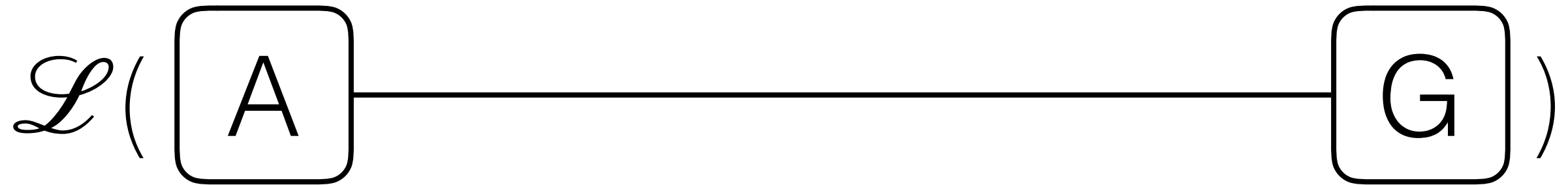
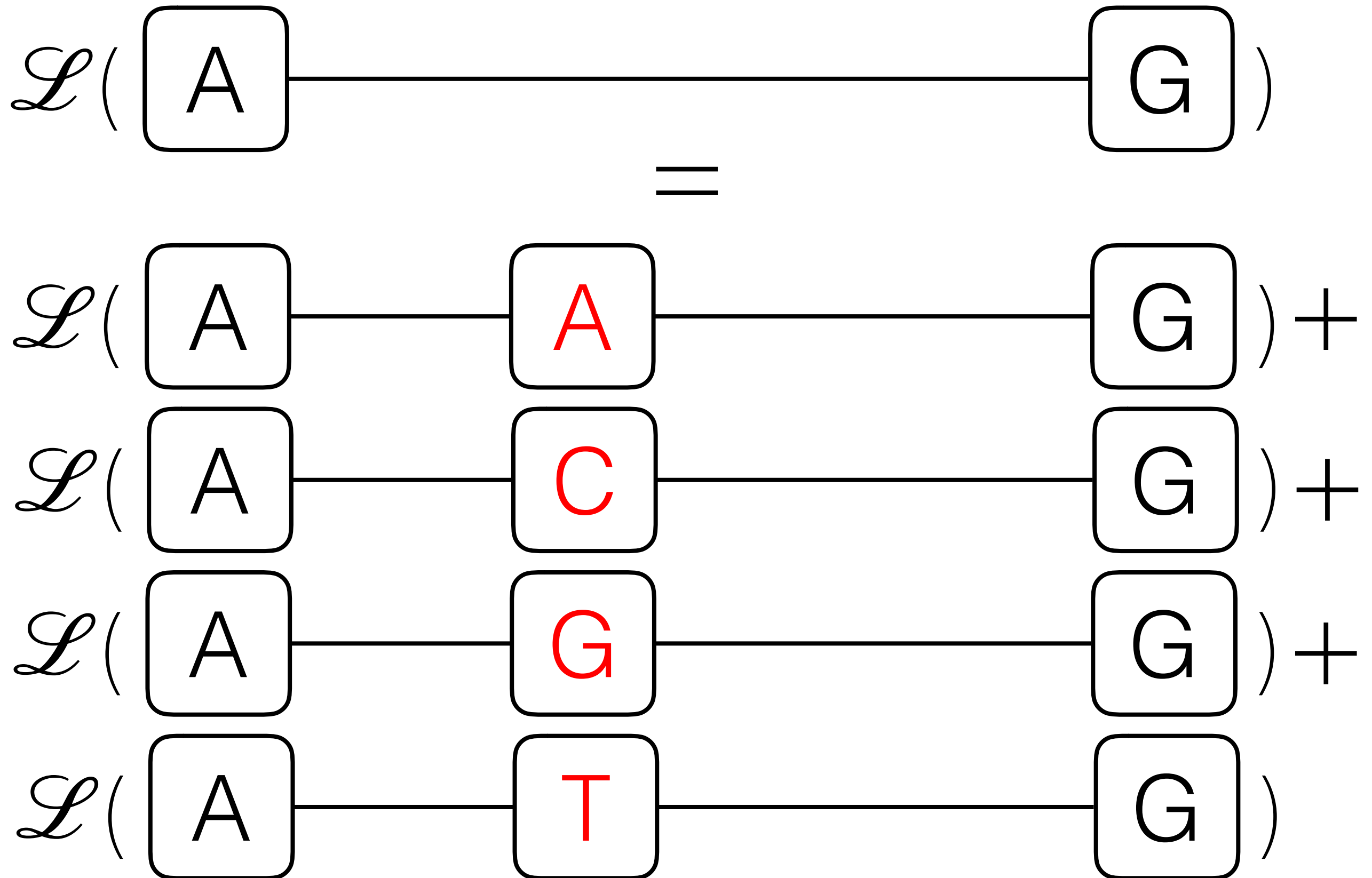


