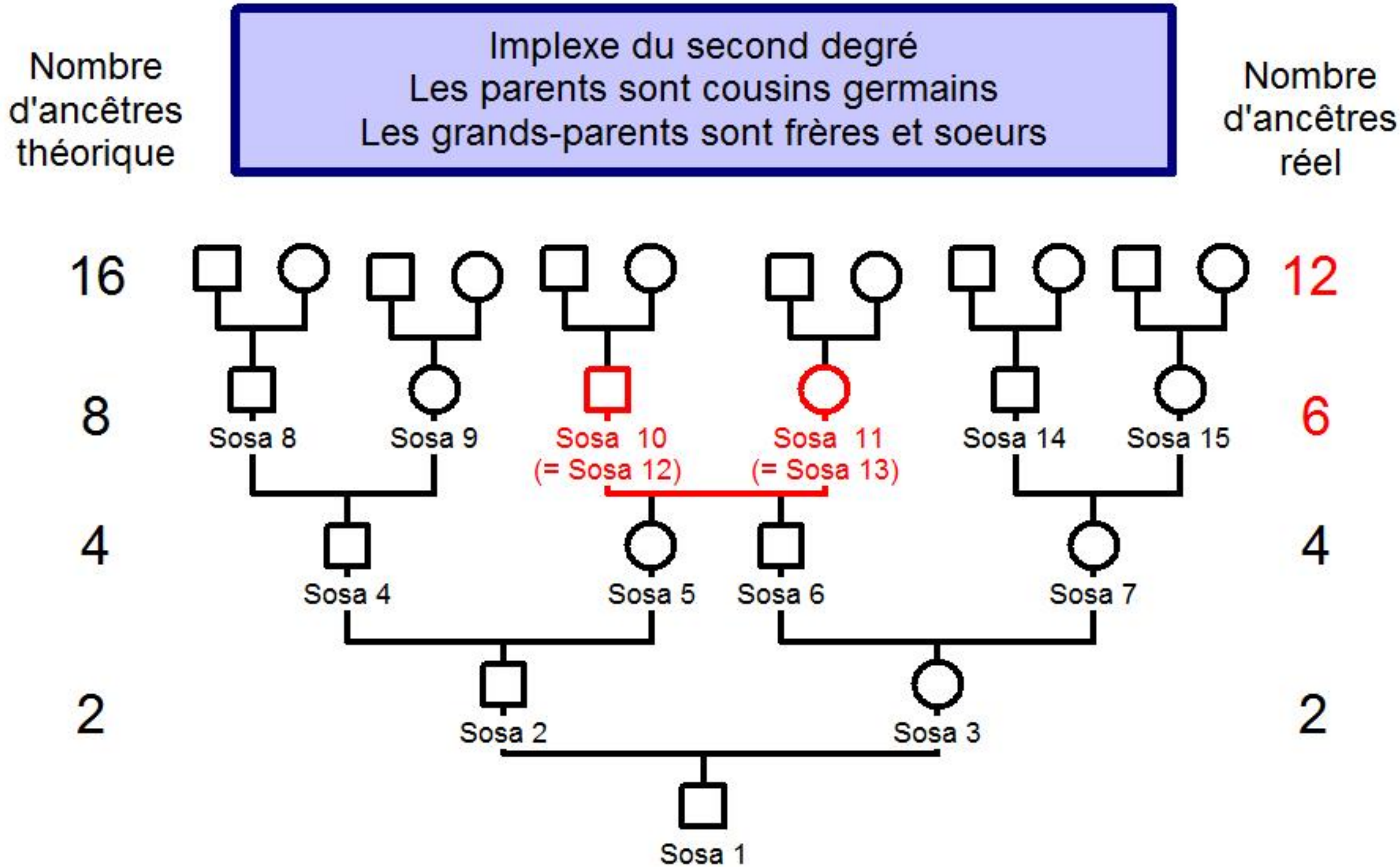

NOTICIA
DE LA GRAN CASA
DE LOS
MARQVESES
DE VILLAFRANCA,
Y SV PARENTESCO CON LAS MAYORES
DE EVROPA,
EN EL
ARBOL GENEALOGICO
DEL EXCELENTISSIMO SEÑOR
D. FADRIQVE
DE TOLEDO OSORIO
SEPTIMO MARQVES
DE ESTA CASA.
POR FRAY GERONIMO DE SOSA,

