Nested coalescents

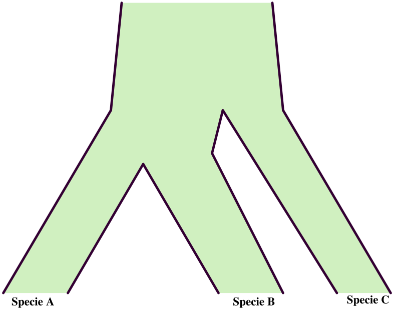
Airam Blancas Benítez
Joint work with: Amaury Lambert and Arno Siri-Jégousse

CIMAT

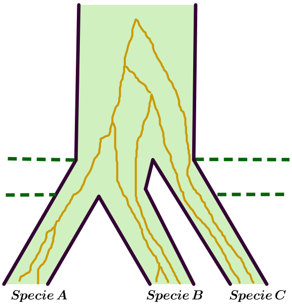
May 16, 2016

Outline

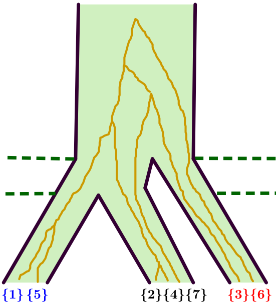
Species trees



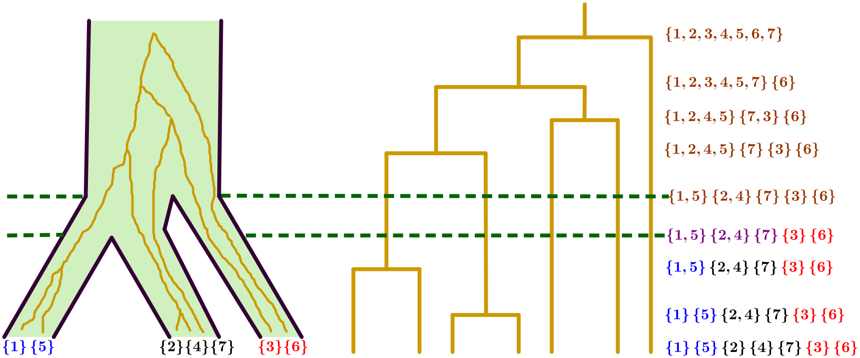
Species trees and Gene trees



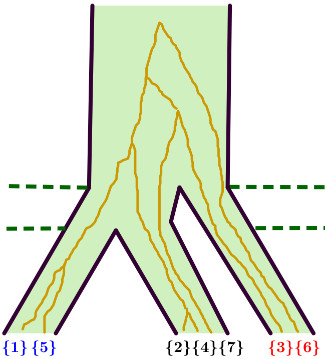
Genes trees

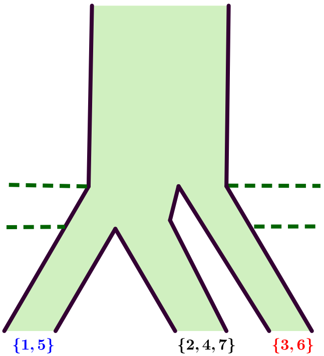


Genes trees

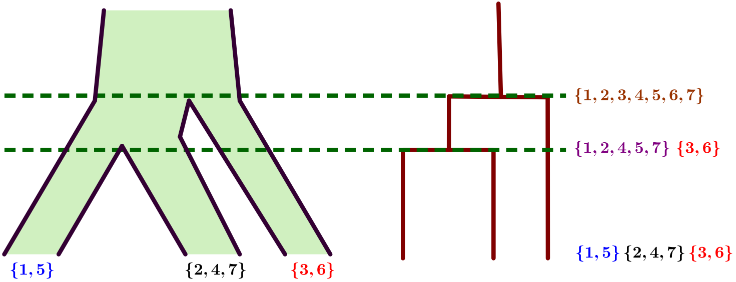


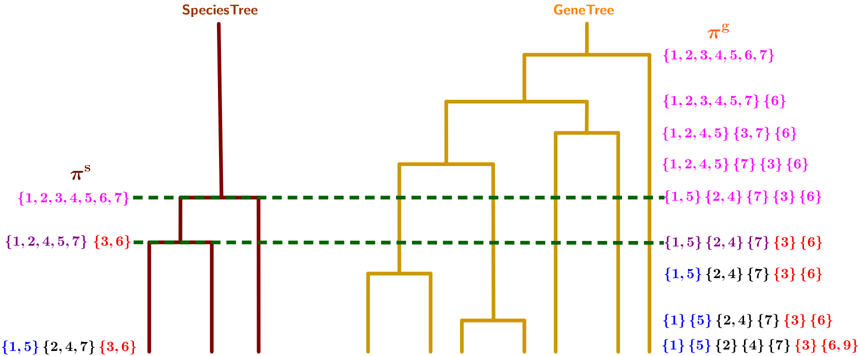
{1} {5}
{2}{4}{7}
{3}{6}
{1, 2, 3, 4, 5, 6, 7}
{1, 2, 3, 4, 5, 7} {6}
{1, 2, 4, 5} {7, 3} {6}
{1, 2, 4, 5} {7} {3} {6}
{1, 5} {2, 4} {7} {3} {6}
{1, 5} {2, 4} {7} {3} {6}
{1, 5} {2, 4} {7} {3} {6}
{1} {5} {2, 4} {7} {3} {6}
{1} {5} {2} {4} {7} {3} {6}