# Calculating Likelihoods on Trees

# Revisiting a Single Branch



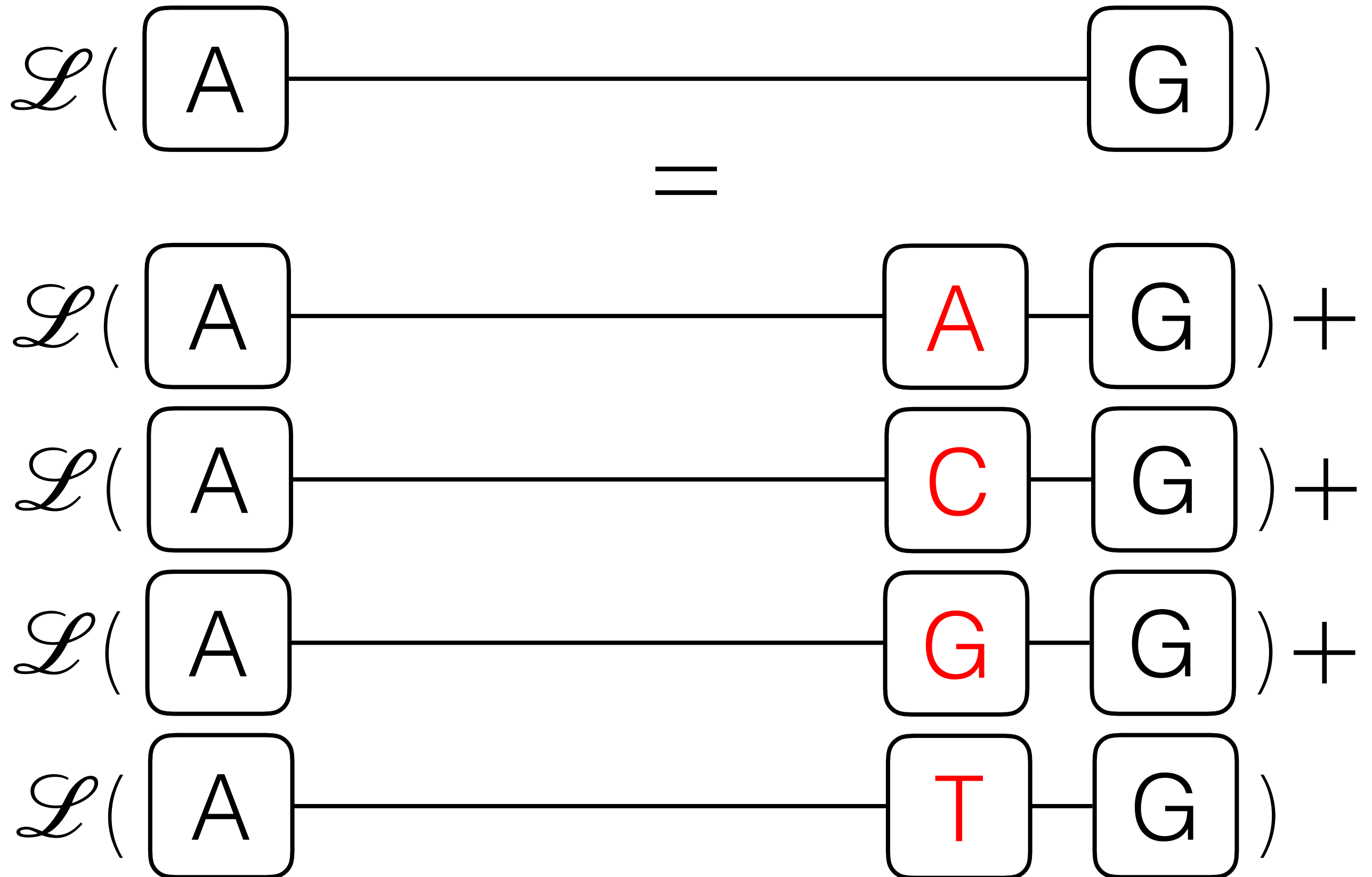
# Revisiting a Single Branch



# Chapman-Kolmogorov Equation

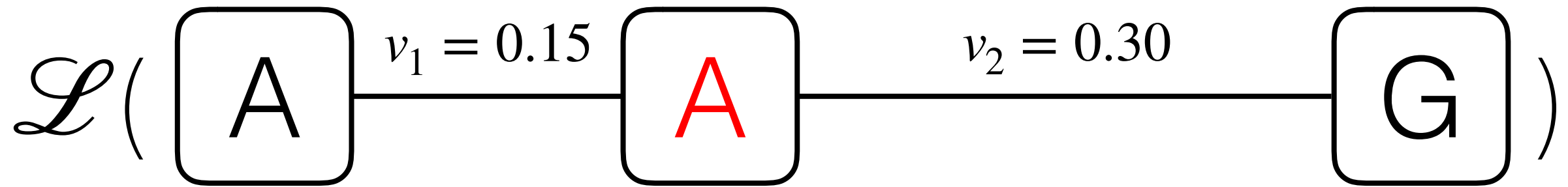
$$\begin{aligned} & \mathcal{L}(\text{A} \text{---} \text{G}) \\ &= \\ & \mathcal{L}(\text{A} \text{---} \text{A} \text{---} \text{G}) + \\ & \mathcal{L}(\text{A} \text{---} \text{C} \text{---} \text{G}) + \\ & \mathcal{L}(\text{A} \text{---} \text{G} \text{---} \text{G}) + \\ & \mathcal{L}(\text{A} \text{---} \text{T} \text{---} \text{G}) \end{aligned}$$