Utilisée dans les logiciels de généalogie (N° SOSA) :

0 : sujet	1 : racine							
1 :parents	2 : père de 1				3 : mère de 1			
2 :aïeuls	4 : père de 2		5 : mère de 2		6 : père de 3		7 : mère de 3	
3 : Bisaïeuls	8 : père de 4	9 : mère de 4	10 : mère de 5	11 : mère de 5	12 : père de 6	13 : mère de 6	14 : père de 7	15 : père de 7
4: Trisaïeuls	SOSA de 16 à 31							

- Degré d'ascendance = partie entière du logarithme en base 2 du SOSA
  ex) SOSA=13 : $\log_2(2^3)=3 < \log_2(13) < \log_2(2^4)=4$ => degré(SOSA=13)=3
- Degré 3 (trisaïeuls) : effectifs théoriques de $16=2^{3+1}$ mais l'effectif réel peut être différent : par l'implexe (doublons) ou bien par incomplétude (ancêtres inconnus)

Exemple d'implexe



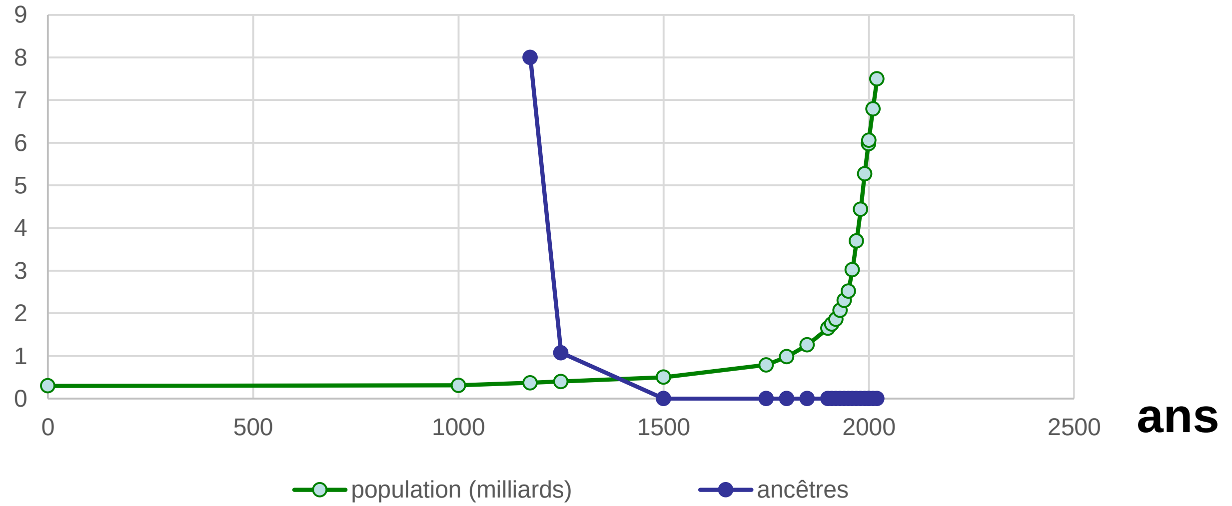
Indice d'implexe (4) = 6/8 = 0,75
Taux d'implexe (4) = (8-6) / 8 = 25%

Ancêtres et population

Population mondiale a une croissance exponentielle
Ancêtres théoriques a une croissance exponentielle inversée (doublement 25/30 ans)

