Lines of descent in a deterministic model with mutation and frequency dependent selection

Fernando Cordero - Bielefeld University *

School on Information and Randomness 2016 - Santiago de Chile

December 2016

*joint work (in progress) with E. Baake and S. Hummel

Structure of the talk

The deterministic mutation and frequency dependent selection model

We consider a 2-type (types in $\{0, 1\}$) population subject to mutation and pairwise dependent selection, where

The deterministic mutation and frequency dependent selection model

We consider a 2-type (types in $\{0, 1\}$) population subject to mutation and pairwise dependent selection, where

- starting with a proportion z_0 of 0-individuals,

The deterministic mutation and frequency dependent selection model

We consider a 2-type (types in $\{0, 1\}$) population subject to mutation and pairwise dependent selection, where

- starting with a proportion z_0 of 0-individuals,
- the proportion $z(t, z_0)$ of 0-individuals at time t follows

The deterministic mutation and frequency dependent selection model

We consider a 2-type (types in $\{0, 1\}$) population subject to mutation and pairwise dependent selection, where

- starting with a proportion z_0 of 0-individuals,
- the proportion $z(t, z_0)$ of 0-individuals at time t follows

$$\frac{dz}{dt}(t) = \alpha(\beta - \gamma z(t))z(t)(1 - z(t)) + u_0(1 - z(t)) - u_1 z(t), \quad (1)$$

The deterministic mutation and frequency dependent selection model

We consider a 2-type (types in $\{0,1\}$) population subject to mutation and pairwise dependent selection, where

- starting with a proportion z_0 of 0-individuals,
- the proportion $z(t,z_0)$ of 0-individuals at time t follows

$$\frac{dz}{dt}(t) = \alpha(\beta - \gamma z(t))z(t)(1-z(t)) + u_0(1-z(t)) - u_1 z(t), \quad (1)$$

- $\alpha\beta$ non-interactive selection, $\alpha|\gamma|$ interactive selection, u_i mutation to type i, $i \in \{0,1\}$, $u := u_0 + u_1$ total rate of mutation.

The deterministic mutation and frequency dependent selection model

We consider a 2-type (types in $\{0,1\}$) population subject to mutation and pairwise dependent selection, where

- starting with a proportion z_0 of 0-individuals,
- the proportion $z(t, z_0)$ of 0-individuals at time t follows

$$\frac{dz}{dt}(t) = \alpha(\beta - \gamma z(t))z(t)(1 - z(t)) + u_0(1 - z(t)) - u_1 z(t), \quad (1)$$

- $\alpha\beta$ non-interactive selection, $\alpha|\gamma|$ interactive selection, u_i mutation to type i, $i \in \{0,1\}$, $u := u_0 + u_1$ total rate of mutation.
- In the sequel, we focus on the case $\beta > 0, \gamma \leq 0$.

The problem

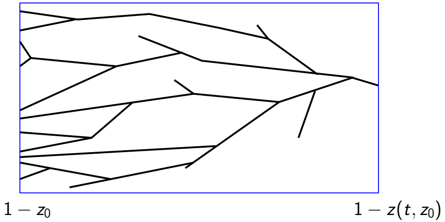
- Forward in time behavior is well-known: stable points, bifurcation structure...

The problem

- Forward in time behavior is well-known: stable points, bifurcation structure...
- Ancestral picture is missing!

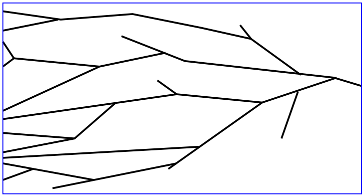
The problem

- Forward in time behavior is well-known: stable points, bifurcation structure...
- Ancestral picture is missing!



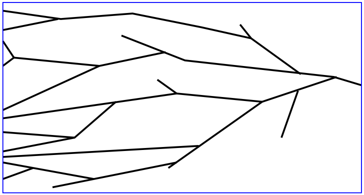
The problem

- Forward in time behavior is well-known: stable points, bifurcation structure...
- Ancestral picture is missing!
- Prob. of sampling a 1−individual at time $t \equiv 1 - z(t, z_0)$.



The problem

- Forward in time behavior is well-known: stable points, bifurcation structure...
- Ancestral picture is missing!
- Prob. of sampling a 1−individual at time $t \equiv 1 - z(t, z_0)$.
- Can we express this probability using an ancestral process (duality!)?



$1 - z_0$ $\qquad$ $1 - z(t, z_0)$

The Moran model ($\beta > 0$, $\gamma \leq 0$)

Population of size N, two types $\{0, 1\}$.

The Moran model ($\beta > 0$, $\gamma \leq 0$)

Population of size N, two types $\{0, 1\}$.

- If an individual reproduces, its single offspring inherits its type and replaces a uniformly chosen individual.

The Moran model ($\beta > 0$, $\gamma \leq 0$)

Population of size N, two types $\{0, 1\}$.

- If an individual reproduces, its single offspring inherits its type and replaces a uniformly chosen individual.
  - Neutral reproduction:
    Each individual reproduces at rate 1.

The Moran model ($\beta > 0$, $\gamma \leq 0$)

Population of size N, two types $\{0, 1\}$.

- If an individual reproduces, its single offspring inherits its type and replaces a uniformly chosen individual.
  - Neutral reproduction:
    Each individual reproduces at rate 1.
  - Non-interactive selective reproduction:
    Each type $0-$individual reproduces at rate $\alpha\beta$.

The Moran model ($\beta > 0$, $\gamma \leq 0$)

Population of size N, two types $\{0,1\}$.

- If an individual reproduces, its single offspring inherits its type and replaces a uniformly chosen individual.
  - Neutral reproduction:
    Each individual reproduces at rate 1.
  - Non-interactive selective reproduction:
    Each type $0-$individual reproduces at rate $\alpha\beta$.
  - Non-interactive selective reproduction:
    Each type $0-$individual picks at random an individual at rate $\alpha|\gamma|$ and checks its type. Type $= 0 \Rightarrow$ reproduction.

The Moran model ($\beta > 0$, $\gamma \leq 0$)

Population of size N, two types $\{0, 1\}$.

- If an individual reproduces, its single offspring inherits its type and replaces a uniformly chosen individual.
  - Neutral reproduction:
    Each individual reproduces at rate 1.
  - Non-interactive selective reproduction:
    Each type $0-$individual reproduces at rate $\alpha\beta$.
  - Non-interactive selective reproduction:
    Each type $0-$individual picks at random an individual at rate $\alpha|\gamma|$ and checks its type. Type $= 0 \Rightarrow$ reproduction.
- Each individual mutates to type i at rate u_i, $i \in \{0, 1\}$.

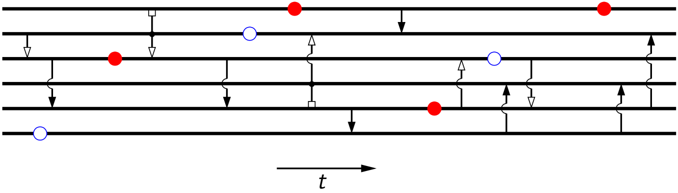
The graphical representation ($\beta > 0, \gamma < 0$)

→ = Neutral reproduction, ⇾ = Non-interactive selective reproduction,

□–•–▷ = Interactive selective reproduction,

○ = Mutation to type 0, ● = Mutation to type 1,

Untyped picture:



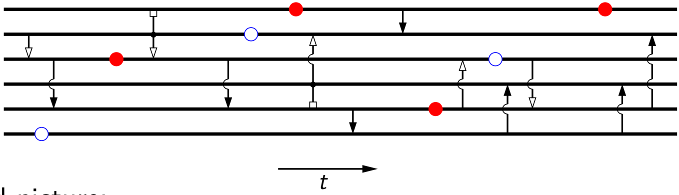
The graphical representation ($\beta > 0, \gamma < 0$)

→ = Neutral reproduction, ⇾ = Non-interactive selective reproduction,

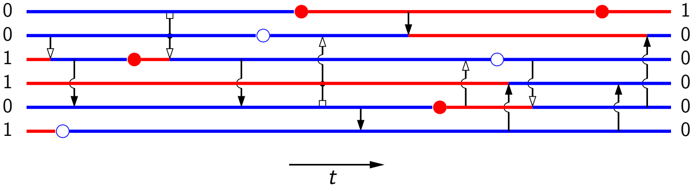
□–•–▷ = Interactive selective reproduction,

○ = Mutation to type 0, ● = Mutation to type 1,

Untyped picture:



Typed picture:



Dynamical law of large numbers

For $t \geq 0$, we denote by X_t^N the number of $0-$individuals at time t in the Moran model of size N.

