The MRCA process in an evolving coalescent

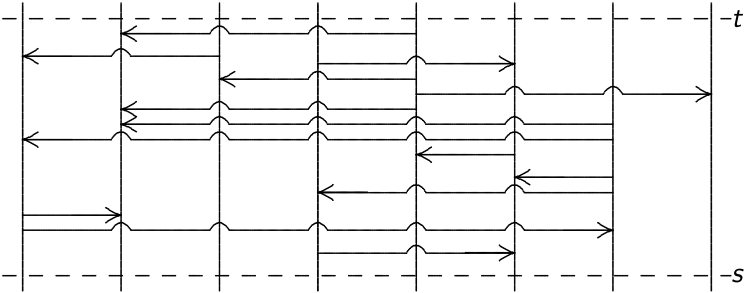
Peter Pfaffelhuber

University of Munich,
Joint work with Anton Wakolbinger

March 14th, 2006

The Moran model

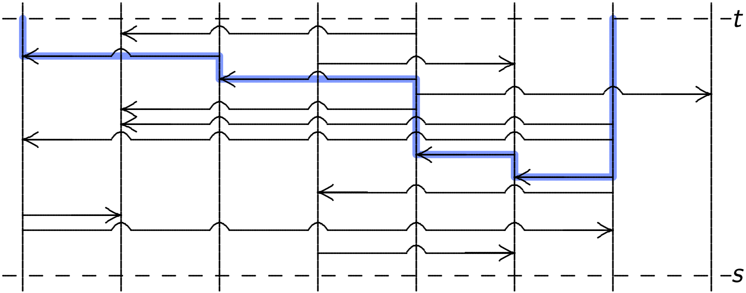
- A population consists of N individuals
- Each pair of individuals *resamples* at rate 1
- Resampling means: one individual dies, the other reproduces



The Moran model

- A population consists of N individuals
- Each pair of individuals *resamples* at rate 1
- Resampling means: one individual dies, the other reproduces

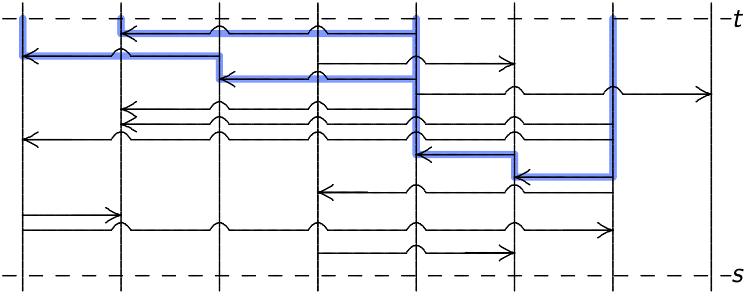
Ancestral lineages coalesce



The Moran model

- A population consists of N individuals
- Each pair of individuals *resamples* at rate 1
- Resampling means: one individual dies, the other reproduces

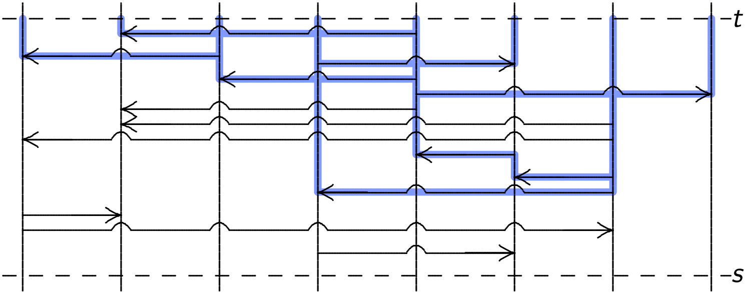
Ancestral lineages coalesce



The Moran model

- A population consists of N individuals
- Each pair of individuals *resamples* at rate 1
- Resampling means: one individual dies, the other reproduces