Vers 1175 : ½ milliard d'habitants et 8 milliards d'ancêtres (25 générations) !
- chaque individu peut être 16 fois en moyenne notre ancêtre
- le croisement des arbres généalogiques freinent la croissance des ancêtres : implexe

Combien de générations avant un implexe en moyenne ?
- au plus 21 générations
- modèle maths : 7 à 8 générations (Stéphane Le Borgne)



Modèle coalescent de Kingman pour modéliser la généalogie d'une population : estimation de l'âge de l'ancêtre commun…(utilisé en génétique)

Probabilités

Que disent les probabilités ?
1 individu né vers 1175 doit être 8milliards/ ½ milliards =16 fois notre ancêtre en moyenne !

En 1180, Philippe Auguste était roi de France, il doit donc être un de nos ancêtre !
- Supposons qu'il ne soit pas un de nos ancêtres.
- Donc chacun de nos ancêtres (les 8 milliards) a une probabilité :

p_0=499 999 999/500 000 000~0,999 999 98 ne pas être Philippe Auguste

- Alors la probabilité que ce roi ne soit pas notre ancêtre est $p_0^{8milliards} \sim 10^{-7}$ très proche de zéro !

Philippe Auguste est donc notre ancêtre « en moyenne » …
pour Charlemagne cela semble presque certain !

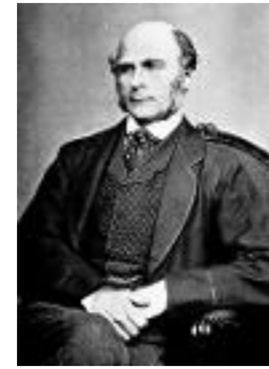
Et nous descendons aussi de Philippe Auguste !
- La probabilité que vos 2 parents de descendent pas du roi est de p_0^2
- Pour les 4 grands parents, on a $p_0^2 \times p_0^2$
- Pour les 8 bisaïeuls, on a $(p_0^2 \times p_0^2) \times (p_0^2 \times p_0^2) = p_0^8$ et ainsi de suite
- Donc la probabilité d'une descendance du roi à la génération 25 est

$p(25)=1-p_0^{T(25)}$ d'où $p(25)=1-2 \times 10^{-7} \sim 1$

Les descendants de Philippe Auguste croissent rapidement : la probabilité passe de quasi 0 (de 1180 à 1800) à 1 rapidement vers 1800/1900. Dans la base universelle Roglo, il y a 929 691 descendants (34 générations) sur 7 500 000 individus !

! nous aurions donc tous du sang bleu !

Générateurs d'arbres généalogiques



- **Modèle malthusien :** croissance exponentielle

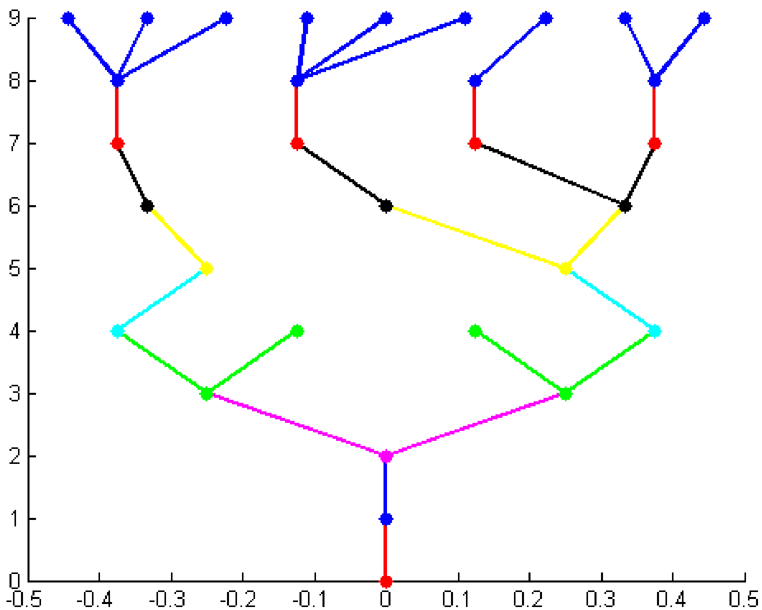
- **Arbre de Galton-Watson** : modèle d'extinction des individus (F. Galton : lords anglais)