# Pulley Principle (Reversibility)



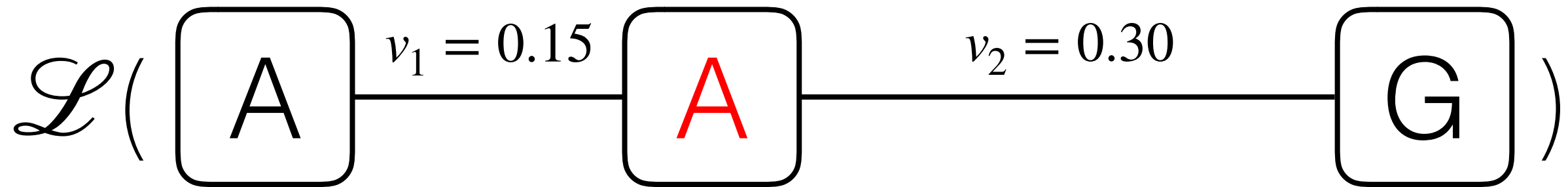
# Pulley Principle (Reversibility)

$$\begin{aligned} \mathcal{L}(\text{A} \text{---} \text{G}) &= \\ \mathcal{L}(\text{A} \text{---} \text{A} \text{---} \text{G}) &+ \\ \mathcal{L}(\text{A} \text{---} \text{C} \text{---} \text{G}) &+ \\ \mathcal{L}(\text{A} \text{---} \text{G} \text{---} \text{G}) &+ \\ \mathcal{L}(\text{A} \text{---} \text{T} \text{---} \text{G}) & \end{aligned}$$



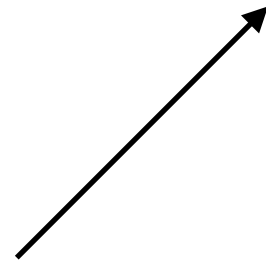
$$P_{AA}(0.15)$$

Probability of starting with an A and ending  
with an A on a branch of length 0.15.