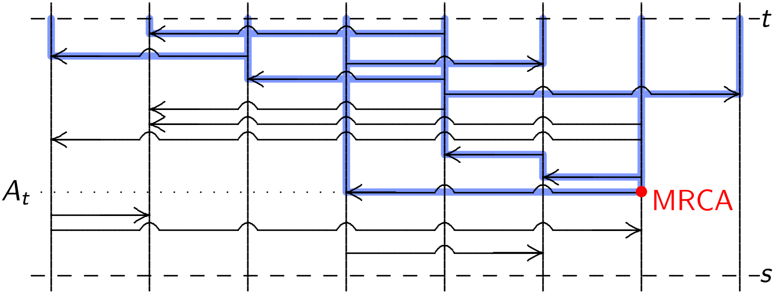
Ancestral lineages coalesce



The Moran model

- A population consists of N individuals
- Each pair of individuals *resamples* at rate 1
- Resampling means: one individual dies, the other reproduces

Ancestral lineages coalesce



The genealogy: Kingman's coalescent

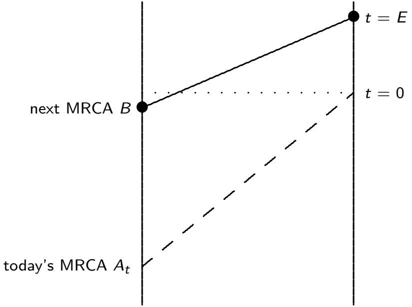
- A population of constant size evolves
- Genealogies are given by a coalescent process $\mathcal{K}$
- Coalescence rate is $\binom{n}{2}$.

When will there be a new MRCA? When will it have lived?

First observation: When one of the two oldest families dies out, the MRCA changes!

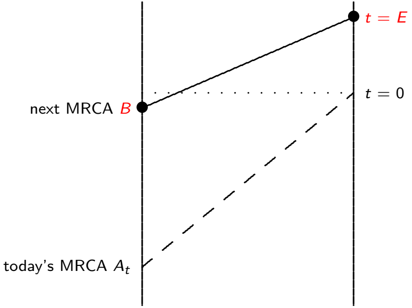
The MRCA-process

- The population at time t has its MRCA at time A_t
- (A_t) ist piecewise constant
- The population at $t = 0$ is in equilibrium. The next jump of the MRCA is at time E and it jumps to B



The MRCA-process

- The population at time t has its MRCA at time A_t
- (A_t) ist piecewise constant
- The population at $t = 0$ is in equilibrium. The next jump of the MRCA is at time E and it jumps to B



What is the joint distribution of E and B?

Exponential waiting times

When does the MRCA change?

- The fraction of the two oldest families is uniformly distributed
- If the MRCA does not jump until time t, this fraction is again uniform

$\Rightarrow$ The waiting time E is memoryless, hence the times $\{E\}$ form a Poisson process

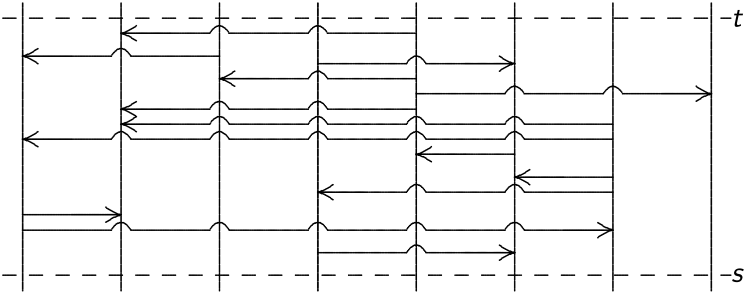
When does it live?