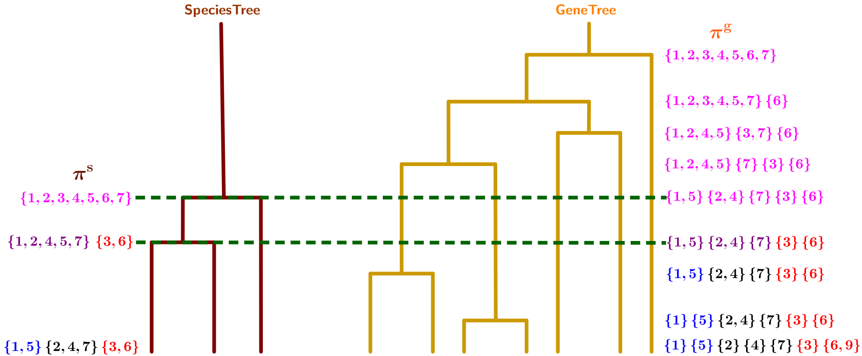

{1} {5}
{2}{4}{7}
{3}{6}


{1, 5}
{2, 4, 7}
{3, 6}

Species trees




SpeciesTree
GeneTree
π^g
$\{1,2,3,4,5,6,7\}$
$\{1,2,3,4,5,7\}\ \{6\}$
$\{1,2,4,5\}\ \{3,7\}\ \{6\}$
$\{1,2,4,5\}\ \{7\}\ \{3\}\ \{6\}$
$\{1,5\}\ \{2,4\}\ \{7\}\ \{3\}\ \{6\}$
$\{1,5\}\ \{2,4\}\ \{7\}\ \{3\}\ \{6\}$
$\{1,5\}\ \{2,4\}\ \{7\}\ \{3\}\ \{6\}$
$\{1\}\ \{5\}\ \{2,4\}\ \{7\}\ \{3\}\ \{6\}$
$\{1\}\ \{5\}\ \{2\}\ \{4\}\ \{7\}\ \{3\}\ \{6,9\}$
π^s
$\{1,2,3,4,5,6,7\}$
$\{1,2,4,5,7\}\ \{3,6\}$
$\{1,5\}\ \{2,4,7\}\ \{3,6\}$



We are interested in coagulating processes with values in $\mathcal{N}_n$:= the subset of **nested partitions** $\pi = (\pi^{\mathrm{s}}, \pi^{\mathrm{g}})$ of $[n]^2$.

Definition

Fix $n \in \bar{\mathbb{N}}$, for every $t \geq 0$ let $\mathcal{R}(t) := ((\mathcal{R}^s(t), \mathcal{R}^g(t)) : t \geq 0)$ be a Markov process with values in $\mathcal{N}_n$. This process is called **simple nested exchangeable coalescent**, snec for short, if

i) For any $t \geq 0$, $\mathcal{R}^g(t)$ and $\mathcal{R}^s(t)$ are exchangeable random partitions.

Definition

Fix $n \in \bar{\mathbb{N}}$, for every $t \geq 0$ let $\mathcal{R}(t) := ((\mathcal{R}^s(t), \mathcal{R}^g(t)) : t \geq 0)$ be a Markov process with values in $\mathcal{N}_n$. This process is called **simple nested exchangeable coalescent**, snec for short, if

i) For any $t \geq 0$, $\mathcal{R}^g(t)$ and $\mathcal{R}^s(t)$ are exchangeable random partitions.

ii) The process $(\mathcal{R}^s(t) : t \geq 0)$ is a simple exchangeable coalescent process.