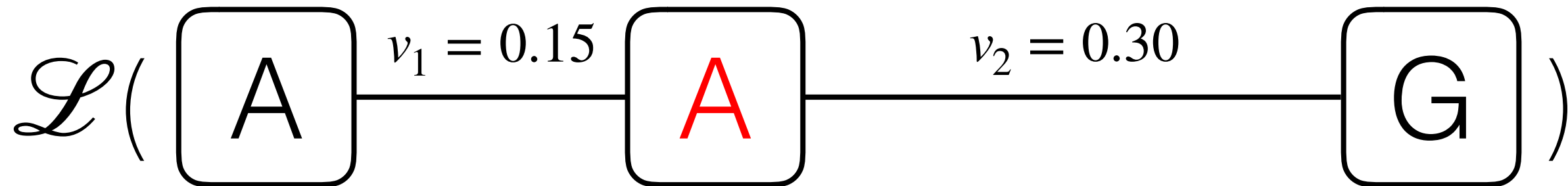
$P_{AA}(0.15)$

$P_{AG}(0.30)$

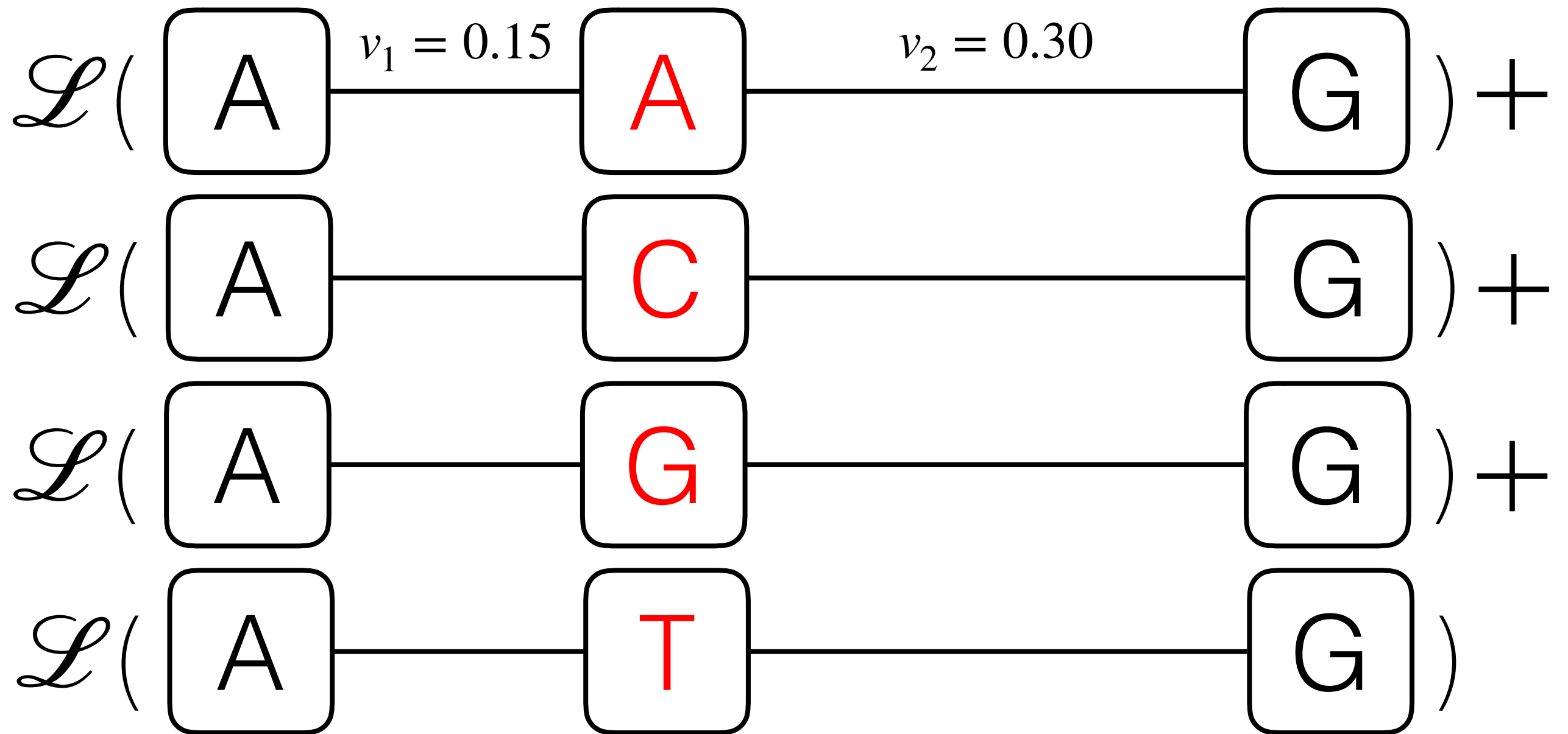


Probability of starting with an A and ending  
with a G on a branch of length 0.30.

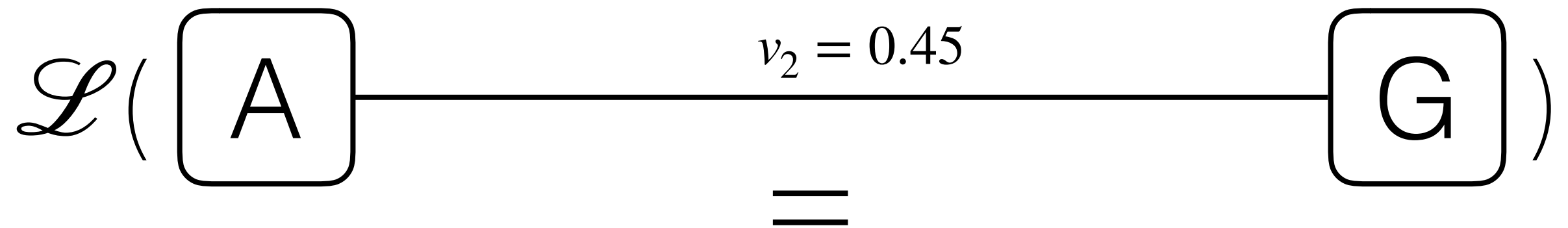




$$P_{AA}(0.15) * P_{AG}(0.30)$$



$$\begin{aligned}
 & \left( P_{AA}(0.15) \right) * P_{AG}(0.30) + \\
 & \left( P_{AC}(0.15) \right) * P_{CG}(0.30) + \\
 & \left( P_{AG}(0.15) \right) * P_{GG}(0.30) + \\
 & \left( P_{AT}(0.15) \right) * P_{TG}(0.30)
 \end{aligned}$$