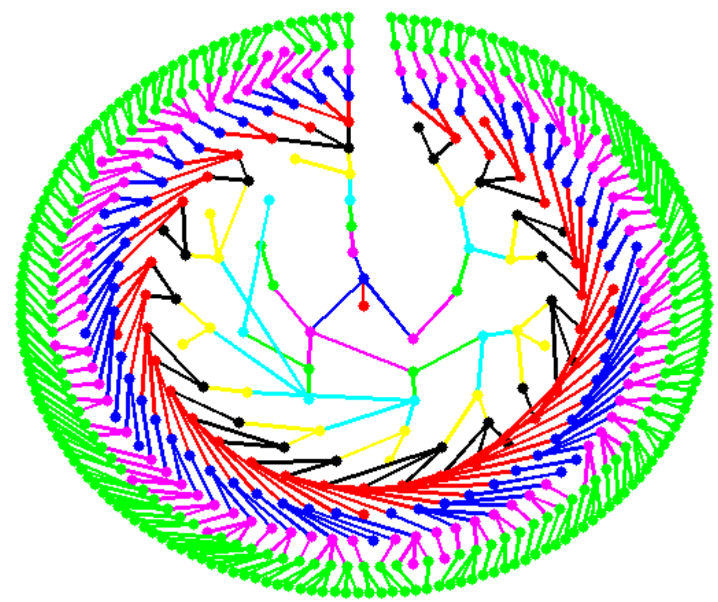
On se donne une distribution de probabilités p_i d'avoir i enfants pour chaque individu
exemple1) $p_0=1/8=p_3$; $p_1=p_2=3/8$ et $p_{i>3}=0$; somme $p_{i=0..infini} = 1$

n(1)=1 … n(10)=9

n(1)=1 …n(12)=183

N-1





Arbre linéaire

Arbre circulaire

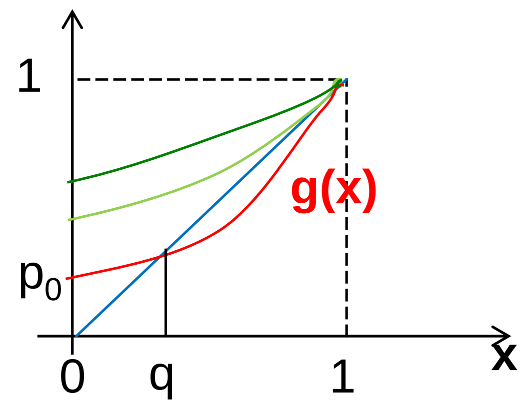
Arbre de Galton-Watson

- **Moyenne ou espérance :**

E = m = nombre moyen d'enfants / individu = $0p_0 + 1p_1 + 2p_2 + 3p_3 + \ldots + np_n + \ldots$
n(N) nbre enfants génération N : $E(n(N)) = m^N$ par récurrence
$E(T) = 1 + m + m^2 + \ldots + m^N + \ldots = 1 / (1-m)$ si $m < 1$: cas extinction sûre
Si m=1 ou m>1 , E(T) diverge vers l'infinie mais extinction reste possible !

