Random Trees and Genealogies

Bernard Derrida
Laboratoire de Physique Statistique,
Ecole Normale Supérieure, Paris

L. Peliti (Naples)

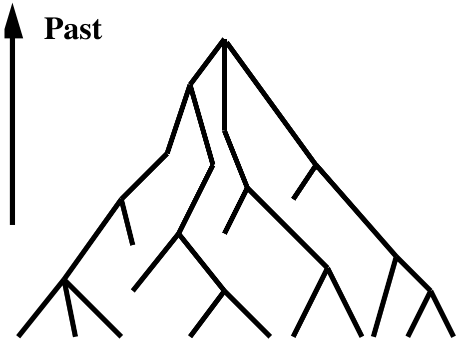
S. Manrubia (Madrid)
D. Zanette (Bariloche Argentina)

E. Brunet (E.N.S. Paris)
D. Simon (E.N.S. Paris)

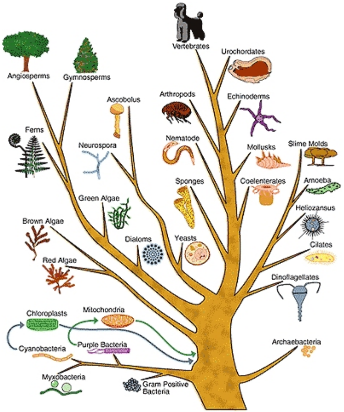
Jerusalem September 2008

OUTLINE

Asexual Reproduction



phylogeny



family names

bacteria

languages ...

Sexual Reproduction

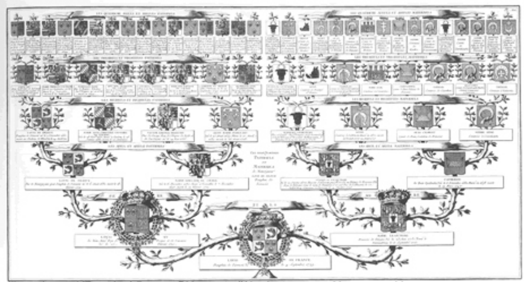
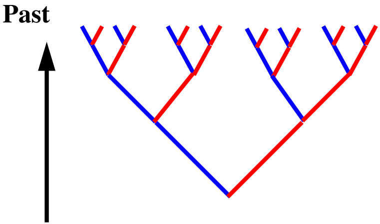
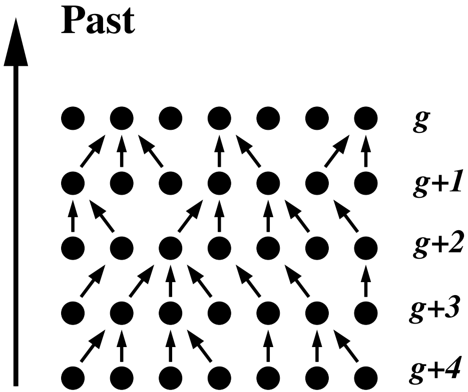


Illustration : *L'Encyclopédie,* par Diderot et d'Alembert - www.blason-armoiries.org



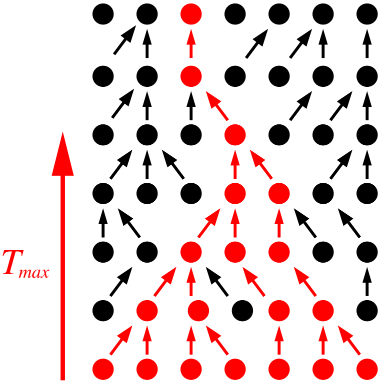
Wright-Fisher model (1930-1931)

- One parent model (asexual reproduction)
- Population of fixed size N
- Each individual has its parent chosen at random in the previous generation (neutrality)



Coalescence times:

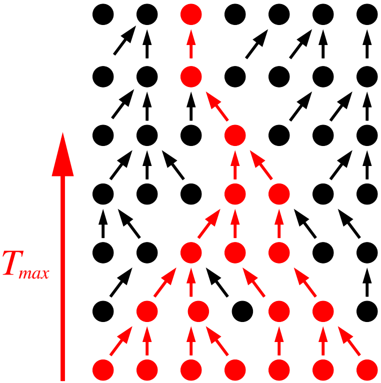
Age of the most recent common ancestor T_{max}