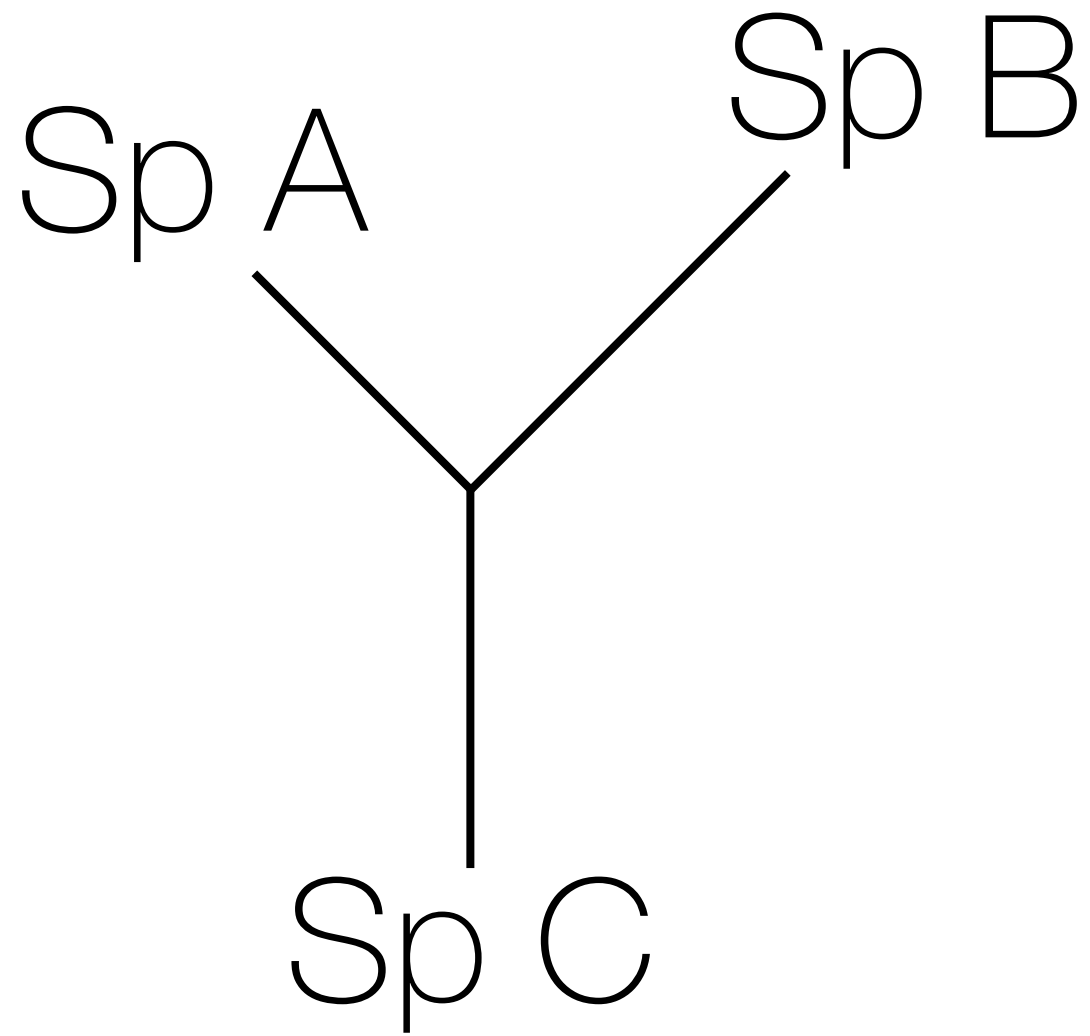


$$P_{AG}(0.45)$$

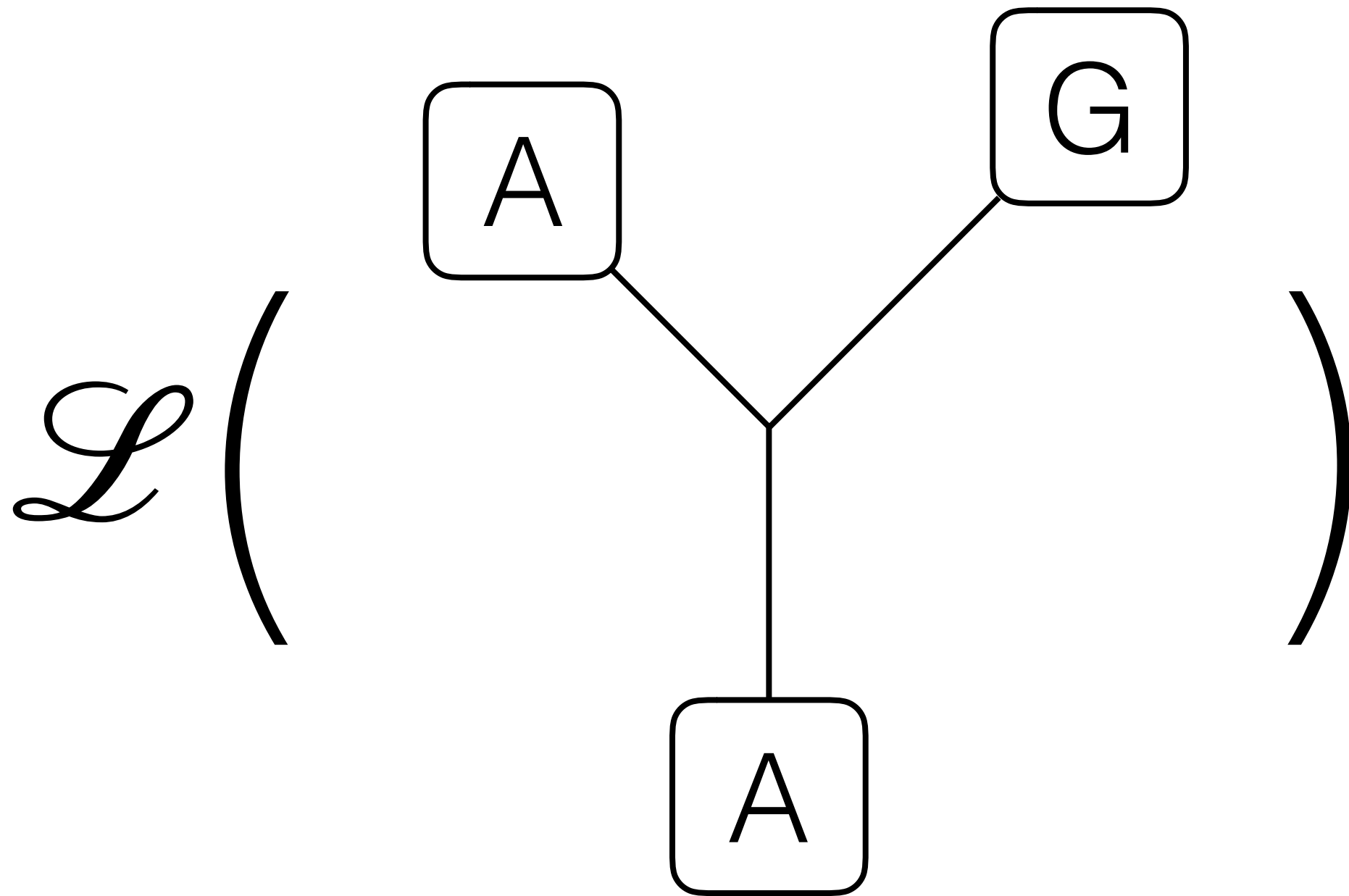
$$=$$

$$\begin{aligned} & \left( P_{AA}(0.15) \right) * P_{AG}(0.30) + \\ & \left( P_{AC}(0.15) \right) * P_{CG}(0.30) + \\ & \left( P_{AG}(0.15) \right) * P_{GG}(0.30) + \\ & \left( P_{AT}(0.15) \right) * P_{TG}(0.30) \end{aligned}$$