Definition

Fix $n \in \bar{\mathbb{N}}$, for every $t \geq 0$ let $\mathcal{R}(t) := ((\mathcal{R}^s(t), \mathcal{R}^g(t)) : t \geq 0)$ be a Markov process with values in $\mathcal{N}_n$. This process is called **simple nested exchangeable coalescent**, snec for short, if

i) For any $t \geq 0$, $\mathcal{R}^g(t)$ and $\mathcal{R}^s(t)$ are exchangeable random partitions.

ii) The process $(\mathcal{R}^s(t) : t \geq 0)$ is a simple exchangeable coalescent process.

iii) Conditional on $(\mathcal{R}^s(t) : t \geq 0)$, the process $(\mathcal{R}^g(t) : t \geq 0)$ restricted to a single species is a simple exchangeable coalescent process.

Outline

Characterization of simple exchangeable coalescents

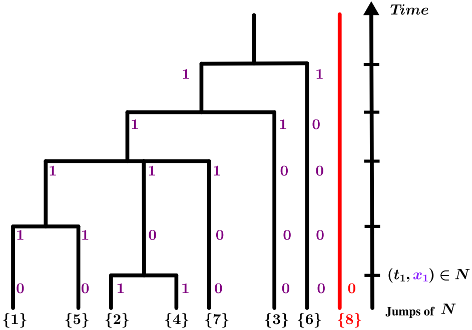
There exists an array of numbers $(\lambda_{b,k})_{2\leq k\leq b}$ which gives us the rate at which any fixed k-tuple of blocks merges when there are b-blocks in total such that

$$\lambda_{b,k} = \lambda_{b+1,k} + \lambda_{b+1,k+1}.$$

Characterization of simple exchangeable coalescents

There exists an array of numbers $(\lambda_{b,k})_{2\leq k\leq b}$ which gives us the rate at which any fixed k-tuple of blocks merges when there are b-blocks in total such that

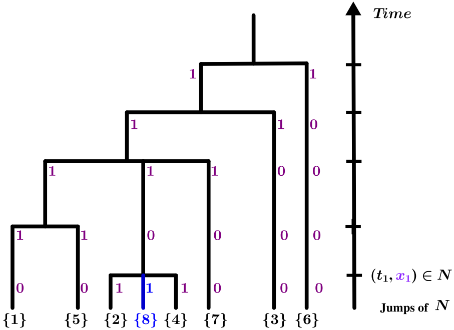
$$\lambda_{b,k} = \lambda_{b+1,k} + \lambda_{b+1,k+1}.$$



Characterization of simple exchangeable coalescents

There exists an array of numbers $(\lambda_{b,k})_{2\leq k\leq b}$ which gives us the rate at which any fixed k-tuple of blocks merges when there are b-blocks in total such that

$$\lambda_{b,k} = \lambda_{b+1,k} + \lambda_{b+1,k+1}.$$



Characterization of snec

- Denote $\mathbf{g} = (g_1, \ldots, g_s)$ and $\mathbf{c} = (c_1, \ldots, c_k)$.
- Let $q(\mathbf{g}, \mathbf{c})$ be the rate at which a k-tuple of those s species **merge** into one, involving c_i genes **in the** i**-th species merging**.

Characterization of snec

- Denote $\mathbf{g} = (g_1, \dots, g_s)$ and $\mathbf{c} = (c_1, \dots, c_k)$.
- Let $q(\mathbf{g}, \mathbf{c})$ be the rate at which a k-tuple of those s species **merge** into one, involving c_i genes **in** the i**-th species merging**.

The jump rate when at least two species merge is given by

$$q(\mathbf{g}, \mathbf{c}) = \int_E x^{k-2}(1-x)^{s-k} \left(\int_{(0,1)} \prod_{i=1}^{k} y_i^{c_i}(1-y_i)^{g_i - c_i} G(dy_1) \cdots G(dy_k) \right) \Sigma(dx, dG), \quad k \geq 2,$$

where Σ is a finite measure on $E := [0,1] \times \mathcal{M}_1[0,1]$.

Characterization of snec

- Denote $\mathbf{g} = (g_1, \ldots, g_s)$ and $\mathbf{c} = (c_1, \ldots, c_k)$.
- Let $q(\mathbf{g}, \mathbf{c})$ be the rate at which a k-tuple of those s species **merge** into one, involving c_i genes **in** the i-**th species merging**.

The jump rate when at least two genes merge in one “activated” specie

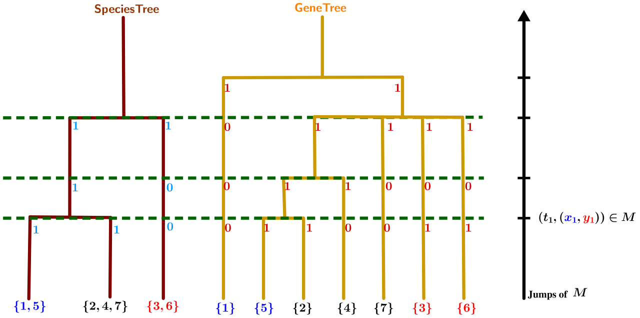
$$q(\mathbf{g}, \mathbf{c}) = \int_E (1-x)^{s-1} \left(\int_{(0,1)} y^{c_1-2}(1-y)^{g_1-c_1} G(dy) \right) \Sigma(dx, dG), \quad c_1 \geq 2,$$

where Σ is a finite measure on $E := [0,1] \times \mathcal{M}_1[0,1]$.