T_{max} depends on g

Coalescence times:

Age of the most recent common ancestor T_{max}



T_{max} depends on g

$$\langle T_{max} \rangle \simeq 2N$$

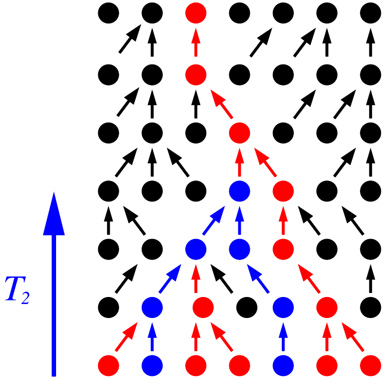
$$\langle {T_{max}}^2 \rangle - \langle T_{max} \rangle^2 \simeq C \; N^2$$

where $C = \frac{4\pi^2}{3} - 12 \simeq 1.6$

T_{max} is a non self-averaging quantity

Coalescence times:

Age T_2



The time T_2 depends on g and on the pair of individuals

Its average $\overline{T}_2$ over the population still depends on g

$$\langle T_2 \rangle \simeq N$$

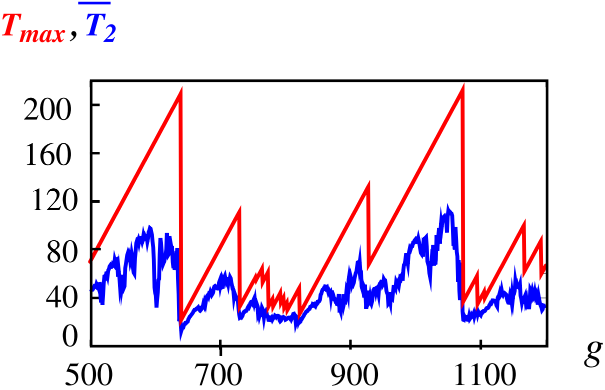
$$\langle {T_2}^2 \rangle - \langle T_2 \rangle^2 \simeq N^2$$

Two sorts of averages: $\begin{cases} \text{over the population} \\ \text{over } g \end{cases}$

Evolution of T_{max} and $\overline{T}_2$

T_{max} = age of the most recent common ancestor

$\overline{T}_2$ = average over the population of T_2

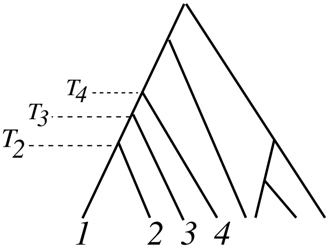


Serva 2005
Simon D. 2006

Coalescence times:

Age T_p

T_p = age of the most recent common ancestor
of p individuals chosen at random



$$\langle T_p \rangle \simeq \frac{2(p-1)}{p} N$$

(e^{-T_2} plays a role very similar to overlaps in spin-glasses)

Kingman's coalescent

Kingman 1982

$Q(k,k')$ = probability that k individuals have k' ancestors at the previous generation

$$Q(k,k) = \left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{k-1}{N}\right)$$

$$Q(k,k') \simeq \begin{cases} 1-\frac{k(k-1)}{2N} & \text{for } k'=k \\ \frac{k(k-1)}{2N} & \text{for } k'=k-1 \\ O\left(\frac{1}{N^2}\right) & \text{for } k'<k-1 \end{cases}$$

Therefore

$$p \underset{t_p}{\longrightarrow} p-1 \underset{t_{p-1}}{\longrightarrow} \cdots 2 \underset{t_2}{\longrightarrow} 1$$

and

$$\boxed{T_p = t_p + t_{p-1} + \ldots t_2}$$

The times t_p are independent and have exponential distibutions with $\langle t_p \rangle = 2N/p(p-1)$