- **Fonction génératrice** :

g(x) pour x dans [0,1], probabilité que tous les descendants vérifie une propriété de probabilité x
$g(x) = p_0 + x\, p_1 + x^2p_2 + \ldots + x^np_n + \ldots$: $g(0) = p_0$; $g(1) = 1$
dérivée $g'(x) = p_1 + 2xp_2 + \ldots + nx^{n-1}p_n + \ldots \geq 0$ donc g(x) croissante
dérivée seconde $g''(x) \geq 0$ donc g(x) convexe



Quelle est la probabilité q d'extinction ? q=g(q)
on remarque que g'(1) = m
si m<1 on a q=1 : extinction certaine
si m>1 on a q<1 (q=1 est éliminé)
calcul de q (méthode point fixe) : $q=\ldots g(\ldots g(g(p_0)))=\lim g^k(p_0)$ quand k tend vers l'infini
si m=1 : cas critique, E(T) infini
si g''(x)>0 alors q=1 : extinction certaine
si g''(x)=0 alors $g(x)= p_0 + x\, p_1$, comme $g'(1)=1= p_1$ alors $p_0=0$: un seul descendant par génération donc q=0

Probabilités de survie

Exemple1) cas d'école

$p_0 = 1/8 = p_3$; $p_1 = p_2 = 3/8$ et $p_{i>3}=0$

$E = m = 0x1/8+1x3/8+2x3/8+3x1/8 = 12/8 = 1{,}5 > 1$

$g(x) = (1 + 3x + 3x^2 + x^3)/8 = (1+x)^3/8$

Quelle est la probabilité d'extinction ?

$g(q) = q<1$ conduit à q solution de $0 = 1 – 5q + 3 q^2 + q^3 = (1-q)(1+4q-q^2)$

1 solution $q = \sqrt{5}-2 \approx 23{,}6\%$ 1 individu / 4 n'a pas de descendance…

Exemple2) population française en 1960

$p_0=10\%$, $p_1=18\%$, $p_2=40\%$, $p_3=22\%$, $p_4=7\%$, $p_5=3\%$, $p_{>5}=0\%$

$E = m = (0x10+1x18+2x40+3x22+4x7+5x3) \approx 2{,}07 > 1$

$g(x) = 0{,}1 + 0{,}18x + 0{,}4x^2 + 0{,}22x^3 + 0{,}07x^4 + 0{,}03x^5$

$g(q) = q<1$: on trouve $q \approx 13{,}09\%$ 1 individu / 8 n'a pas de descendance…

Problème de Francis Galton (1873) : disparition des noms de famille des lords anglais
la transmission est de père en fils donc $m=2{,}07/2=1{,}035$: proche du cas critique
Soit la distribution p_n : probabilité qu'un individu ait n fils
$p_0=32{,}28\%$, $p_1=39{,}47\%$, $p_2=21{,}81\%$, $p_3=5{,}44\%$, $p_4=0{,}40\%$, $p_5=0{,}09\%$, $p_{>5}=0\%$
Puis on calcule la probabilité d'extinction : $q \approx 92\%$

En conclusion, la probabilité que la descendance d'une famille survive est forte (86%) mais que le nom de famille survive est faible (8%)

Généalogies des mathématiciens

Site : **Mathematics Genealogy Project** www.genealogy.math.ndsu.nodak.edu

Avec T=234 346 (15/09/2018)

Leonard Euler (PhD en 1726) : son directeur de thèse était Johann Bernoulli et à eu 6 étudiants et au moins 110 195 descendants !

Cédric Villani (1998) < Pierre Louis Lions (1979) < Haïm Brézis (1972) < Gustave Choquet & Jacques-Louis Lions *grand-pair* de son fils Pierre Louis

Hervé Stève (1988) < Alain Dervieux (1981) < Roland Glowinski (1970) < Jacques-Louis Lions (1954) < Laurent Schwartz (1943) < Georges Valiron (1914) < Emile Borel (1893) < Gaston Darboux (1866) < Michel Chasles (1814) < Siméon Denis Poisson (1800) < Joseph Louis Lagrange & Pierre-Simon Laplace étudiant de Jean Le Rond d'Alembert

Joseph Louis Lagrange (1754) < Leonhard Euler (1726) < Johann Bernoulli (1690) < Jacob Bernoulli & Nikolaus Eglinger