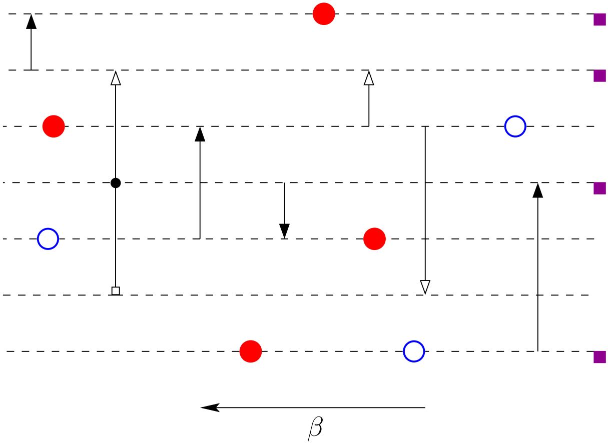
Proposition (Law of large numbers)

For $z_0 \in [0,1]$, let $z(z_0,\cdot)$ be the solution of (1) *with $z(0,z_0) = z_0$. Assume that $\lim_{N\to\infty} \frac{X_0^N}{N} = z_0 \in [0,1]$. Then, for all $\varepsilon > 0$, we have*

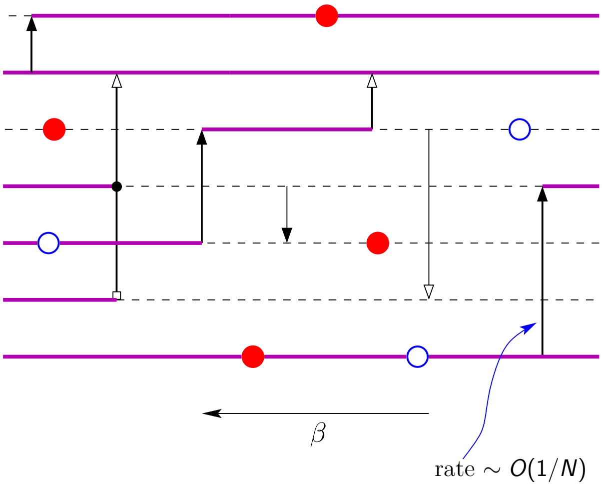
$$\lim_{N\to\infty} P\left(\sup_{t\leq T}\left|\frac{X_t^N}{N} - z(t,z_0)\right| > \varepsilon\right) = 0,$$

i.e. X^N/N converges to $z(\cdot,z_0)$ uniformly in compacts in probability.

The ASG (N large)



The ASG (N large)

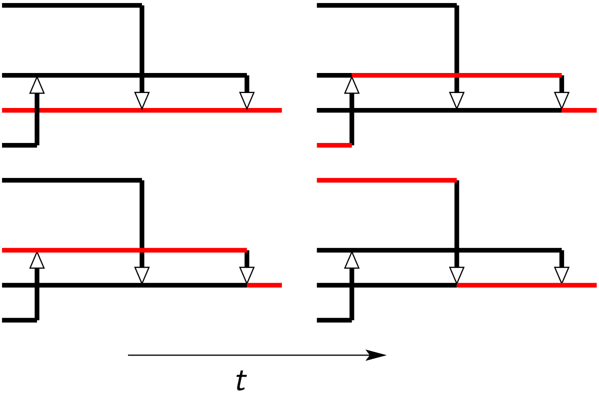


The asymptotic ASG ($\gamma = 0$)

- When $\gamma = 0$ and $N \to \infty$, only bifurcation and mutation events survive.

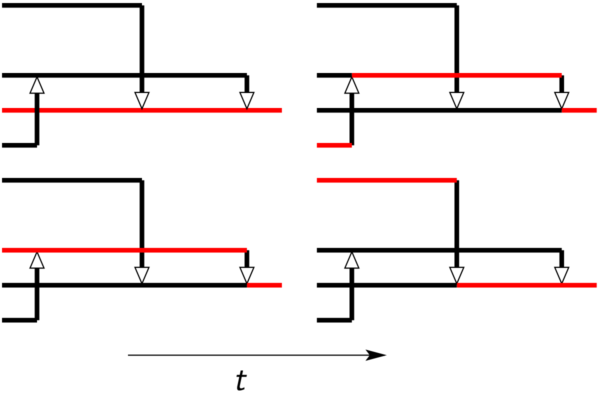
The asymptotic ASG ($\gamma = 0$)

- When $\gamma = 0$ and $N \to \infty$, only bifurcation and mutation events survive.
- In addition, if there is no mutation



The asymptotic ASG ($\gamma = 0$)

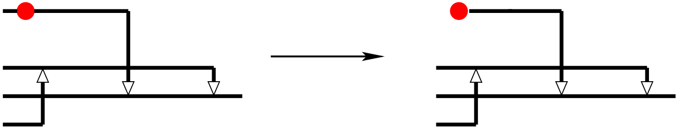
- When $\gamma = 0$ and $N \to \infty$, only bifurcation and mutation events survive.
- In addition, if there is no mutation

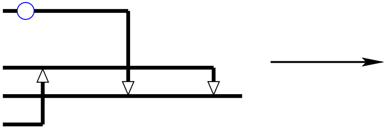


The probability of sampling an individual of type 0 at (forward) time t depends only on the number of lines in the ASG at time 0.

The killed ASG ($\gamma = 0$)

- With mutations:

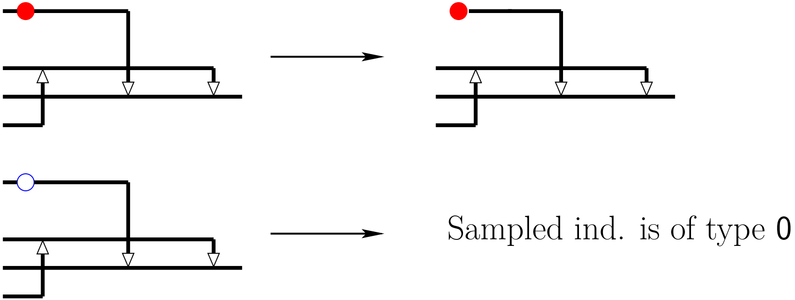




Sampled ind. is of type 0

The killed ASG ($\gamma = 0$)

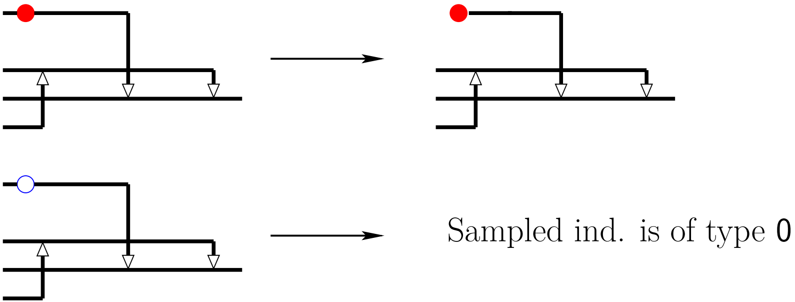
- With mutations:



Mutation to type 1 $\rightarrow$ erase one line.

The killed ASG ($\gamma = 0$)

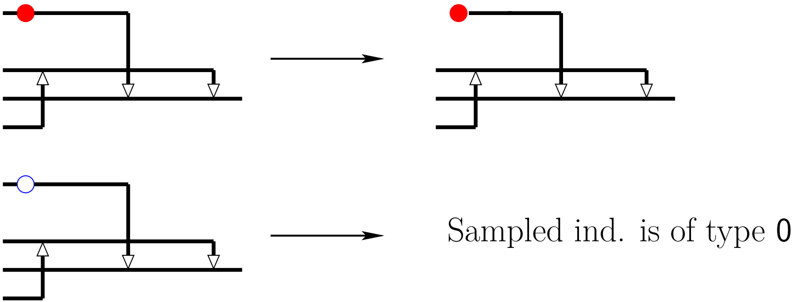
- With mutations:



Mutation to type 1 $\rightarrow$ erase one line.
Mutation to type 0 $\rightarrow$ stop! the sampled individual is of type 0.