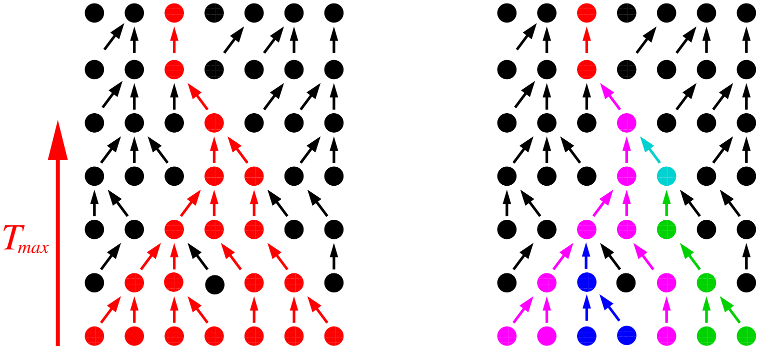
Beyond mean-field models

- One parent model
- Population of fixed size N
- Individuals are on the sites of a cube of $N = L^d$ lattice sites
- Each individual has its parent chosen at random in the previous generation in its neighborhood

Brunet D. Simon 2008

	$d \geq 2$	$d = 1$
$\frac{T_3}{T_2}$	$\frac{4}{3}$	$\frac{7}{5}$
$\frac{T_4}{T_2}$	$\frac{3}{2}$	$\frac{8}{5}$
$\frac{T_p}{T_2}$	$\frac{2(p-1)}{p}$	$\frac{2(p-1)(p+4)}{(p+1)(p+2)}$
$T_p \sim$	N	N^2

Genetic diversities



The sampling theory of selectively neutral alleles Ewens 1972

Summary on neutral models of asexual reproduction

- The coalescence times T_p are not selfaveraging
- Kingman's coalescent
- $T_p \sim N$
- Universal ratios $\frac{T_p}{T_2}$
- More is different

Summary on neutral models of asexual reproduction

- The coalescence times T_p are not selfaveraging
- Kingman's coalescent
- $T_p \sim N$
- Universal ratios $\frac{T_p}{T_2}$
- ~~More is different~~
- $d = 1$ is different (non-mean field)

Tomorrow's lecture: Evolution with **selection**

	neutral	selection
T_3/T_2	$4/3$	$5/4$
T_4/T_2	$3/2$	$25/18$
T_p/T_2	$2(p-1)/p$	long expression
Trees	Kingman 82	Parisi (79-80) Bolthausen-Sznitman (98)
$T_p \sim$	N	$(\log N)^3$

Sexual Reproduction

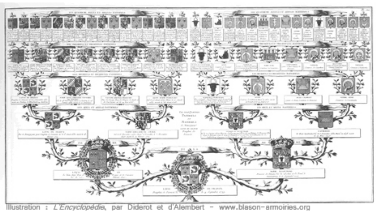
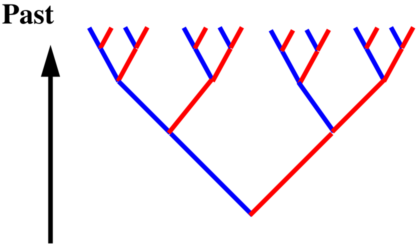


Illustration : *L'Encyclopédie*, par Diderot et d'Alembert – www.blason-armoiries.org