Outline

Poisson construction

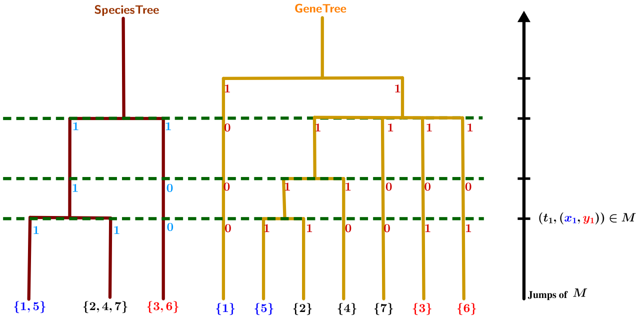
Let $(t_i, (x_i, y_i))_i$ be atoms of a Poisson point process M on $(0, \infty) \times (0,1)^2$ of intensity $dt \otimes \nu_{sg}(dx)$.



- At time t_i each *species* tosses a coin with heads probability x_i.
- All the *species* present at time t_i^- getting *heads* **merge** into one.

Poisson construction of snec

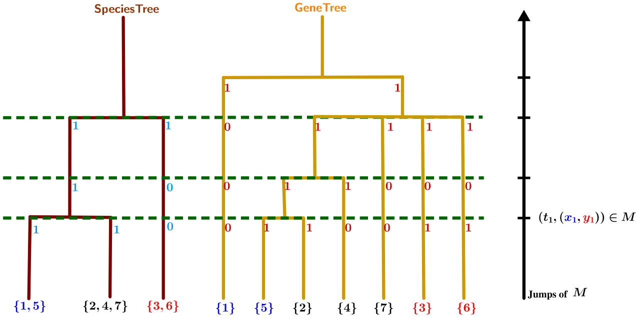
Let $(t_i, (x_i, y_i))_i$ be atoms of a Poisson point process M on $(0, \infty) \times (0, 1)^2$ of intensity $dt \otimes \nu_{sg}(dx, dy)$.



- At time t_i each *gene* tosses a coin with heads probability y_i.
- If the *genes* are **inside** *species* getting *heads*, then all genes getting *heads* **merge** into one.

Poisson construction of snec

Let $(t_i, (x_i, y_i))_i$ be atoms of a Poisson point process M on $(0, \infty) \times (0, 1)^2$ of intensity $dt \otimes \nu_{sg}(dx, dy)$.



- At time t_i each *gene* tosses a coin with heads probability y_i.
- If the *genes* are **inside** *species* getting *heads*, then all genes getting *heads* **merge** into one.
- To get non-degenerate construction we assume $\int_{[0,1]^2} (x^2 + xy^2) \nu_{sg}(dx, dy) < \infty$.

Binary merging in the snec

Recall $\mathbf{g} = (g_1, \dots, g_s)$ and $\mathbf{c} = (c_1, \dots, c_k)$.
Let us suppose that $\Sigma(dx, dG) = \delta_0 \times \delta_{\delta_0}$. Then

- $q(\mathbf{g}, \mathbf{c}) = \mathbb{1}_{\{k=2\}} \mathbb{1}_{\{c_1=0, c_2=0\}} + \mathbb{1}_{\{k=1\}} \mathbb{1}_{\{c_1=2\}}$

Binary merging in the snec

Recall $\mathbf{g} = (g_1, \ldots, g_s)$ and $\mathbf{c} = (c_1, \ldots, c_k)$.
Let us suppose that $\Sigma(dx, dG) = \delta_0 \times \delta_{\delta_0}$. Then

- $q(\mathbf{g}, \mathbf{c}) = \mathbb{1}_{\{k=2\}} \mathbb{1}_{\{c_1=0, c_2=0\}} + \mathbb{1}_{\{k=1\}} \mathbb{1}_{\{c_1=2\}}$

We do not have simultaneous binary merging in the species and binary merging in the genes.

Further questions

1. Which are classical measures in $(0,1)^2$ for defining "classical snec"?

Further questions

1. Which are classical measures in $(0,1)^2$ for defining "classical snec"?
2. To find some forward in time evolutionary models with (limit) genealogies being snec processes.
   - Discrete models
   - Valued measure models
   - Partition flows

Thank you!