The killed ASG ($\gamma = 0$)

- With mutations:



Mutation to type 1 $\rightarrow$ erase one line.
Mutation to type 0 $\rightarrow$ stop! the sampled individual is of type 0.
Killed ASG (E. Baake, U. Lenz, A. Wakolbinger)

Duality

Let $R = (R_t)_{\geq 0}$ be the line-counting process in the killed-ASG, i.e. the continuous time Markov chain with rates

$$q_R(k,j) := \begin{cases} ks & \text{if } j = k+1, \\ ku_1 & \text{if } j = k-1, \\ ku_0 & \text{if } j = \infty. \end{cases}$$

Theorem (Duality)

For $t \geq 0$ and $z_0 \in (0,1)$, we have

$$1 - z(t, z_0) := E_1\left[(1-z_0)^{R_t}\right].$$

Absorption probability

Define $h_k := P(R \text{ absorbs in } 0 | R_0 = k)$

Absorption probability

Define $h_k := P(R \text{ absorbs in } 0 | R_0 = k)$

- $h_0 = 1$ and $h_\infty = 0$.

Absorption probability

Define $h_k := P(R \text{ absorbs in } 0 | R_0 = k)$

- $h_0 = 1$ and $h_\infty = 0$.
- First step analysis $\Rightarrow sh_{k+1} - (u+s)h_k + u_1 h_{k-1} = 0$.

Absorption probability

Define $h_k := P(R \text{ absorbs in } 0 | R_0 = k)$

- $h_0 = 1$ and $h_\infty = 0$.
- First step analysis $\Rightarrow sh_{k+1} - (u+s)h_k + u_1 h_{k-1} = 0$.
- $h_k = y_*^k$, where

$$y_* := \begin{cases} \frac{u+s-\sqrt{(u+s)^2 - 4u_1 s}}{2s} & \text{if } s > 0, \\ \frac{u_1}{u} & \text{if } s = 0. \end{cases}$$

Absorption probability

Define $h_k := P(R \text{ absorbs in } 0 | R_0 = k)$

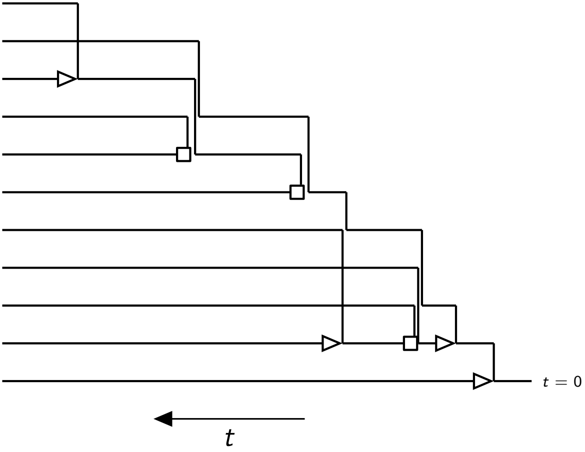
- $h_0 = 1$ and $h_\infty = 0$.
- First step analysis $\Rightarrow sh_{k+1} - (u+s)h_k + u_1 h_{k-1} = 0$.
- $h_k = y_*^k$, where

$$y_* := \begin{cases} \frac{u+s-\sqrt{(u+s)^2-4u_1 s}}{2s} & \text{if } s > 0, \\ \frac{u_1}{u} & \text{if } s = 0. \end{cases}$$

- Duality $\Rightarrow y_*$ is also the proportion of 1's at stationarity.

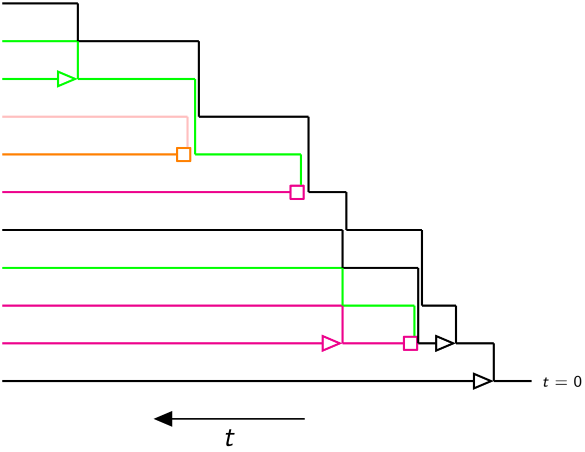
The Asymptotic ASG with interactions and no mutations

The asymptotic ASG consists of bifurcations and trifurcations.

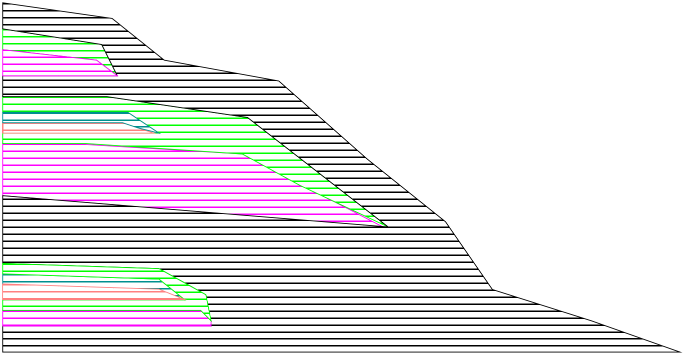


The Asymptotic ASG with interactions and no mutations

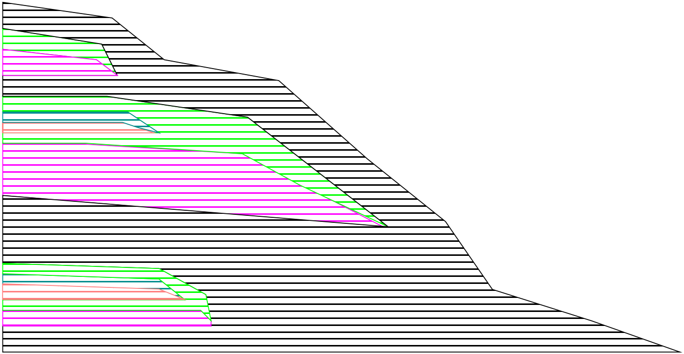
The asymptotic ASG consists of bifurcations and trifurcations.



Ancestral selection graph

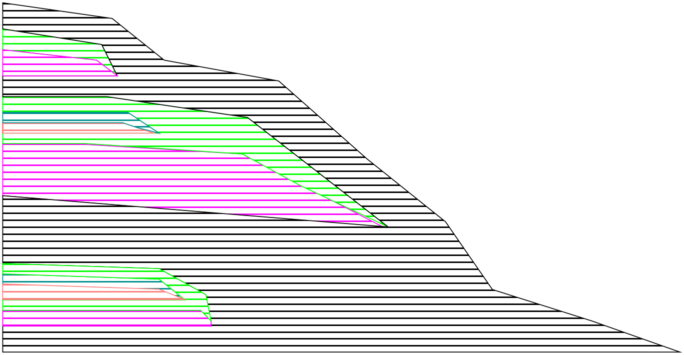


Ancestral selection graph



How to encode this information in order to determine the probability of sampling an individual of type 1?

Ancestral selection graph



How to encode this information in order to determine the probability of sampling an individual of type 1? Ternary trees with marked leaves!

The line-counting tree without mutation

- We start with ①.

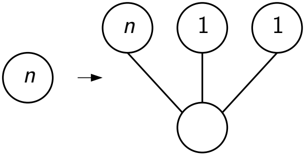
The line-counting tree without mutation

- We start with (1).
- Before the first trifurcation: the root increases its level by 1 at each bifurcation.

The line-counting tree without mutation

- We start with (1).
- Before the first trifurcation: the root increases its level by 1 at each bifurcation.

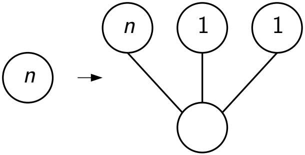
- At the first trifurcation: the root produces three children, the left child inherits the mark of the root and the other two are marked with 1.



The line-counting tree without mutation

- We start with (1).
- Before the first trifurcation: the root increases its level by 1 at each bifurcation.