- The occurrence of MRCAs to come is also memoryless

$\Rightarrow$ The times $\{B\}$ form a Poisson-process

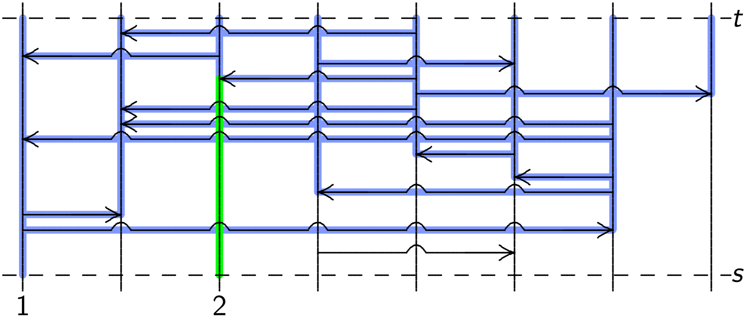
Persistence

- Look forwards in time
- Order individuals according to length of their ine of ascent



Persistence

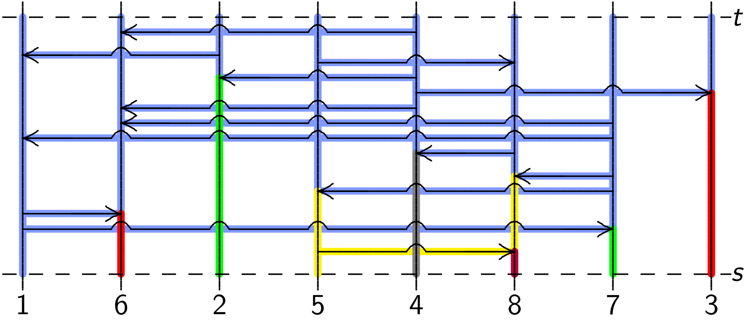
- We can reorder the Moran model by the level of *persistence*
- Individuals at a lower level live longer



Persistence

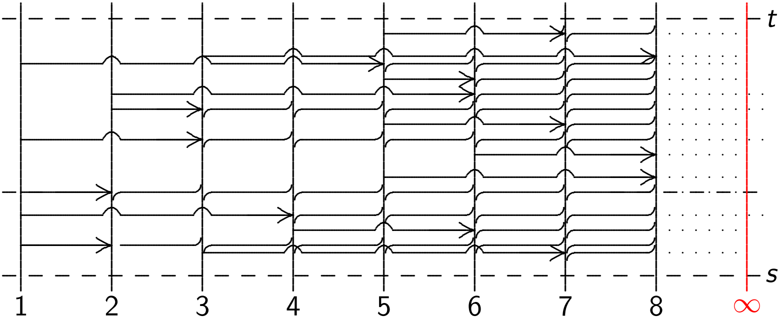
- We can reorder the Moran model by the level of *persistence*
- Individuals at a lower level live longer

The order by persistence is given by the lookdown process



The Lookdown-process

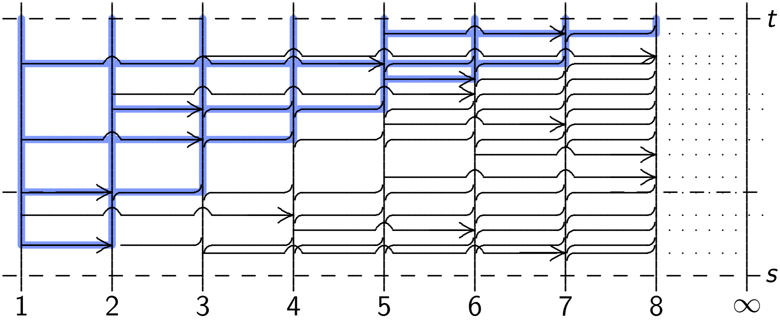
- Every individual at *level j looks down* to $i < j$ at rate 1
- Simultaneously levels $j, j+1, \dots$ increase by 1
- Individuals die at ∞



The Lookdown-process

- Every individual at *level j looks down* to $i < j$ at rate 1
- Simultaneously levels $j, j+1, \ldots$ increase by 1
- Individuals die at ∞

