

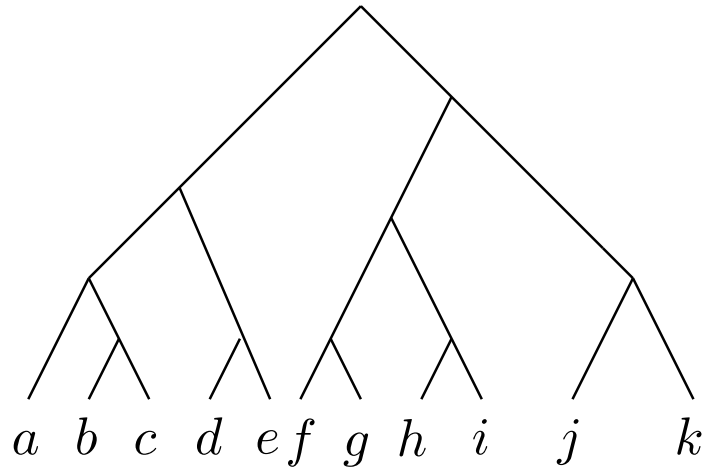
The Age of a Vertex

Tanja Gernhard

February 10, 2006

The Age of a Vertex

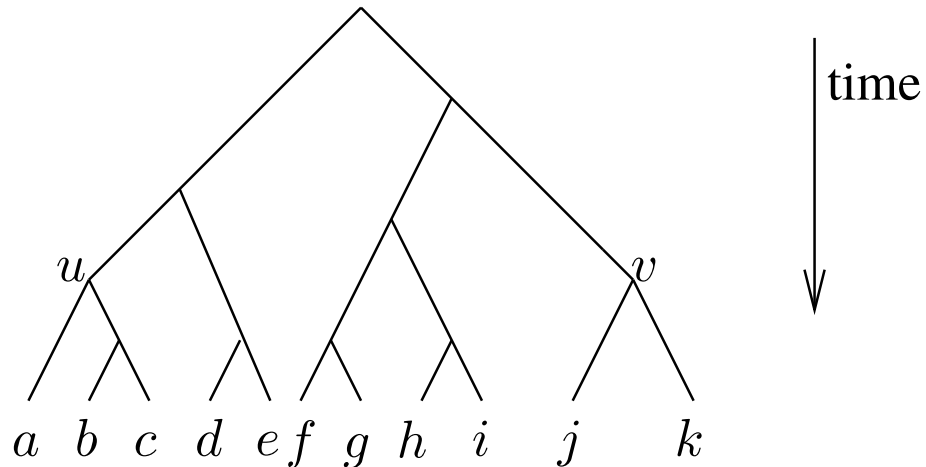
A tree displays the relationships between species.



The Age of a Vertex

A tree displays the relationships between species.

Can we say something about the relative time?

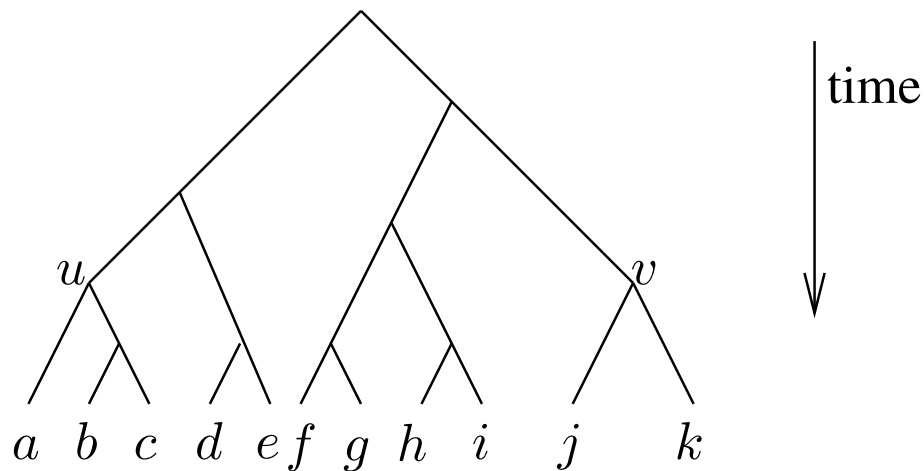


The Age of a Vertex

A tree displays the relationships between species.

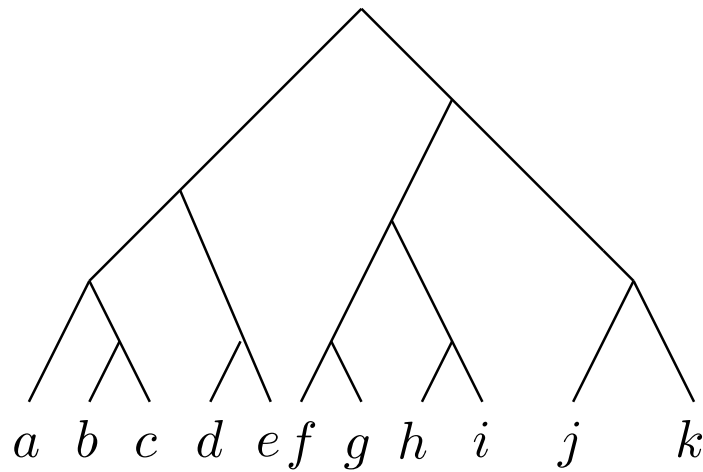
Can we say something about the relative time?