- At the first trifurcation: the root produces three children, the left child inherits the mark of the root and the other two are marked with 1.

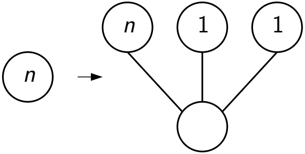


  - Left-child $\equiv$ the region before trifurcation.

The line-counting tree without mutation

- We start with (1).
- Before the first trifurcation: the root increases its level by 1 at each bifurcation.

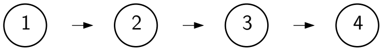
- At the first trifurcation: the root produces three children, the left child inherits the mark of the root and the other two are marked with 1.



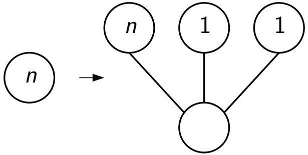
  - Left-child ≡ the region before trifurcation.
  - Middle-child ≡ the checking line.

The line-counting tree without mutation

- We start with (1).
- Before the first trifurcation: the root increases its level by 1 at each bifurcation.



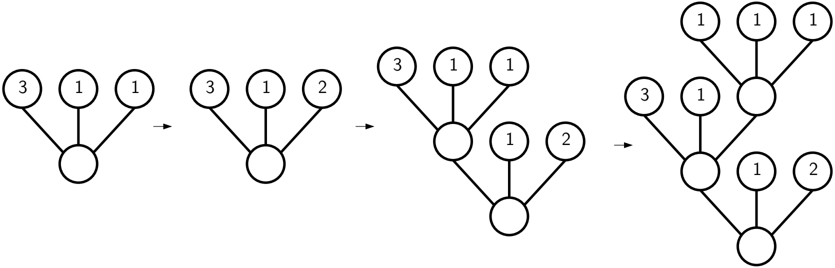
- At the first trifurcation: the root produces three children, the left child inherits the mark of the root and the other two are marked with 1.



  - Left-child $\equiv$ the region before trifurcation.
  - Middle-child $\equiv$ the checking line.
  - Right-child $\equiv$ the incoming line.

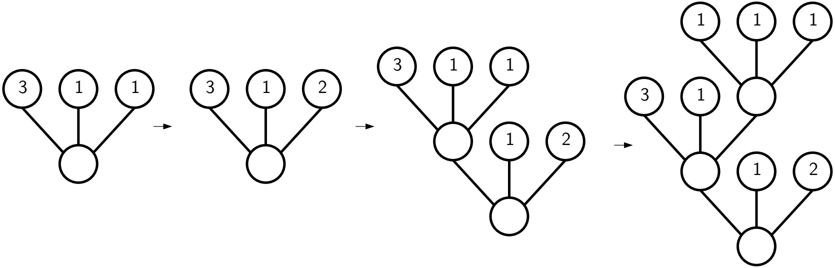
The line-counting tree without mutation

- Leaves evolve independently, increasing its mark by 1 at each bifurcation and producing three children at a trifurcation.



The line-counting tree without mutation

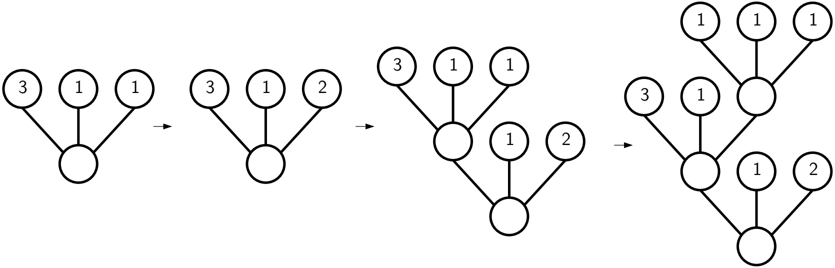
- Leaves evolve independently, increasing its mark by 1 at each bifurcation and producing three children at a trifurcation.



- Thus, for $\mathcal{T} = (\tau, m)$ and $\ell \in L_\tau$, we have transitions

The line-counting tree without mutation

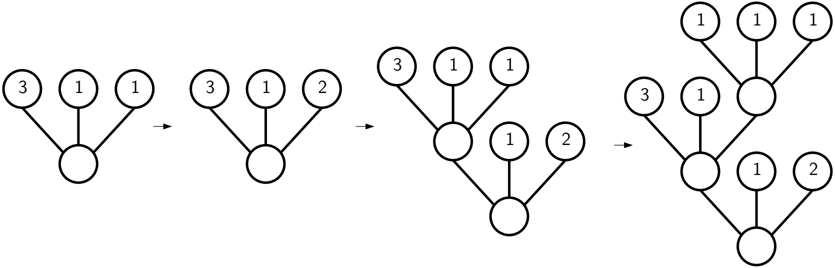
- Leaves evolve independently, increasing its mark by 1 at each bifurcation and producing three children at a trifurcation.



- Thus, for $\mathcal{T} = (\tau, m)$ and $\ell \in L_\tau$, we have transitions
  1. $\mathcal{T} \mapsto \mathcal{T}^{\curlyvee,\ell}$ at rate $\alpha\beta m(\ell)$.

The line-counting tree without mutation

- Leaves evolve independently, increasing its mark by 1 at each bifurcation and producing three children at a trifurcation.



- Thus, for $\mathcal{T} = (\tau, m)$ and $\ell \in L_\tau$, we have transitions
  1. $\mathcal{T} \mapsto \mathcal{T}^{\curlyvee,\ell}$ at rate $\alpha\beta m(\ell)$.
  2. $\mathcal{T} \mapsto \mathcal{T}^{\Psi,\ell}$ at rate $\alpha|\gamma| m(\ell)$.

Assigning types

- Each leaf ℓ represents a region with $m(\ell)$ lines in the ASG.

Assigning types

- Each leaf ℓ represents a region with $m(\ell)$ lines in the ASG.
- The rules are "locally" as in the non-interactive case.

Assigning types

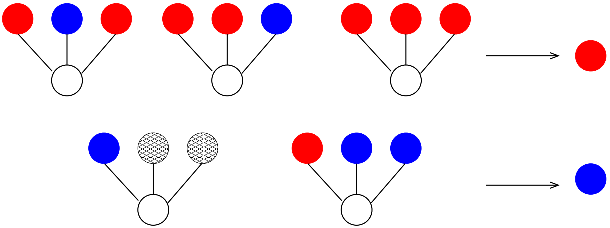
- Each leaf ℓ represents a region with $m(\ell)$ lines in the ASG.
- The rules are "locally" as in the non-interactive case.
- Determining the type of the sampled ind. is equivalent to the following coloring problem.

Assigning types

- Each leaf ℓ represents a region with $m(\ell)$ lines in the ASG.
- The rules are "locally" as in the non-interactive case.
- Determining the type of the sampled ind. is equivalent to the following coloring problem.
    1. If one of the lines in the leaf ℓ is of type 0, then we color it blue, otherwise it is colored red.

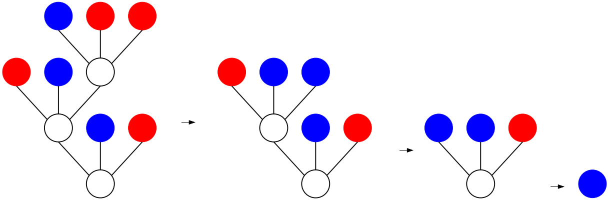
Assigning types

- Each leaf ℓ represents a region with $m(\ell)$ lines in the ASG.
- The rules are "locally" as in the non-interactive case.
- Determining the type of the sampled ind. is equivalent to the following coloring problem.
  1. If one of the lines in the leaf ℓ is of type 0, then we color it blue, otherwise it is colored red.
  2. We propagate the colors as follows



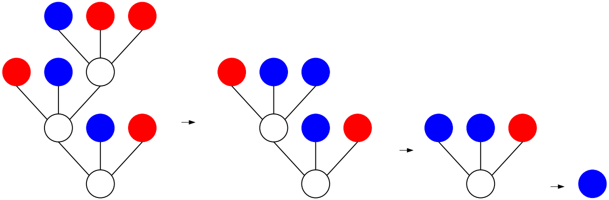
Assigning types

- Each leaf ℓ represents a region with $m(\ell)$ lines in the ASG.
- The rules are "locally" as in the non-interactive case.
- Determining the type of the sampled ind. is equivalent to the following coloring problem.
  1. If one of the lines in the leaf ℓ is of type 0, then we color it blue, otherwise it is colored red.
  2. We propagate the colors as follows



Assigning types

- Each leaf ℓ represents a region with $m(\ell)$ lines in the ASG.
- The rules are "locally" as in the non-interactive case.
- Determining the type of the sampled ind. is equivalent to the following coloring problem.
    1. If one of the lines in the leaf ℓ is of type 0, then we color it blue, otherwise it is colored red.
    2. We propagate the colors as follows



    3. The sampled ind. is of type 1 iff the root is colored red.

Computing the sampling probability

- Assume that the current proportion of $1-$individuals of is y and that the line-counting tree is at state $\mathcal{T} = (\tau, m)$.