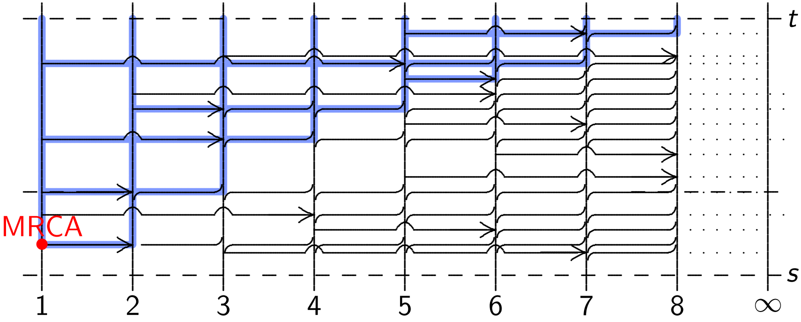
Genealogies are (modulo reordering) as in $\mathcal{K}$



The Lookdown-process

- levels $j+1, j+2\dots$ are pushed whenever j is pushed
- Lines are ordered by persistence

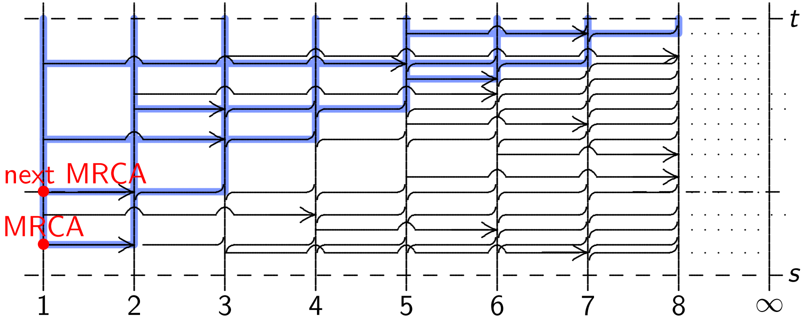
MRCAs always have level 1



The Lookdown-process

- every lookdown $1 \to 2$ defines the birth of a new MRCA
- It will be the MRCA when all individuals descend from him

  Already at t it has some descendant

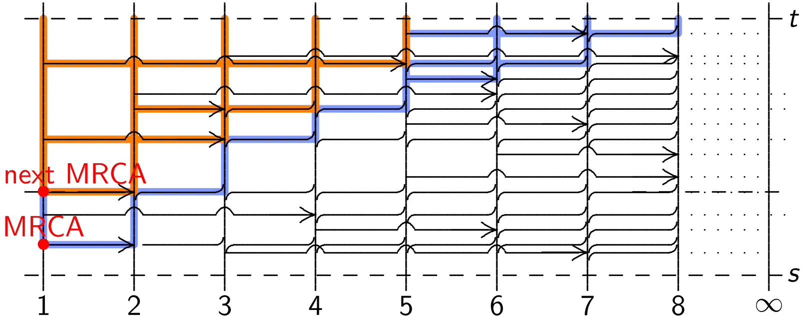


The Lookdown-process

- every lookdown $1 \longrightarrow 2$ defines the birth of a new MRCA
- It will be the MRCA when all individuals descend from him

  Already at t it has some descendant

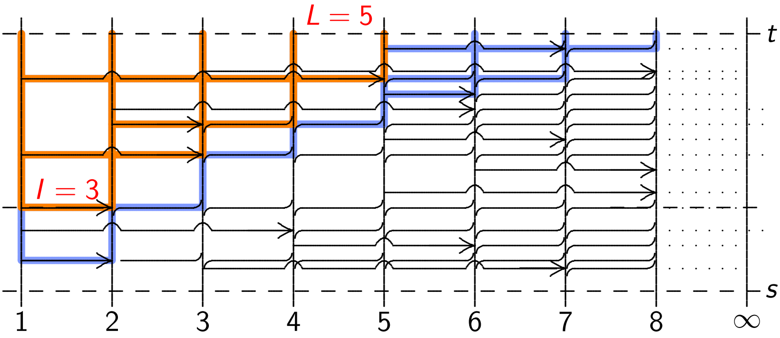
  fixation curves determine the takeover of the new MRCA