Is an inner vertex an early or late event compared to the other vertices?



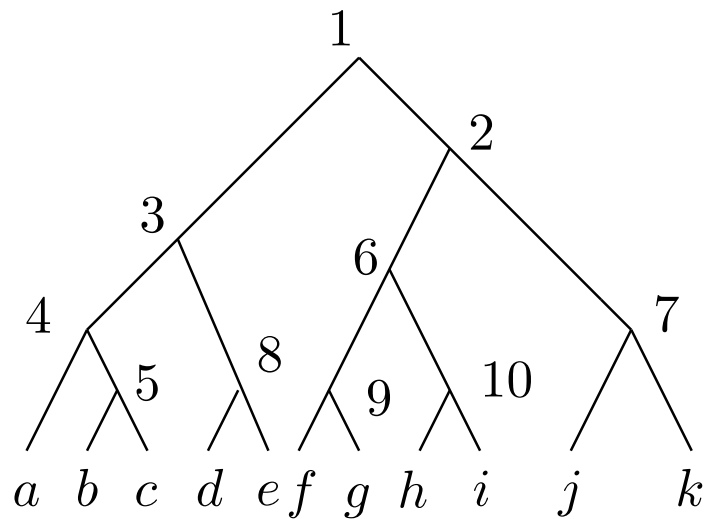
The Rank Function

We want to put the inner vertices in a linear order.



The Rank Function

We want to put the inner vertices in a linear order.



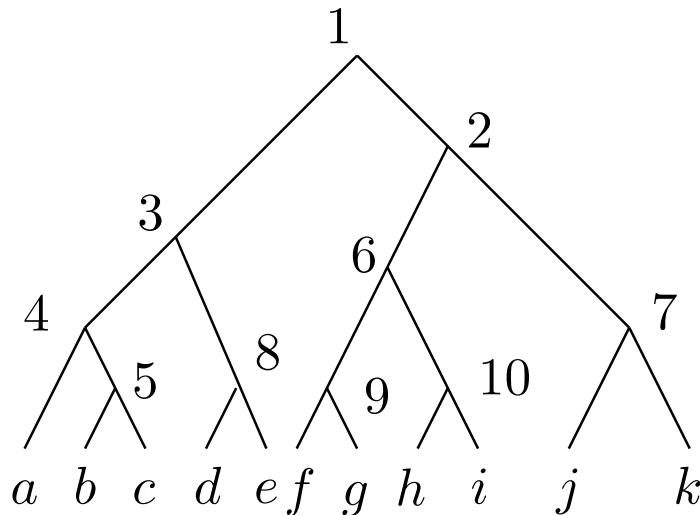
The Rank Function

We want to put the inner vertices in a linear order.

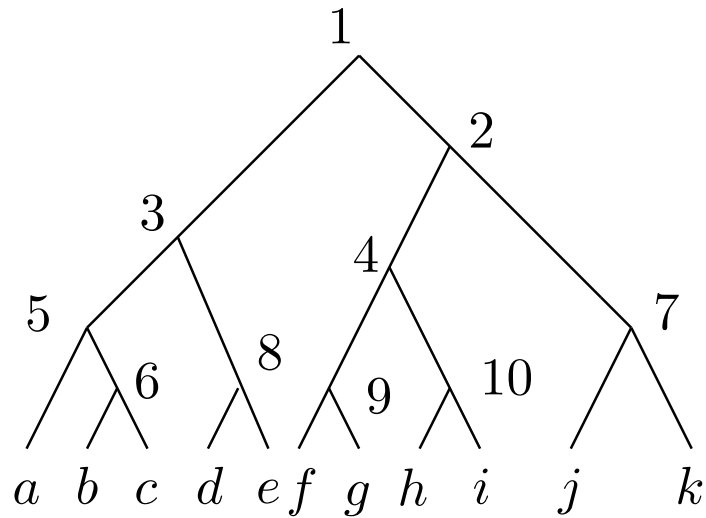
A rank function r on a tree induces a linear order on the inner vertices:

$$r : \overset{\circ}{V} \rightarrow \{1, \dots, |\overset{\circ}{V}|\}$$

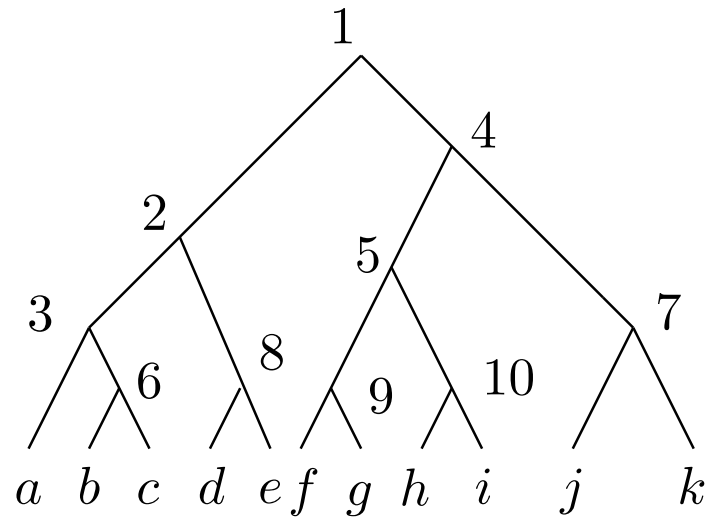
r bijection with $r(v) < r(w)$ if vertex v is an ancestor of w .



The Rank Function



The Rank Function



The Yule Model

We need a stochastic model for speciation - we chose the Yule model.

The Yule Model

We need a stochastic model for speciation - we chose the Yule model.

Under the Yule model, each pendant edge is equally likely to split.

The Yule Model

We need a stochastic model for speciation - we chose the Yule model.

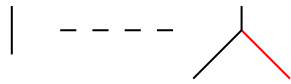
Under the Yule model, each pendant edge is equally likely to split.



The Yule Model

We need a stochastic model for speciation - we chose the Yule model.

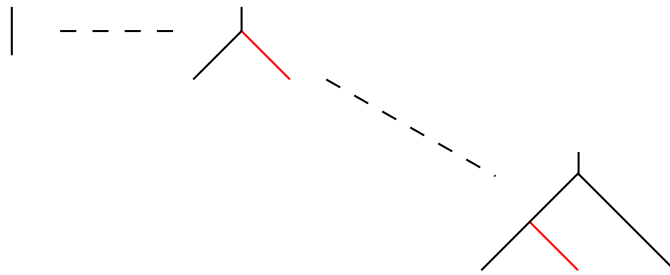
Under the Yule model, each pendant edge is equally likely to split.



The Yule Model

We need a stochastic model for speciation - we chose the Yule model.

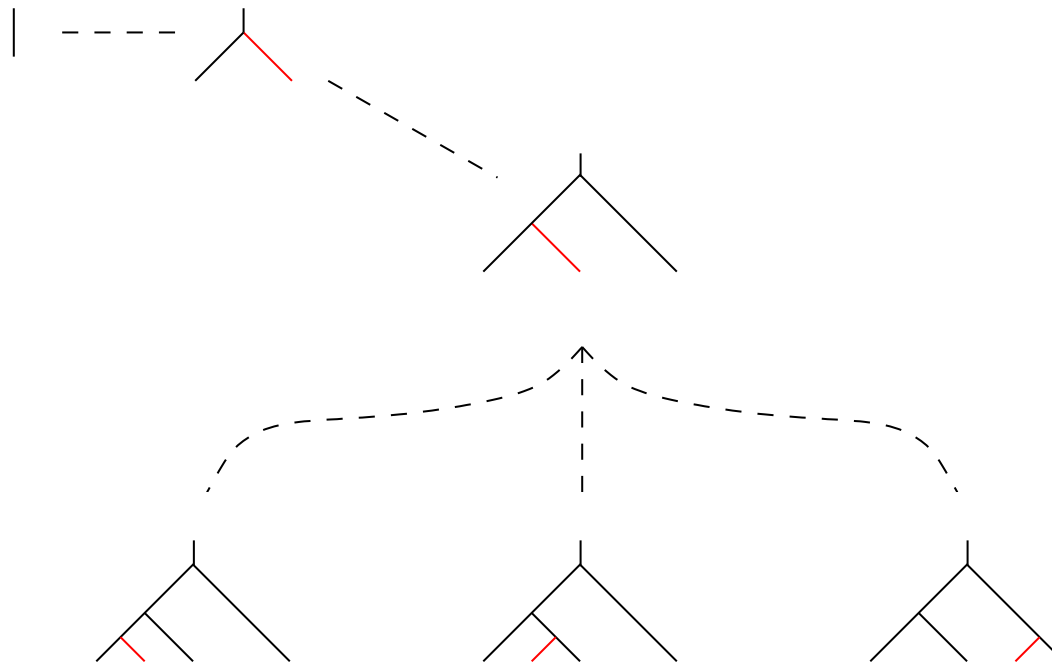
Under the Yule model, each pendant edge is equally likely to split.