Computing the sampling probability

- Assume that the current proportion of $1-$individuals of is y and that the line-counting tree is at state $\mathcal{T} = (\tau, m)$.
- $H(\mathcal{T}, y) :=$ the probability of sampling a $1-$individual.

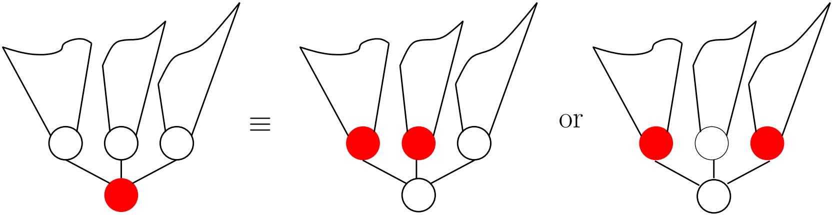
Computing the sampling probability

- Assume that the current proportion of $1-$individuals of is y and that the line-counting tree is at state $\mathcal{T} = (\tau, m)$.
- $H(\mathcal{T}, y) :=$ the probability of sampling a $1-$individual.
- $H(\textcircled{n}, y) = y^n$.

Computing the sampling probability

- Assume that the current proportion of 1−individuals of is y and that the line-counting tree is at state $\mathcal{T} = (\tau, m)$.
- $H(\mathcal{T}, y) :=$ the probability of sampling a 1−individual.
- $H(\textcircled{n}, y) = y^n$.
- If v_1, v_2, v_3 are the left, middle and right children of the root:

$$H(\mathcal{T}, y) = H(\mathcal{T}_{v_1}, y)\left[H(\mathcal{T}_{v_2}, y) + H(\mathcal{T}_{v_3}, y) - H(\mathcal{T}_{v_2}, y)H(\mathcal{T}_{v_3}, y)\right].$$



Pruning and grafting: mutation to type 1

Mutation to type 1 happens at a leaf ℓ at rate $u_1 m(\ell)$.

Pruning and grafting: mutation to type 1

Mutation to type 1 happens at a leaf ℓ at rate $u_1 m(\ell)$.

- If $m(\ell) > 1$, then the effect is local (we decrease in 1 the mark of the leaf).

Pruning and grafting: mutation to type 1

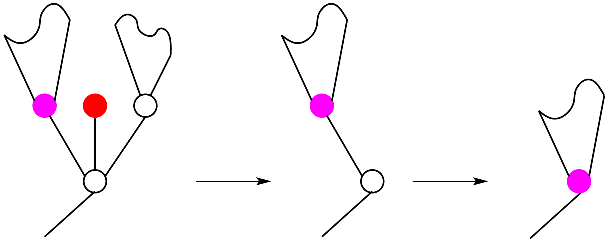
Mutation to type 1 happens at a leaf ℓ at rate $u_1 m(\ell)$.

- If $m(\ell) > 1$, then the effect is local (we decrease in 1 the mark of the leaf).
- The same happens if $m(\ell) = 1$ and ℓ is the left-child of its parent.

Pruning and grafting: mutation to type 1

Mutation to type 1 happens at a leaf ℓ at rate $u_1 m(\ell)$.

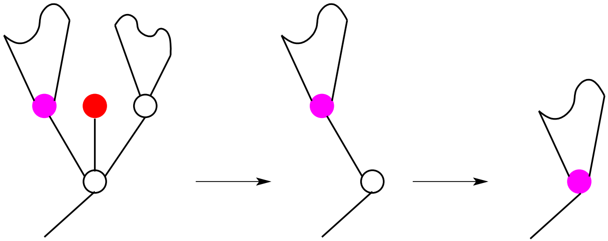
- If $m(\ell) > 1$, then the effect is local (we decrease in 1 the mark of the leaf).
- The same happens if $m(\ell) = 1$ and ℓ is the left-child of its parent.
- If not, then



Pruning and grafting: mutation to type 1

Mutation to type 1 happens at a leaf ℓ at rate $u_1 m(\ell)$.

- If $m(\ell) > 1$, then the effect is local (we decrease in 1 the mark of the leaf).
- The same happens if $m(\ell) = 1$ and ℓ is the left-child of its parent.
- If not, then



This transition is denoted by $\mathcal{T} \mapsto \mathcal{T}^{\bullet,\ell}$.

Pruning and grafting: mutation to type 0

Mutation to type 0 happens at a leaf ℓ at rate $u_0 m(\ell)$.

Pruning and grafting: mutation to type 0

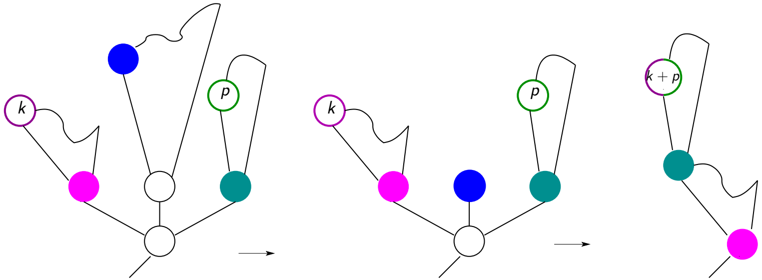
Mutation to type 0 happens at a leaf ℓ at rate $u_0 m(\ell)$.

- If ℓ is the most-left leaf of $\mathcal{T}$: stop! The sampled indiv. is of type 0. We set $\mathcal{T}^{\circ,\ell} = \bowtie$ (cemetery point).

Pruning and grafting: mutation to type 0

Mutation to type 0 happens at a leaf ℓ at rate $u_0 m(\ell)$.

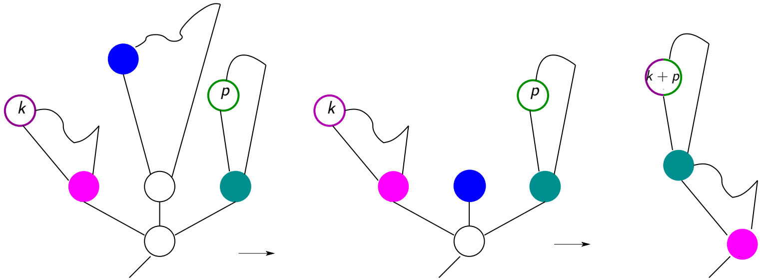
- If ℓ is the most-left leaf of $\mathcal{T}$: stop! The sampled indiv. is of type 0. We set $\mathcal{T}^{\circ,\ell} = \bowtie$ (cemetery point).
- If not, there are two options



Pruning and grafting: mutation to type 0

Mutation to type 0 happens at a leaf ℓ at rate $u_0 m(\ell)$.

- If ℓ is the most-left leaf of $\mathcal{T}$: stop! The sampled indiv. is of type 0. We set $\mathcal{T}^{\circ,\ell} = \bowtie$ (cemetery point).
- If not, there are two options



This transition is denoted by $\mathcal{T} \mapsto \mathcal{T}^{\circ,\ell}$.