Jacob Bernoulli (1676) < Nicolas Malebranche (1672) < Gottfried Wilhelm Leibniz (1666) < Jakob Thomasius (1643) < Friedrich Leibniz (1622) < inconnu

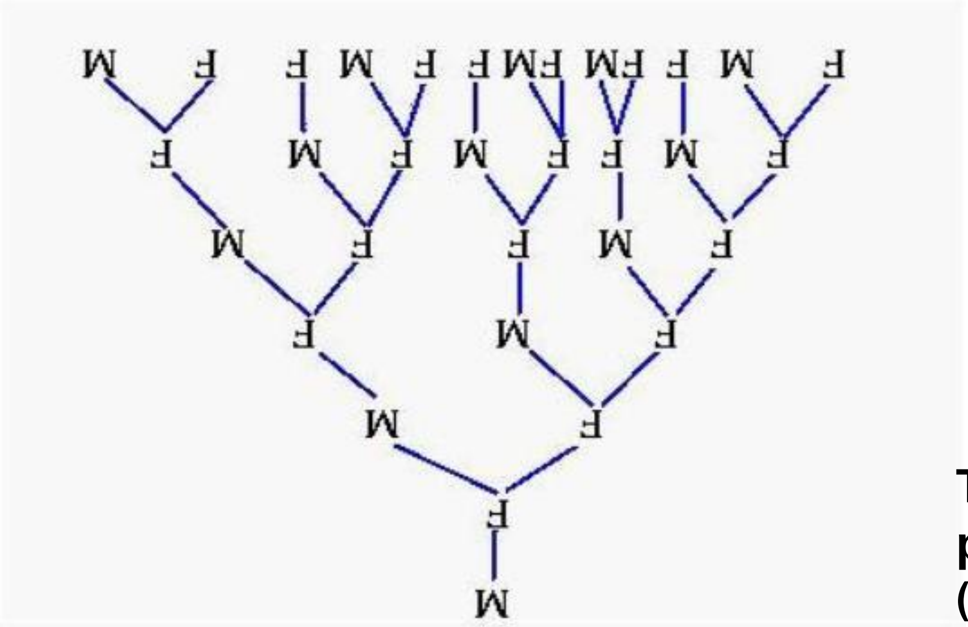
Par Nikolaus Eglinger , on remonte aux mathématiciens arabes …

Généalogie des abeilles

Familles des Hyménoptères (guêpes, fourmis, …)

- ✓ Un mâle M provient d’une femelle F (œuf non fécondé)
- ✓ Une femelle F d’un mâle M et d’une femelle F(œuf fécondé)

Taille de l’arbre en partant d’un mâle M : T(1)=1(M), T(2)=1(F), T(3)=2(F+M), T(4)=3(2F+M), T(5)=5 (3F+2M), T(6)=8 (5F+3M), T(7)=13 (8F+5M), …
T(N)=T(N-1)+T(N-2) **suite de Fibonacci** avec T(N-1) nombre de F et T(N-2) nombre de M



Leonardo Fibonacci (1175-1250)

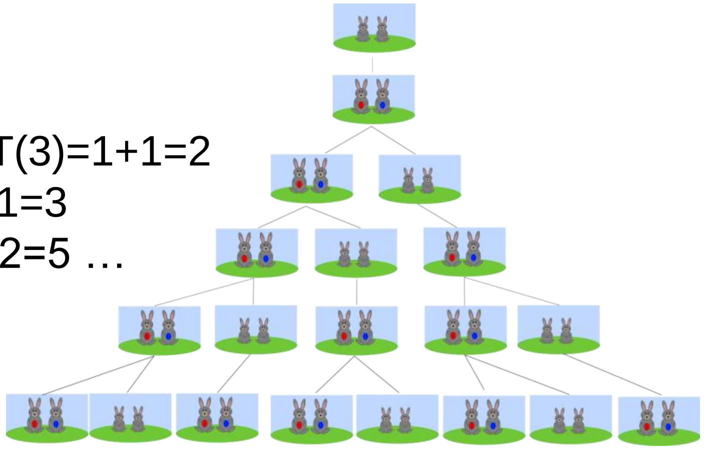
T(N)/T(N-1) tend vers le nombre d’Or proche de 1,618… solution de $x^2-x-1=0$
(1,618039… x 1,618039… = 2,618039… !)