The Yule Model

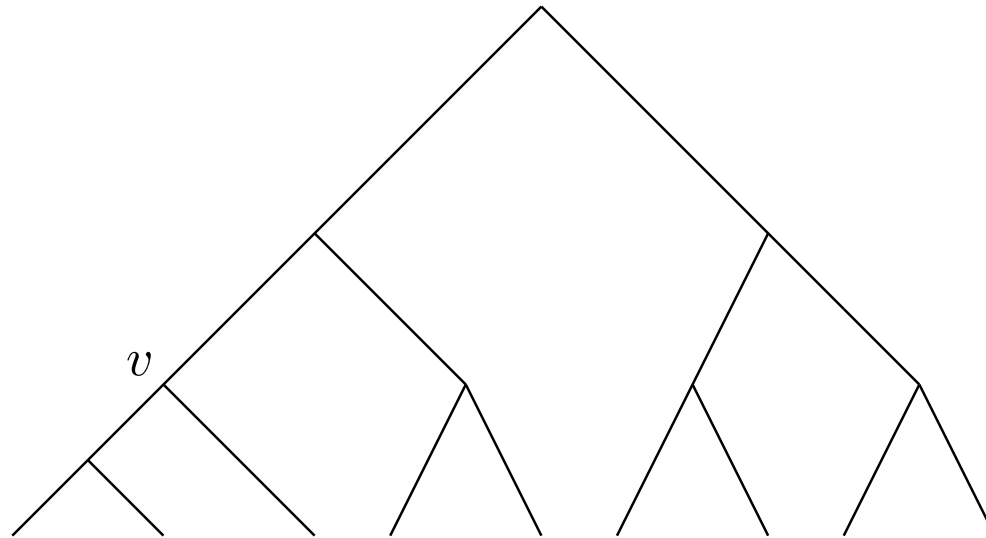
We need a stochastic model for speciation - we chose the Yule model.

Under the Yule model, each pendant edge is equally likely to split.



Rank of a Vertex

Under the Yule model, how likely is it that vertex v has rank 5, i.e. $r(v) = 5$?



Rank of a Vertex

- Under the Yule model, each ranked phylogenetic tree is equally likely.

Rank of a Vertex

- Under the Yule model, each ranked phylogenetic tree is equally likely.
- That means, $\mathbb{P}[r(v) = i] = \frac{|\{r:r(v)=i,r \in r(\mathcal{T})\}|}{|r(\mathcal{T})|}$
where $r(\mathcal{T})$ is the set of all rank functions on \mathcal{T} .

Rank of a Vertex

- Under the Yule model, each ranked phylogenetic tree is equally likely.
- That means, $\mathbb{P}[r(v) = i] = \frac{|\{r:r(v)=i,r \in r(\mathcal{T})\}|}{|r(\mathcal{T})|}$
where $r(\mathcal{T})$ is the set of all rank functions on \mathcal{T} .
- It is known that $|r(\mathcal{T})| = \frac{|\mathring{V}|!}{\prod_{v \in \mathring{V}} \lambda_v} =: R_{\mathcal{T}}$.

Rank of a Vertex

- Under the Yule model, each ranked phylogenetic tree is equally likely.
- That means, $\mathbb{P}[r(v) = i] = \frac{|\{r:r(v)=i,r \in r(\mathcal{T})\}|}{|r(\mathcal{T})|}$
where $r(\mathcal{T})$ is the set of all rank functions on \mathcal{T} .
- It is known that $|r(\mathcal{T})| = \frac{|\mathring{V}|!}{\prod_{v \in \mathring{V}} \lambda_v} =: R_{\mathcal{T}}$.
- We developed a quadratic algorithm to calculate $|\{r : r(v) = i, r \in r(\mathcal{T})\}| =: \alpha_{\mathcal{T},v}(i)$.

Rank of a Vertex

- Under the Yule model, each ranked phylogenetic tree is equally likely.
- That means, $\mathbb{P}[r(v) = i] = \frac{|\{r:r(v)=i,r \in r(\mathcal{T})\}|}{|r(\mathcal{T})|}$
where $r(\mathcal{T})$ is the set of all rank functions on \mathcal{T} .
- It is known that $|r(\mathcal{T})| = \frac{|\mathring{V}|!}{\prod_{v \in \mathring{V}} \lambda_v} =: R_{\mathcal{T}}$.
- We developed a quadratic algorithm to calculate $|\{r : r(v) = i, r \in r(\mathcal{T})\}| =: \alpha_{\mathcal{T},v}(i)$.

Algorithm RANKCOUNTBIN

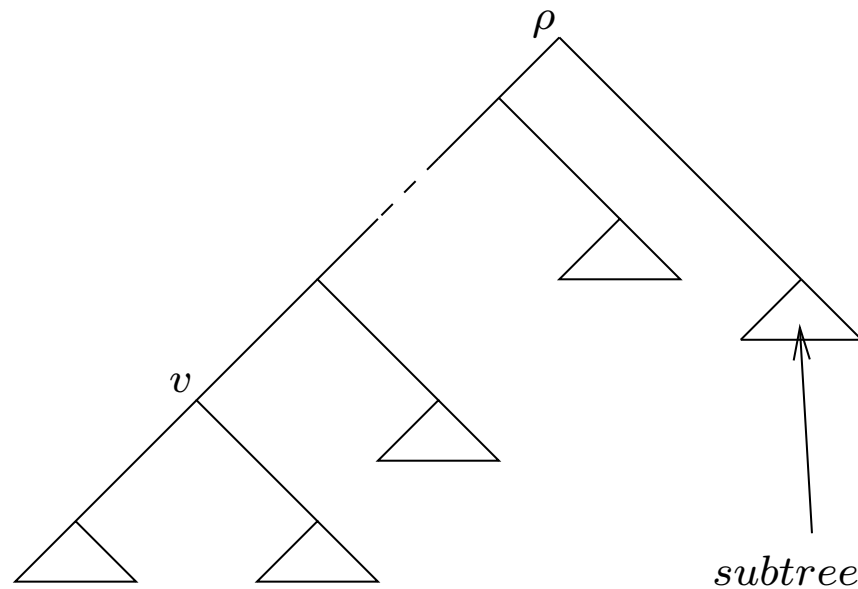
- **Input:** A rooted binary phylogenetic tree \mathcal{T} and an interior vertex v .