The line-counting tree

Definition

The line-counting tree process $\mathcal{T}^* := (\mathcal{T}^*(t))_{t \geq 0}$ is the jump process with values on $\Upsilon \cup \{\bowtie\}$ starting at $\textcircled{1}$ and with the following transitions. If $\mathcal{T}^*$ is currently in state $\mathcal{T} := (\tau, m)$: for each $\ell \in L_\tau$,

1. $\mathcal{T} \longrightarrow \mathcal{T}^{\curlyvee,\ell}$ at rate $\alpha\beta\, m(\ell)$.
2. $\mathcal{T} \longrightarrow \mathcal{T}^{\Psi,\ell}$ at rate $\alpha|\gamma|\, m(\ell)$.
3. $\mathcal{T} \longrightarrow \mathcal{T}^{\bullet,\ell}$ at rate $u_1\, m(\ell)$.
4. $\mathcal{T} \longrightarrow \mathcal{T}^{\circ,\ell}$ at rate $u_0\, m(\ell)$.

The states $\textcircled{0}$ and $\bowtie$ are absorbing states of $\mathcal{T}^*$.

Duality I

We extend the definition of H to the cemetery point by setting $H(\bowtie, y) := 0$.

Duality I

We extend the definition of H to the cemetery point by setting $H(\bowtie, y) := 0$.

Theorem (Duality)

For all $t > 0$ and $z_0 \in (0,1)$, we have

$$1 - z(t, z_0) = E_{\textcircled{1}}[H(\mathcal{T}^*(t), 1 - z_0)].$$

Duality I

We extend the definition of H to the cemetery point by setting $H(\bowtie, y) := 0$.

Theorem (Duality)

For all $t > 0$ and $z_0 \in (0,1)$, we have

$$1 - z(t, z_0) = E_{①}[H(\mathcal{T}^*(t), 1 - z_0)].$$

Proof.

Consider the det. Markov process $y(\cdot, y_0) = 1 - z(\cdot, 1 - y_0)$, $t \geq 0, y_0 \in (0,1)$.

Duality I

We extend the definition of H to the cemetery point by setting $H(\bowtie, y) := 0$.

Theorem (Duality)

For all $t > 0$ *and* $z_0 \in (0,1)$, *we have*

$$1 - z(t, z_0) = E_{\textcircled{1}}[H(\mathcal{T}^*(t), 1 - z_0)].$$

Proof.

Consider the det. Markov process $y(\cdot, y_0) = 1 - z(\cdot, 1 - y_0)$, $t \geq 0, y_0 \in (0,1)$. We show by induction on the size of the underlying trees that

$$A_y H(\mathcal{T}, \cdot)(y_0) = A_{\mathcal{T}^*} H(\cdot, y_0)(\mathcal{T}).$$

Duality I

We extend the definition of H to the cemetery point by setting $H(\bowtie, y) := 0$.

Theorem (Duality)

For all $t > 0$ and $z_0 \in (0,1)$, we have

$$1 - z(t, z_0) = E_{\textcircled{1}}[H(\mathcal{T}^*(t), 1 - z_0)].$$

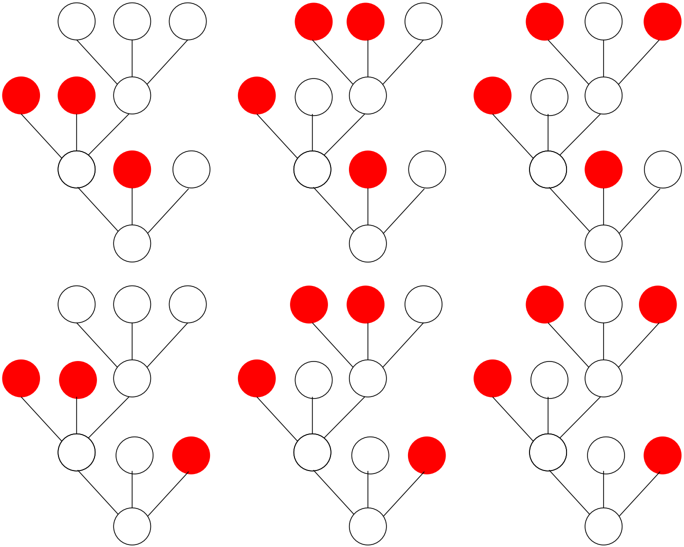
Proof.

Consider the det. Markov process $y(\cdot, y_0) = 1 - z(\cdot, 1 - y_0)$, $t \geq 0, y_0 \in (0,1)$. We show by induction on the size of the underlying trees that

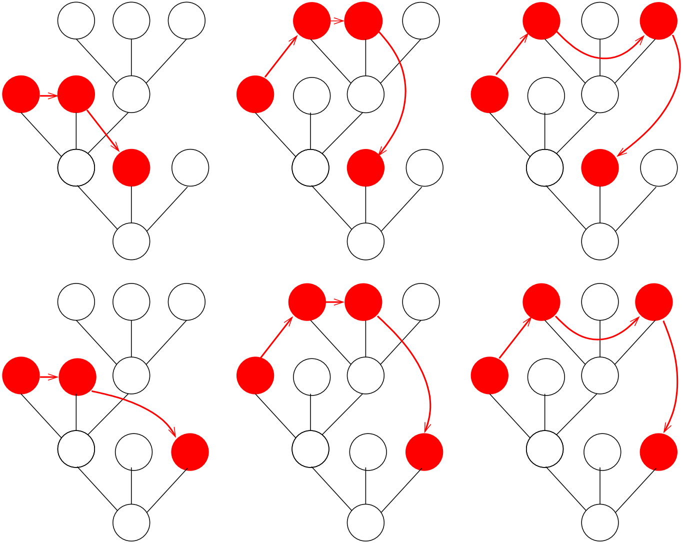
$$A_y H(\mathcal{T}, \cdot)(y_0) = A_{\mathcal{T}^*} H(\cdot, y_0)(\mathcal{T}).$$

The result follows from classical duality results. □

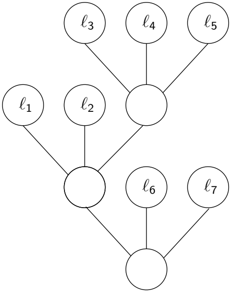
Minimal red-colorings



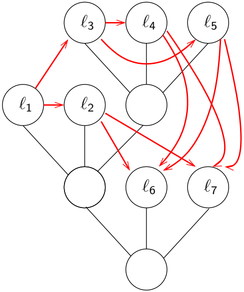
Minimal red-colorings



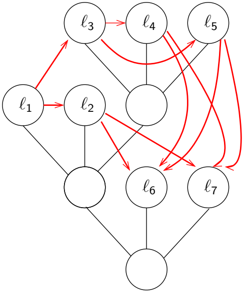
From ternary trees to directed graphs (digraphs)

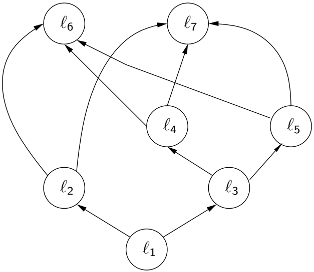


From ternary trees to directed graphs (digraphs)

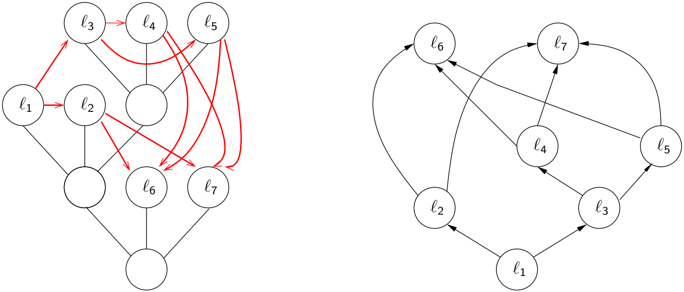


From ternary trees to directed graphs (digraphs)



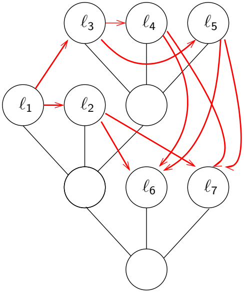


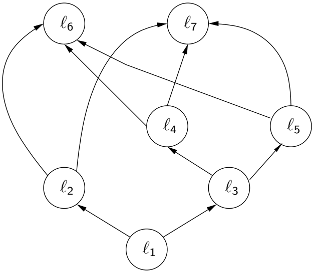
From ternary trees to directed graphs (digraphs)



- Most left-leave $\mapsto$ Source.