Généalogie des lapins

Fibonacci : liber Abaci en 1202, croissance d'une population de lapins
Chaque mois, un couple de lapins donnent naissance à un couple de lapins : progression exponentielle de couples adultes … mais 1 lapereau met 1 mois pour devenir adulte et donc procréer :

- En janvier, on a un jeune couple de lapins $T(1)=1$
- En février, ce couple met bat $T(2)=1$
- En mars, on a 2 couples adultes et le 1er couple met bat $T(3)=1+1=2$
- En avril, on a 3 couples et 2 couples mettent bat $T(4)=2+1=3$
- En mai, on a 5 couples et 3 couples mettent bat $T(5)=3+2=5$ …

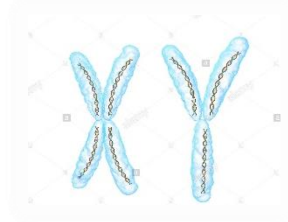
A nouveau la **suite de Fibonacci :** $T(N) = T(N-1) + T(N-2)$
($T(N-2)$ couples devenant adulte)



Mais les lapins ne vivent que 3 mois :
$T(1)=T(2)=T(3)=1$, $T(4)=T(5)=2$, $T(6)=3$, $T(7)=4$, $T(8)=5$, $T(9)=7$, $T(10)=9$, $T(11)=12$, …
on obtient **la suite de Padovan** : $T(N+1) = T(N-1) + T(N-2)$
avec $T(N+1)/T(N)$ qui tend vers 1,324… solution réelle de $x^3-x-1=0$

Les lapins finissent par manquer de nourriture :
modèle avec la suite logistique : $T(N) = r\,T(N-1)\,(1-T(N-1)/K)$
convergence ($r<3$), régime périodique, chaos, divergence

Généalogie génétique



- La généalogie génétique nécessite l'usage de tests ADN qui mesurent le niveau de rapports génétiques entre des individus. Les gènes se transmettent entre générations, des comparaisons génétiques permettent d'établir <u>un degré de parenté</u> plus ou moins proche entre individus
- Georges Darwin fils de Charles. En 1875, il trouva entre 2,25 % et 4,5 % de mariages entre cousins germains dans la population de la Grande-Bretagne suivant les classes sociales
- A partir de 1990 : travaux sur le génome, carte du chromosome Y qui se transmet de père en fils
- En 2001 : 7 groupes humains majeurs ancêtres des Européens
- Depuis 2010 : tests en vente aux US

    Test ligne paternelle ADN-Y (récente ou ancienne)

    Test ligne maternelle ADN-mitochondrial (héritage par l'ovule)

Généa…logique

- **Marcel** est le père de **Caroline** et l'oncle de **Patrick**.
- **Jacques** est le fils de **Solange** et le cousin de **Jean**.
- **Raymond** et **Marcel** sont les oncles de **Patrick**.
- **Irène** est la belle sœur **d'Edith** et la mère de **Pierre**.
- **Solange** est la bru de **Marie** et la mère de **Valérie**.
- **Brigitte** et **Pierre** sont jumeaux.
- **Françoise** est la cousine de **Valérie** et la fille de **Raymond**.
- **Jean** est le petit fils de **Charles** et le neveu **d'Irène**.
- **Patrick** est le frère de **Paul** et le neveu **d'Irène**.
- **Edith** est la belle sœur **d'André** et la mère de **Jean**.
- **Marcel** et **André** sont deux frères.
- **Irène** est la fille de **Marie** et la mère de **Françoise**.
- **Paul** est le cousin de **Jean**

Généa…logique (solution)