Algorithm RANKCOUNTBIN

- **Input:** A rooted binary phylogenetic tree \mathcal{T} and an interior vertex v .
- **Output:** The values of $\alpha_{\mathcal{T},v}(i)$ for $i = 1, \dots, |\mathring{V}|$.

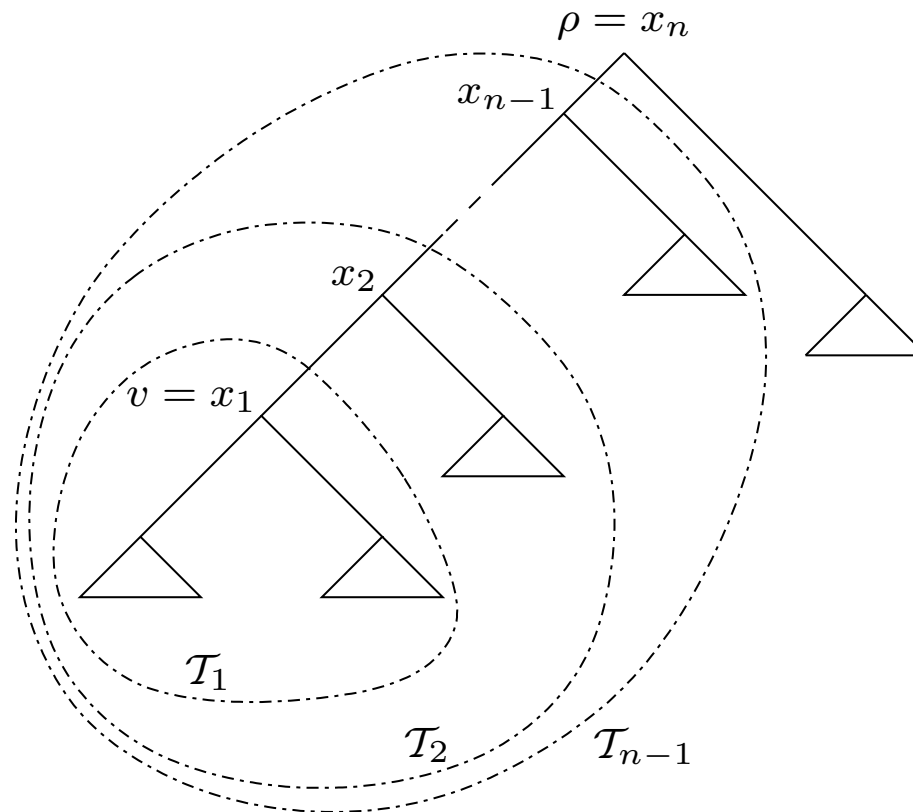
Algorithm RANKCOUNTBIN

- **Input:** A rooted binary phylogenetic tree \mathcal{T} and an interior vertex v .
- **Output:** The values of $\alpha_{\mathcal{T},v}(i)$ for $i = 1, \dots, |\dot{V}|$.