The Lookdown-process

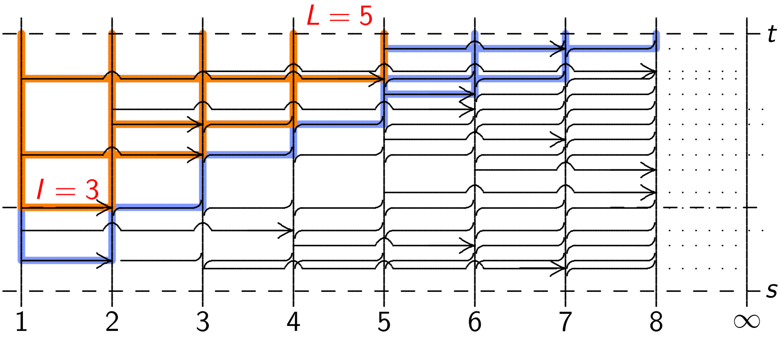
A special role is played by

- I: number of ancestors at the time of the new MRCA
- L: position of the fixation curve at time t



The Lookdown-process

- L: position of the fixation curve at time t
- L individuals still have offspring at the time of the next MRCA change



When does the new MRCA live?

The joint distribution of I and L

- Some combinatorics gives:

$$\mathbf{P}[L = \ell, I = i] = \begin{cases} \frac{\ell!(\ell-1)}{3} \frac{1}{\binom{i+\ell}{i}}, & \ell \geq 2, i \geq 3 \\ \frac{1}{3}, & \ell = 1, i = \infty. \end{cases}$$

Coalescence-time distributions

$$\begin{aligned} S_i^j &: \text{time used for } \mathcal{K} \text{ to come from } j \text{ down to } i \\ &= \text{time, the fixation curve needs to go from } i \text{ to } j \end{aligned}$$

$$S_i^j = \sum_{k=i+1}^{j} T_k \text{ for } T_k \sim \exp\left(\binom{k}{2}\right).$$

For $d > 0$ let $R_{i,d} \sim S_i^\infty$ given $S_1^i + S_i^\infty = d$.

When does the new MRCA live?

Theorem Let (L, I) be as above. Given $A_0 = -d$, the next MRCA will be found at time E and it will have lived at time B, where

$$(E, B) = \begin{cases} (S_1^2 + S_2^\infty, S_1^2) & \text{if } L = 1, \\ (S_L^\infty, -R_{I,d}) & \text{if } L > 1. \end{cases}$$

Corollary Let $X \sim \exp(1)$. Then

$$\mathbf{P}[X \in dt] = \sum_{\ell=1}^{\infty} \frac{2}{(\ell+1)(\ell+2)} \mathbf{P}[S_\ell^\infty \in dt].$$

MRCA processes

Fixation curves

- In the lookdown every lookdown $2 \to 1$ determines the start of a new fixation curve
- Fixation curves move from k to $k+1$ at rate $\binom{k+1}{2}$.

The MRCA processes

- The MRCA process $\mathcal{A} = (A_t)_t$ gives times of MRCAs at all times
- The point process $\mathcal{F} = \{(E, B)\}$ of MRCA times B and MRCA jump times E is called the MRCA point process
- Both can be constructed from the lookdown process

Number of MRCAs in the past

- Denote the number of fixation curves in equilibrium by Z
- Z gives the number of MRCAs that are established in the future which live in the past
- The probability generating function of Z is

$$\mathbf{E}[u^Z] = \frac{1}{3} \exp\Big(\sum_{i=2}^{\infty} \log\Big(\frac{i(i+1)+2(u-1)}{(i+2)(i-1)} \Big) \Big).$$

- $$\begin{aligned} \mathbf{E}[Z] &= 1, & \mathbf{Var}[Z] &= 14 - \tfrac{4}{3}\pi^2 \approx 0.84052, \\ \mathbf{P}[Z=0] &= \frac{1}{3}, & \mathbf{P}[Z=1] &= \frac{11}{27} \approx 0.40740, \\ \mathbf{P}[Z=2] &= \frac{107}{243} - \frac{2}{81}\pi^2 \approx 0.19664. \end{aligned}$$