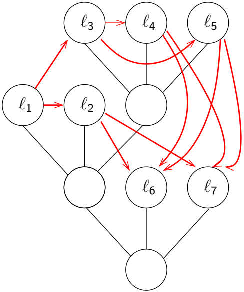
From ternary trees to directed graphs (digraphs)

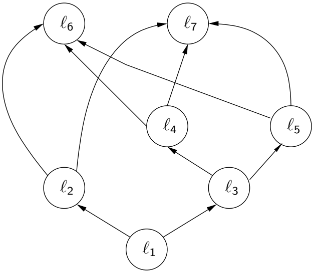




- Most left-leave $\mapsto$ Source.
- Leaves $\mapsto$ Nodes.

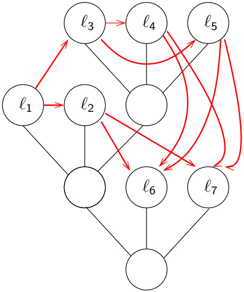
From ternary trees to directed graphs (digraphs)

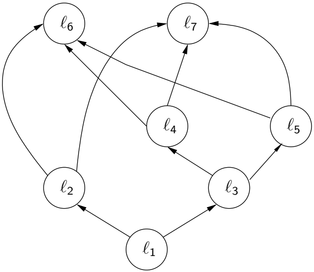




- Most left-leave $\mapsto$ Source.
- Leaves $\mapsto$ Nodes.
- Marks on the leaves $\mapsto$ Marks on the nodes.

From ternary trees to directed graphs (digraphs)





- Most left-leave $\mapsto$ Source.
- Leaves $\mapsto$ Nodes.
- Marks on the leaves $\mapsto$ Marks on the nodes.
- Minimal colorings $\mapsto$ Strings (paths from source to a sink).

Transforming ternary trees in directed graphs (digraphs)

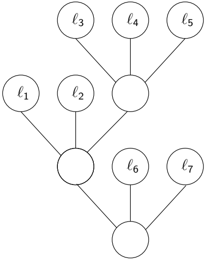
- $\rho(\textcircled{n}) = \textcircled{n}$.

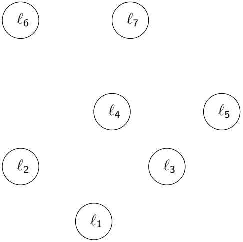
Transforming ternary trees in directed graphs (digraphs)

- $\rho(\textcircled{n}) = \textcircled{n}$.
- Denote by v_1, v_2, v_3 the left, middle and right children of the root of $\mathcal{T}$.

Transforming ternary trees in directed graphs (digraphs)

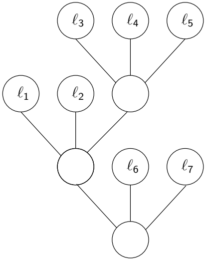
- $\rho(\textcircled{n}) = \textcircled{n}$.
- Denote by v_1, v_2, v_3 the left, middle and right children of the root of $\mathcal{T}$.
- $\rho(\mathcal{T})$ is defined from $\rho(\mathcal{T}_{v_1})$, $\rho(\mathcal{T}_{v_2})$ and $\rho(\mathcal{T}_{v_3})$ by adding, for each sink v of $\rho(\mathcal{T}_{v_1})$, two edges, one from v to the source of $\rho(\mathcal{T}_{v_2})$ and the other from v to the source of $\rho(\mathcal{T}_{v_3})$.

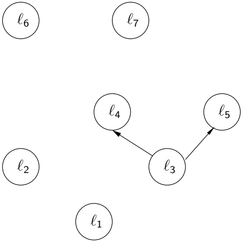




Transforming ternary trees in directed graphs (digraphs)

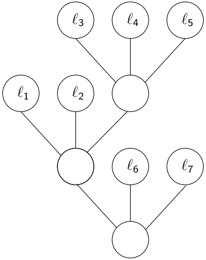
- $\rho(\textcircled{n}) = \textcircled{n}$.
- Denote by v_1, v_2, v_3 the left, middle and right children of the root of $\mathcal{T}$.
- $\rho(\mathcal{T})$ is defined from $\rho(\mathcal{T}_{v_1})$, $\rho(\mathcal{T}_{v_2})$ and $\rho(\mathcal{T}_{v_3})$ by adding, for each sink v of $\rho(\mathcal{T}_{v_1})$, two edges, one from v to the source of $\rho(\mathcal{T}_{v_2})$ and the other from v to the source of $\rho(\mathcal{T}_{v_3})$.

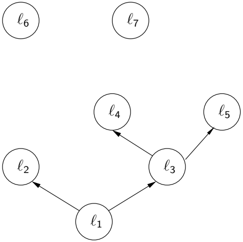




Transforming ternary trees in directed graphs (digraphs)

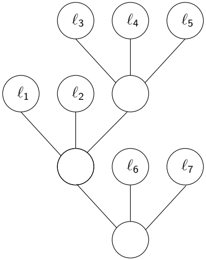
- $\rho(\textcircled{n}) = \textcircled{n}$.
- Denote by v_1, v_2, v_3 the left, middle and right children of the root of $\mathcal{T}$.
- $\rho(\mathcal{T})$ is defined from $\rho(\mathcal{T}_{v_1})$, $\rho(\mathcal{T}_{v_2})$ and $\rho(\mathcal{T}_{v_3})$ by adding, for each sink v of $\rho(\mathcal{T}_{v_1})$, two edges, one from v to the source of $\rho(\mathcal{T}_{v_2})$ and the other from v to the source of $\rho(\mathcal{T}_{v_3})$.

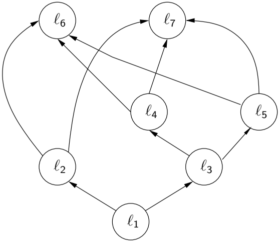




Transforming ternary trees in directed graphs (digraphs)

- $\rho(\textcircled{n}) = \textcircled{n}$.
- Denote by v_1, v_2, v_3 the left, middle and right children of the root of $\mathcal{T}$.
- $\rho(\mathcal{T})$ is defined from $\rho(\mathcal{T}_{v_1})$, $\rho(\mathcal{T}_{v_2})$ and $\rho(\mathcal{T}_{v_3})$ by adding, for each sink v of $\rho(\mathcal{T}_{v_1})$, two edges, one from v to the source of $\rho(\mathcal{T}_{v_2})$ and the other from v to the source of $\rho(\mathcal{T}_{v_3})$.





An inclusion-exclusion formula for the duality function

- Let Σ_* be the set of labelled acyclic digraphs with exactly one source.

An inclusion-exclusion formula for the duality function

- Let Σ_* be the set of labelled acyclic digraphs with exactly one source.
- If $(G, m) \in \Sigma_*$ and $\{\sigma_1, ..., \sigma_n\}$ is the set of strings of G, we define
$$\mathcal{H}(\mathcal{G}, y) := \sum_{\emptyset \neq I \subset [n]} (-1)^{|I|-1} \, y^{m(\sigma_I)},$$
where $m(\sigma_I) :=$ sum of the marks of all nodes covered by $\sigma_I = \cup_{i \in I} \sigma_i$.

An inclusion-exclusion formula for the duality function

- Let Σ_* be the set of labelled acyclic digraphs with exactly one source.
- If $(G, m) \in \Sigma_*$ and $\{\sigma_1, \ldots, \sigma_n\}$ is the set of strings of G, we define

$$\mathcal{H}(\mathcal{G}, y) := \sum_{\emptyset \neq I \subset [n]} (-1)^{|I|-1} \, y^{m(\sigma_I)},$$

where $m(\sigma_I)$:= sum of the marks of all nodes covered by $\sigma_I = \cup_{i \in I} \sigma_i$.

Proposition

For each $\mathcal{T} \in \Upsilon$ and $y \in [0, 1]$, we have

$$H(\mathcal{T}, y) = \mathcal{H}(\rho(\mathcal{T}), y).$$

The line-counting digraph process

Definition

The line-counting digraph process is the continuous time Markov chain $\mathcal{G} := (\mathcal{G}(t))_{t \geq 0}$ with values in $\Sigma_* \cup \{\bowtie\}$, defined by $\mathcal{G}(t) = \rho(\mathcal{T}^*(t))$, $t \geq 0$.

The line-counting digraph process

Definition

The line-counting digraph process is the continuous time Markov chain $\mathcal{G} := (\mathcal{G}(t))_{t \geq 0}$ with values in $\Sigma_* \cup \{\bowtie\}$, defined by $\mathcal{G}(t) = \rho(\mathcal{T}^*(t))$, $t \geq 0$.