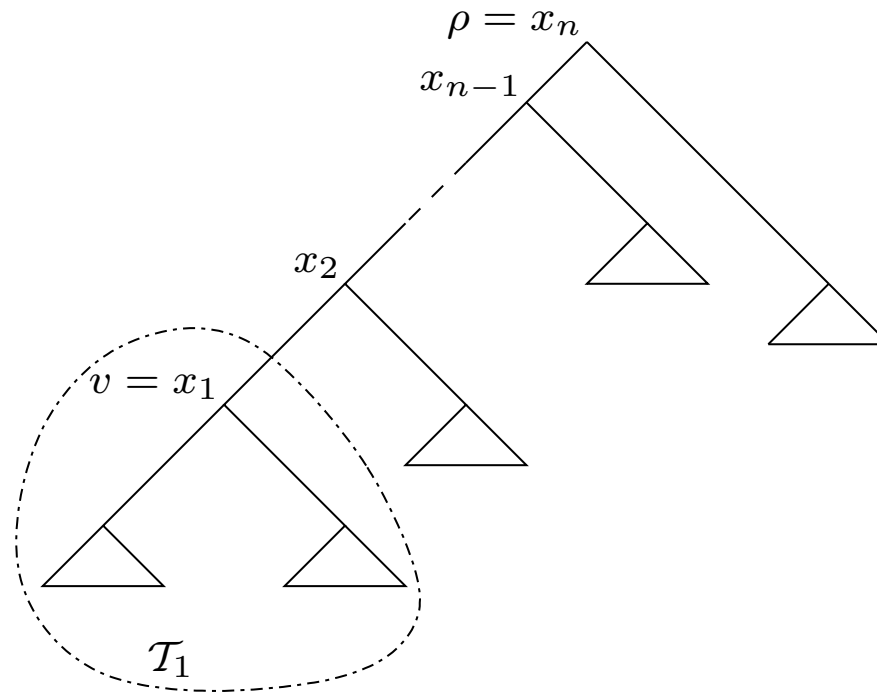
Algorithm RANKCOUNTBIN

- **Input:** A rooted binary phylogenetic tree \mathcal{T} and an interior vertex v .
- **Output:** The values of $\alpha_{\mathcal{T},v}(i)$ for $i = 1, \dots, |\mathring{V}|$.



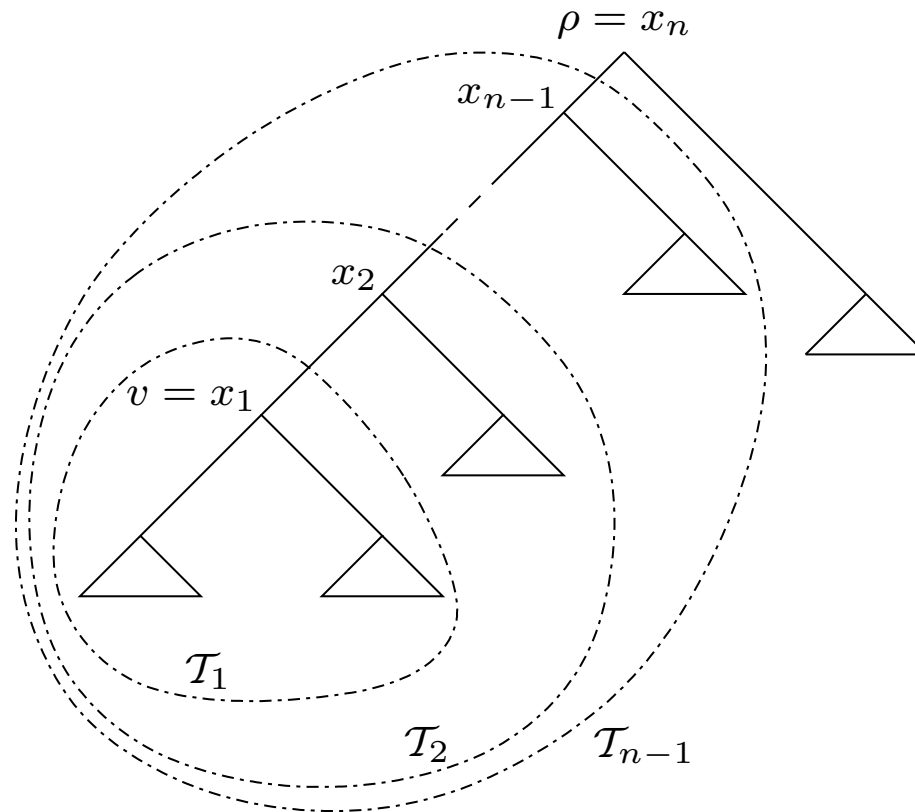
Algorithm RANKCOUNTBIN - The Recursion

$$\alpha_{\mathcal{T}_1, v}(1) := \frac{|\dot{V}_{\mathcal{T}_1}|!}{\prod_{v \in \dot{V}_{\mathcal{T}_1}} \lambda_v}$$