2 parents
4 grand parents
...
2^g ancestors g generations ago

At Roman times: $g \sim 76$

$$2^g \sim 10^{23}$$

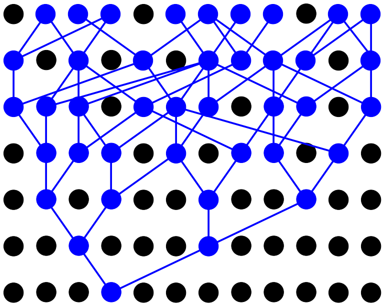


$$10^{23} \sim \frac{\text{mass of earth}}{\text{mass of man (of my size)}}$$

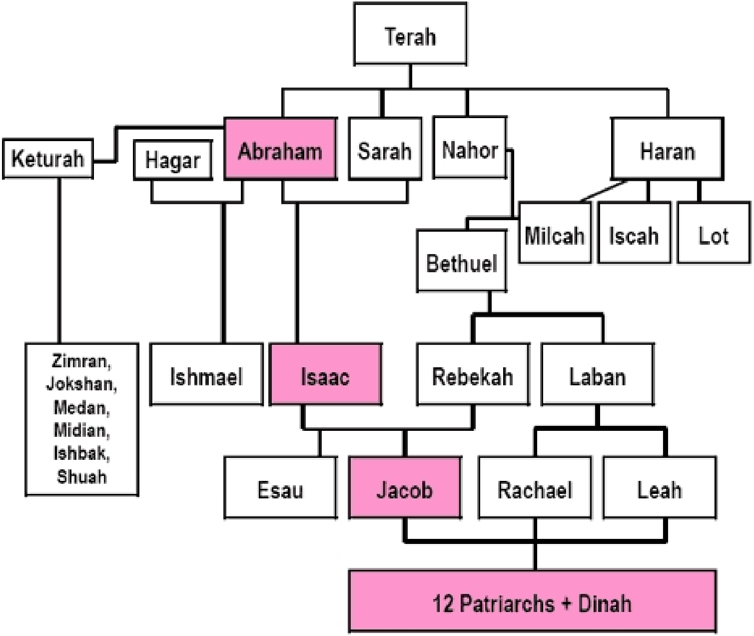
$\sim$ Avogadro's number

Neutral model

- Two parent model (sexual reproduction)
- Population of fixed size N
- Each individual has its two parents chosen at random in the previous generation (neutrality)

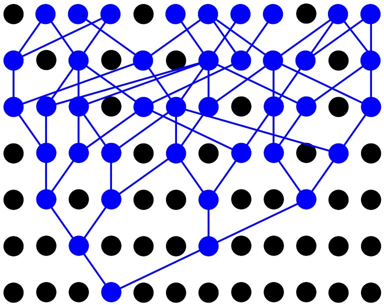


Abraham's Family Tree



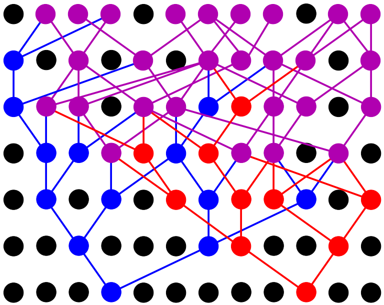
Neutral model

- Two parent model (sexual reproduction)
- Population of fixed size N
- Each individual has its two parents chosen at random in the previous generation (neutrality)



Neutral model

- Two parent model (sexual reproduction)
- Population of fixed size N
- Each individual has its two parents chosen at random in the previous generation (neutrality)



In the long term:

$$\begin{cases} \text{About} \sim .203N \text{ have no long term descendance} \\ \text{About} \sim .797N \text{ become ancestors of every body} \end{cases}$$

Chang 1999

Ages of ancestors: $N \sim 5 \times 10^9$

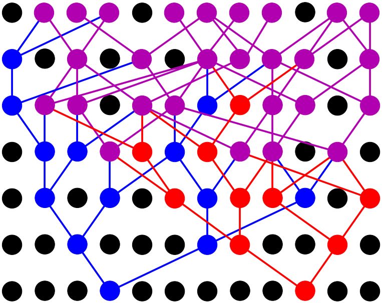
- Number of generations to at least one common ancestor of the whole population

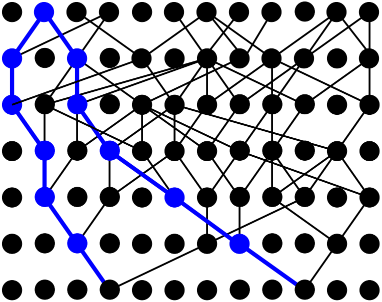
$$1.44 \log N \simeq 32$$

- Number of generations for all individuals to have all their ancestors in common

$$2.55 \log N \simeq 57$$

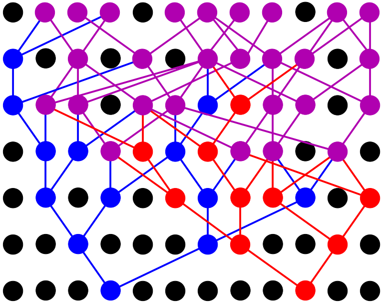
Neutral model

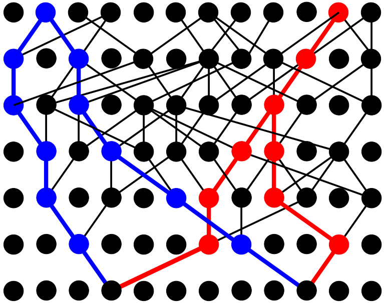




Abraham

Neutral model





Abraham

Lucy

An artist's view