- The line-counting digraph process $\mathcal{G}$ is in duality with the deterministic process $1 - z(\cdot, z_0)$ w.r.t $\mathcal{H}$.

The line-counting digraph process

Definition

The line-counting digraph process is the continuous time Markov chain $\mathcal{G} := (\mathcal{G}(t))_{t\geq 0}$ with values in $\Sigma_* \cup \{\bowtie\}$, defined by $\mathcal{G}(t) = \rho(\mathcal{T}^*(t))$, $t \geq 0$.

- The line-counting digraph process $\mathcal{G}$ is in duality with the deterministic process $1 - z(\cdot, z_0)$ w.r.t $\mathcal{H}$.
- The transitions of $\mathcal{T}^*$ can be translated in a natural way to transitions of $\mathcal{G}$.

The line-counting digraph process

Definition

The line-counting digraph process is the continuous time Markov chain $\mathcal{G} := (\mathcal{G}(t))_{t\geq 0}$ with values in $\Sigma_* \cup \{\bowtie\}$, defined by $\mathcal{G}(t) = \rho(\mathcal{T}^*(t))$, $t \geq 0$.

- The line-counting digraph process $\mathcal{G}$ is in duality with the deterministic process $1 - z(\cdot, z_0)$ w.r.t $\mathcal{H}$.
- The transitions of $\mathcal{T}^*$ can be translated in a natural way to transitions of $\mathcal{G}$.
- $\mathcal{G}$ can be constructed independently of $\mathcal{T}^*$.

A series representation

Let $T_{\circledcirc}$ be the time to absorption of $\mathcal{T}^*$ at $\textcircled{0}$ and set $a_0(t) = P(T_{\circledcirc} \leq t)$.

A series representation

Let $T_{\circledcirc}$ be the time to absorption of $\mathcal{T}^*$ at ⓪ and set $a_0(t) = P(T_{\circledcirc} \leq t)$.

Corollary

For $z_0 \in (0, 1)$, we have

$$1 - z(t, z_0) = \sum_{k=0}^{\infty} a_k(t)(1 - z_0)^k,$$

where the coefficients $(a_k(t))_{k\geq 1}$ satisfy

A series representation

Let $T_{\circledcirc}$ be the time to absorption of $\mathcal{T}^*$ at $\textcircled{0}$ and set $a_0(t) = P(T_{\circledcirc} \le t)$.

Corollary

For $z_0 \in (0,1)$, we have

$$1 - z(t, z_0) = \sum_{k=0}^{\infty} a_k(t)(1 - z_0)^k,$$

where the coefficients $(a_k(t))_{k\ge1}$ satisfy

- $a_0'(t) = u_1 a_1(t)$,

A series representation

Let $T_{\circledcirc}$ be the time to absorption of $\mathcal{T}^*$ at $\textcircled{0}$ and set $a_0(t) = P(T_{\circledcirc} \leq t)$.

Corollary

For $z_0 \in (0,1)$, we have

$$1 - z(t, z_0) = \sum_{k=0}^{\infty} a_k(t)(1 - z_0)^k,$$

where the coefficients $(a_k(t))_{k\geq 1}$ satisfy

- $a_0'(t) = u_1 a_1(t)$,
- $a_1'(t) = 2u_1 a_2(t) - (\alpha(\beta - \gamma) + u)a_1(t)$,

A series representation

Let $T_{\circledcirc}$ be the time to absorption of $\mathcal{T}^*$ at $\textcircled{0}$ and set $a_0(t) = P(T_{\circledcirc} \leq t)$.

Corollary

For $z_0 \in (0,1)$, we have

$$1 - z(t, z_0) = \sum_{k=0}^{\infty} a_k(t)(1 - z_0)^k,$$

where the coefficients $(a_k(t))_{k \geq 1}$ satisfy

- $a_0'(t) = u_1 a_1(t)$,
- $a_1'(t) = 2u_1 a_2(t) - (\alpha(\beta - \gamma) + u)a_1(t)$,
- $a_2'(t) = 3u_1 a_3(t) - 2(\alpha(\beta - \gamma) + u)a_2(t) + (\alpha(\beta + 2\gamma))a_1(t)$,

A series representation

Let $T_{\circledcirc}$ be the time to absorption of $\mathcal{T}^*$ at $\textcircled{0}$ and set $a_0(t) = P(T_{\circledcirc} \leq t)$.

Corollary

For $z_0 \in (0,1)$, we have

$$1 - z(t, z_0) = \sum_{k=0}^{\infty} a_k(t)(1 - z_0)^k,$$

where the coefficients $(a_k(t))_{k \geq 1}$ satisfy

- $a_0'(t) = u_1 a_1(t)$,
- $a_1'(t) = 2u_1 a_2(t) - (\alpha(\beta - \gamma) + u)a_1(t)$,
- $a_2'(t) = 3u_1 a_3(t) - 2(\alpha(\beta - \gamma) + u)a_2(t) + (\alpha(\beta + 2\gamma))a_1(t)$,
- $a_k'(t) = (k+1)u_1 a_{k+1}(t) - k(\alpha(\beta - \gamma) + u)a_k(t)$,
  $+(k-1)(\alpha(\beta + 2\gamma))a_{k-1}(t) - (k-2)\alpha\gamma a_{k-2}(t), \quad k \geq 3.$

Absorption probability

- Set $h_0 = 1$, $h_\infty = 0$ and

$$h_k := P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k}\right), \quad k \geq 1.$$

Absorption probability

- Set $h_0 = 1$, $h_\infty = 0$ and

$$h_k := P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k}\right), \quad k \geq 1.$$

- A first step analysis leads to

$$\begin{aligned}(u + \alpha(\beta - \gamma))h_k &= \alpha\beta h_{k+1} + u\nu_1 h_{k-1} \\ &\quad - \alpha\gamma P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k} \otimes \textcircled{1} \otimes \textcircled{1}\right).\end{aligned}$$

Absorption probability

- Set $h_0 = 1$, $h_\infty = 0$ and

$$h_k := P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k}\right), \quad k \geq 1.$$

- A first step analysis leads to

$$\begin{aligned}(u + \alpha(\beta - \gamma))h_k &= \alpha\beta h_{k+1} + u\nu_1 h_{k-1} \\ &\quad - \alpha\gamma P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k} \otimes \textcircled{1} \otimes \textcircled{1}\right).\end{aligned}$$

- One can show that

$$P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k} \otimes \textcircled{1} \otimes \textcircled{1}\right) = h_k(2h_1 - h_1^2),$$

Absorption probability

- Set $h_0 = 1$, $h_\infty = 0$ and

$$h_k := P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k}\right), \quad k \geq 1.$$

- A first step analysis leads to

$$\begin{aligned}(u + \alpha(\beta - \gamma))h_k &= \alpha\beta h_{k+1} + u\nu_1 h_{k-1} \\ &\quad - \alpha\gamma P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k} \otimes \textcircled{1} \otimes \textcircled{1}\right).\end{aligned}$$

- One can show that

$$P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k} \otimes \textcircled{1} \otimes \textcircled{1}\right) = h_k(2h_1 - h_1^2),$$

$\Rightarrow$ h_1 is the root in $[0, 1]$ of a cubic polynomial.

Absorption probability

- Set $h_0 = 1$, $h_\infty = 0$ and

$$h_k := P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k}\right), \quad k \geq 1.$$

- A first step analysis leads to

$$\begin{aligned}(u + \alpha(\beta - \gamma))h_k = \alpha\beta h_{k+1} + u\nu_1 h_{k-1} \\ - \alpha\gamma P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k} \otimes \textcircled{1} \otimes \textcircled{1}\right).\end{aligned}$$

- One can show that

$$P\left(T_{\circledcirc} < \infty | \mathcal{T}^*(0) = \textcircled{k} \otimes \textcircled{1} \otimes \textcircled{1}\right) = h_k(2h_1 - h_1^2),$$

$\Rightarrow$ h_1 is the root in $[0, 1]$ of a cubic polynomial.

- Duality $\Rightarrow$ h_1 = stationary proportion of $1'$s in the forward model.

Thank you for your attention!