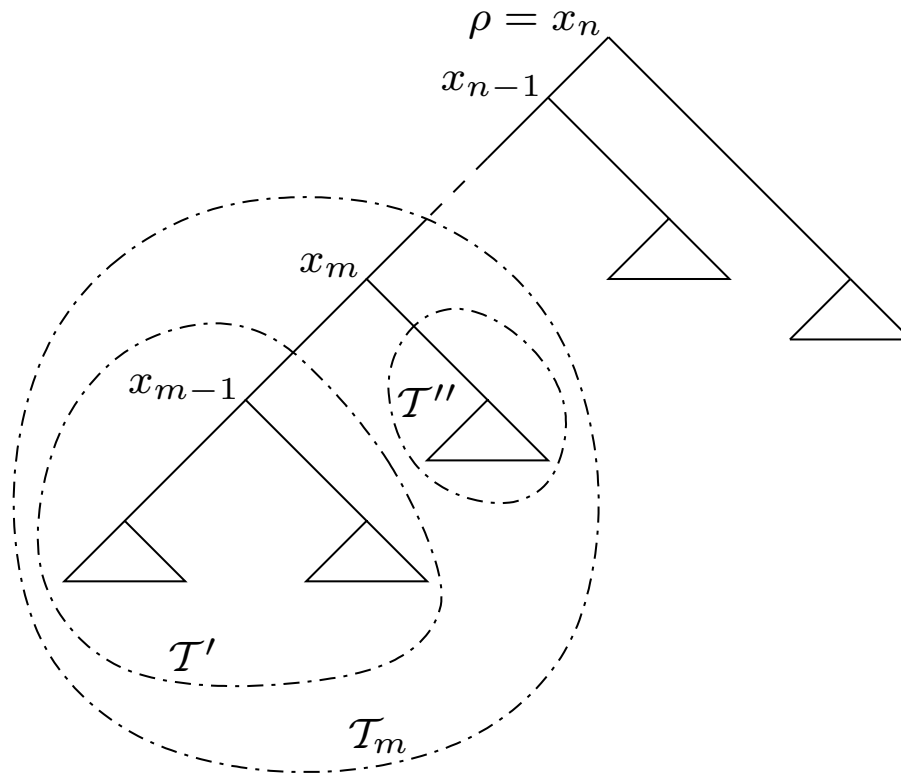
Algorithm RANKCOUNTBIN - The Recursion

$$\alpha_{\mathcal{T}_m, v}(i) := \sum_{j=0}^{\min\{i-2, |\dot{V}_{\mathcal{T}''}|\}} \alpha_{\mathcal{T}', v}(i-j-1) R_{\mathcal{T}''} \left(\begin{array}{c} |\dot{V}_{\mathcal{T}'}| + |\dot{V}_{\mathcal{T}''}| - (i-1) \\ |\dot{V}_{\mathcal{T}''}| - j \end{array} \right) \binom{i-2}{j}$$



Algorithm RANKCOUNTBIN - The Recursion

$$\alpha_{\mathcal{T}_m, v}(i) := \sum_{j=0}^{\min\{i-2, |\dot{V}_{\mathcal{T}''}|\}} \alpha_{\mathcal{T}', v}(i-j-1) R_{\mathcal{T}''} \left(\begin{matrix} |\dot{V}_{\mathcal{T}'}| + |\dot{V}_{\mathcal{T}''}| - (i-1) \\ |\dot{V}_{\mathcal{T}''}| - j \end{matrix} \right) \binom{i-2}{j}$$



Probability of $r(v) = i$

Our goal was to calculate $\mathbb{P}[r(v) = i]$.

Probability of $r(v) = i$

Our goal was to calculate $\mathbb{P}[r(v) = i]$.

Since each ranked tree is equally likely, we have

$$\mathbb{P}[r(v) = i] = \frac{\alpha_{\mathcal{T},v}(i)}{R_{\mathcal{T}}}.$$

Probability of $r(v) = i$

Our goal was to calculate $\mathbb{P}[r(v) = i]$.

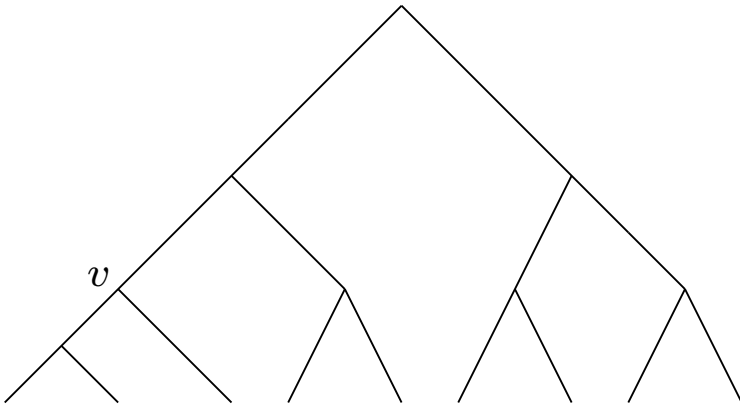
Since each ranked tree is equally likely, we have

$$\mathbb{P}[r(v) = i] = \frac{\alpha_{\mathcal{T},v}(i)}{R_{\mathcal{T}}}.$$

Example

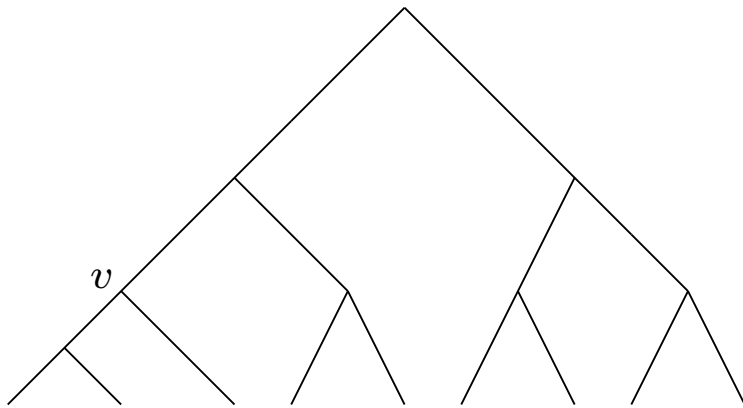
Example

Consider the following tree:



Example

Consider the following tree:



$$\mathbb{P}[r(v) = 1] = 0$$

$$\mathbb{P}[r(v) = 2] = 0$$

$$\mathbb{P}[r(v) = 3] = \frac{20}{93}$$

$$\mathbb{P}[r(v) = 4] = \frac{16}{93}$$

$$\mathbb{P}[r(v) = 5] = \frac{27}{93}$$

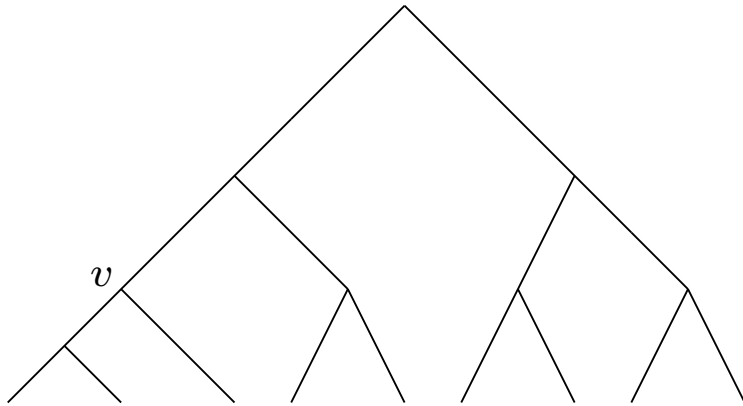
$$\mathbb{P}[r(v) = 6] = \frac{20}{93}$$

$$\mathbb{P}[r(v) = 7] = \frac{10}{93}$$

$$\mathbb{P}[r(v) = 8] = 0$$

Example

Consider the following tree:



$$\mu_{r(v)} = \sum_{i=1}^8 i \mathbb{P}[r(v) = i] \approx 4.83$$

$$\sigma_{r(v)}^2 = \sum_{i=1}^8 i^2 \mathbb{P}[r(v) = i] - \mu_{r(v)}^2 \approx 1.65$$

Extensions of RANKCOUNTBIN

- RANKCOUNTBIN can be generalized to RANKCOUNT for non-binary trees.

Extensions of RANKCOUNTBIN

- RANKCOUNTBIN can be generalized to RANKCOUNT for non-binary trees.
The runtime is $O(|V|^d)$.

Extensions of RANKCOUNTBIN

- RANKCOUNTBIN can be generalized to RANKCOUNT for non-binary trees.
The runtime is $O(|V|^d)$.
- Algorithm COMPARE compares two inner vertices - which one is more likely to be earlier?

Extensions of RANKCOUNTBIN

- RANKCOUNTBIN can be generalized to RANKCOUNT for non-binary trees.
The runtime is $O(|V|^d)$.
- Algorithm COMPARE compares two inner vertices - which one is more likely to be earlier?
COMPARE runs in $O(|V|^2)$ and uses RANKCOUNTBIN and RANKCOUNT.

Testing a tree to be Yule

RANKCOUNTBIN works for trees which evolved under the Yule model.

Testing a tree to be Yule

RANKCOUNTBIN works for trees which evolved under the Yule model.

We test the Yule model against the uniform model (PDA model) with the log-likelihood ratio test, i.e.

- $\ln \frac{\mathbb{P}_Y[\mathcal{T}]}{\mathbb{P}_U[\mathcal{T}]} > 0 \Rightarrow \text{accept Yule.}$
- $\ln \frac{\mathbb{P}_Y[\mathcal{T}]}{\mathbb{P}_U[\mathcal{T}]} \leq 0 \Rightarrow \text{reject Yule.}$

Testing a tree to be Yule

RANKCOUNTBIN works for trees which evolved under the Yule model.

We test the Yule model against the uniform model (PDA model) with the log-likelihood ratio test, i.e.

- $\ln \frac{\mathbb{P}_Y[\mathcal{T}]}{\mathbb{P}_U[\mathcal{T}]} > 0 \Rightarrow \text{accept Yule.}$
- $\ln \frac{\mathbb{P}_Y[\mathcal{T}]}{\mathbb{P}_U[\mathcal{T}]} \leq 0 \Rightarrow \text{reject Yule.}$

The Type I and Type II error of this test can be obtained by simulations.

Testing a tree to be Yule

RANKCOUNTBIN works for trees which evolved under the Yule model.

We test the Yule model against the uniform model (PDA model) with the log-likelihood ratio test, i.e.

- $\ln \frac{\mathbb{P}_Y[\mathcal{T}]}{\mathbb{P}_U[\mathcal{T}]} > 0 \Rightarrow \text{accept Yule.}$
- $\ln \frac{\mathbb{P}_Y[\mathcal{T}]}{\mathbb{P}_U[\mathcal{T}]} \leq 0 \Rightarrow \text{reject Yule.}$

The Type I and Type II error of this test can be obtained by simulations.

We found analytical bounds by introducing martingales and applying the Azuma inequality.

Thank you - Dankeschön