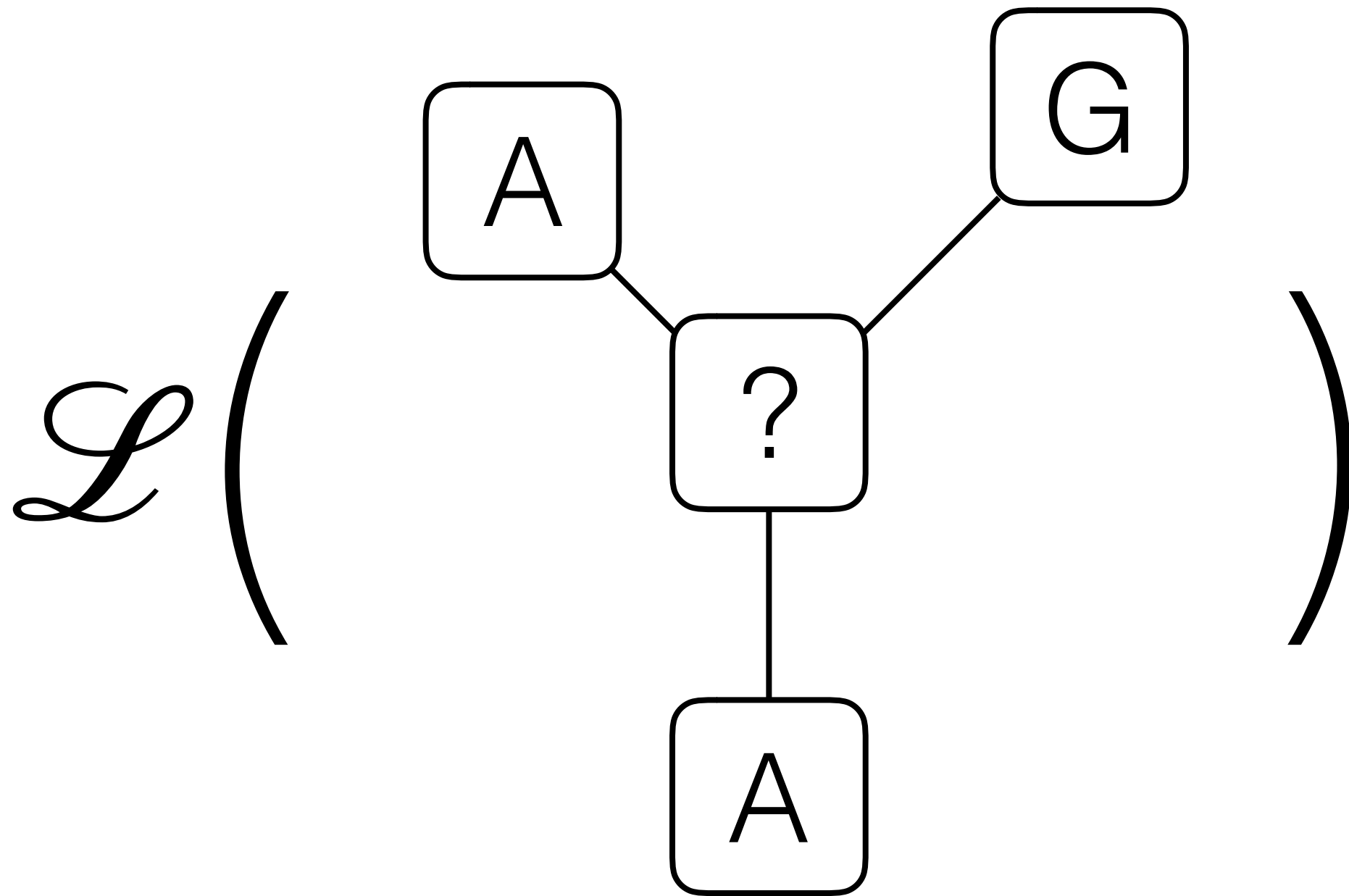
# Three-Taxon Unrooted Tree



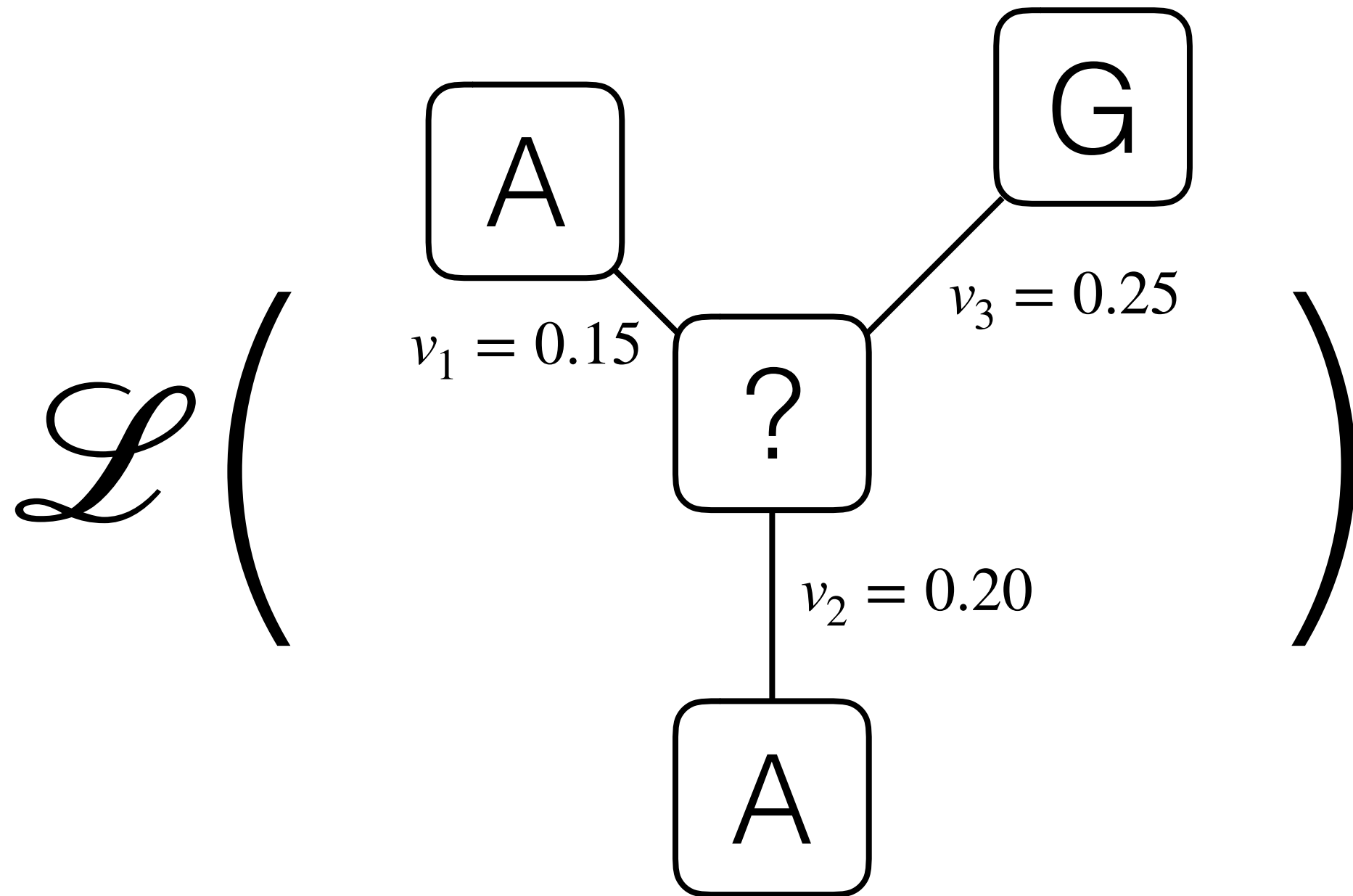
# Three-Taxon Unrooted Tree

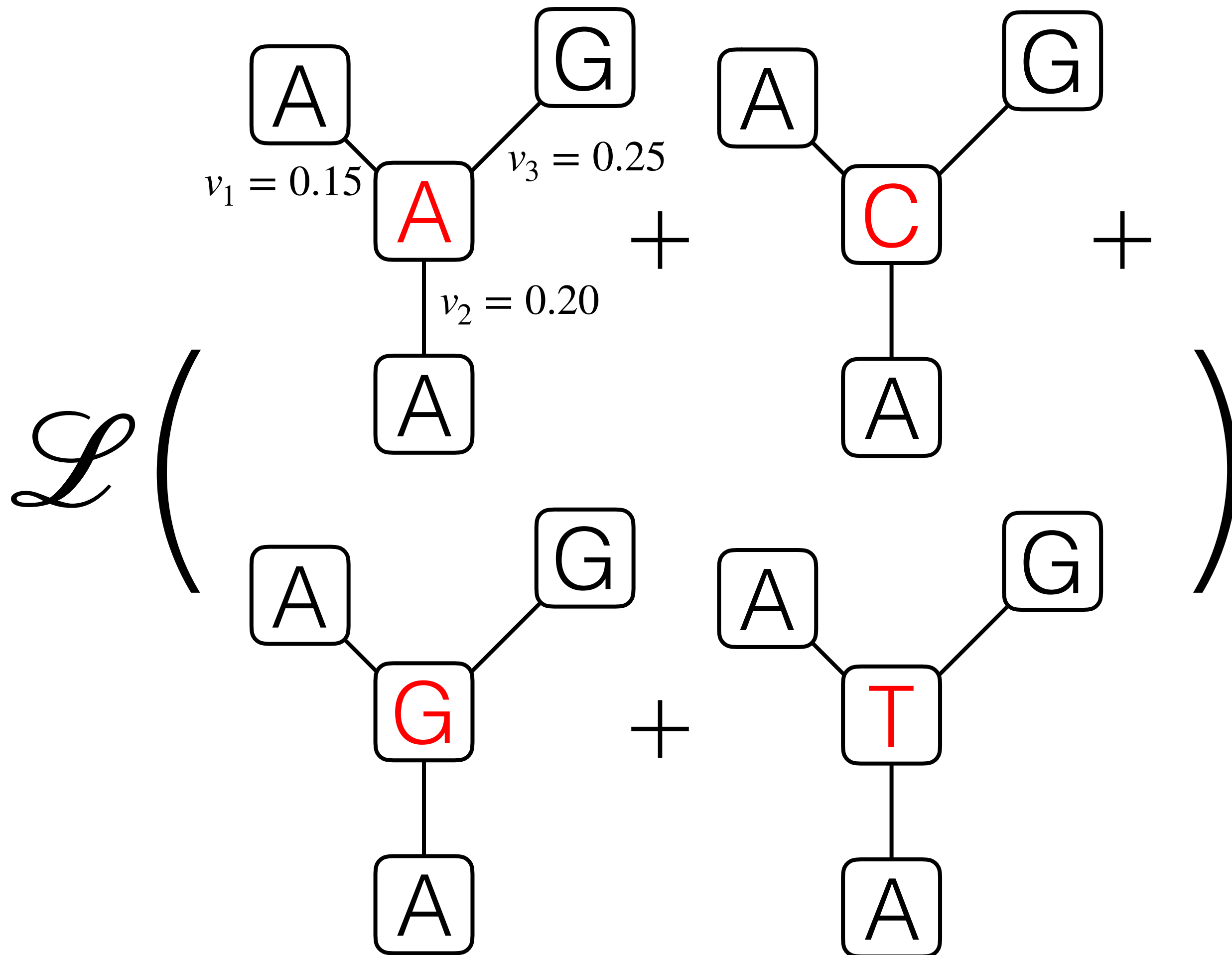


# Three-Taxon Unrooted Tree



# Three-Taxon Unrooted Tree





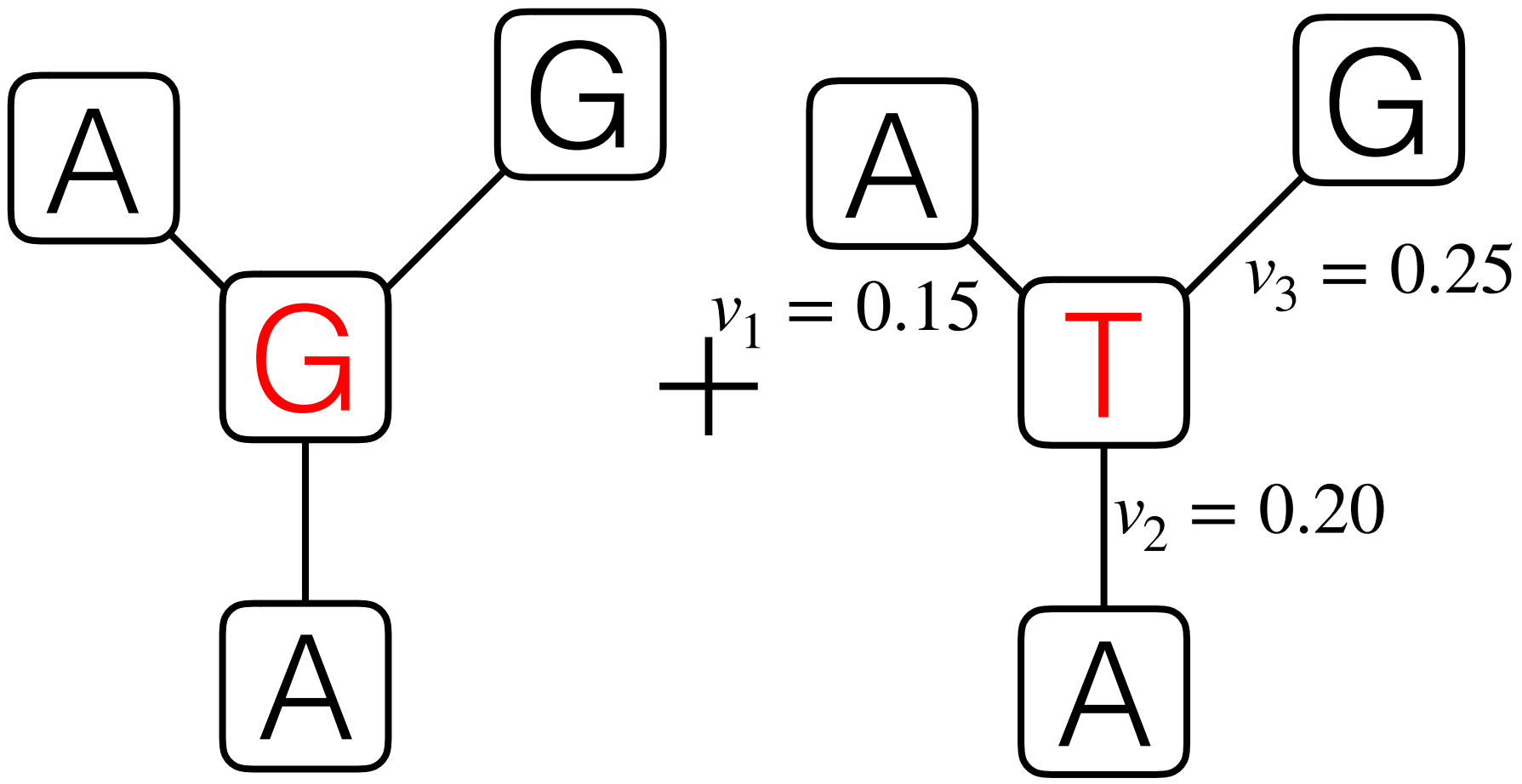


$$\mathcal{L} \left( P_{AA}(0.15)P_{AG}(0.25)P_{AA}(0.20) + \begin{array}{c} \boxed{A} \quad \boxed{G} \\ \diagdown \quad \diagup \\ \boxed{C} \\ | \\ \boxed{A} \end{array} + \begin{array}{c} \boxed{A} \quad \boxed{G} \\ \diagdown \quad \diagup \\ \boxed{G} \\ | \\ \boxed{A} \end{array} + \begin{array}{c} \boxed{A} \quad \boxed{G} \\ \diagdown \quad \diagup \\ \boxed{T} \\ | \\ \boxed{A} \end{array} \right)$$

$v_1 = 0.15$        $v_2 = 0.20$        $v_3 = 0.25$

$$P_{AA}(0.15)P_{AG}(0.25)P_{AA}(0.20) + P_{CA}(0.15)P_{CG}(0.25)P_{CA}(0.20) +$$

$\mathcal{L} ($



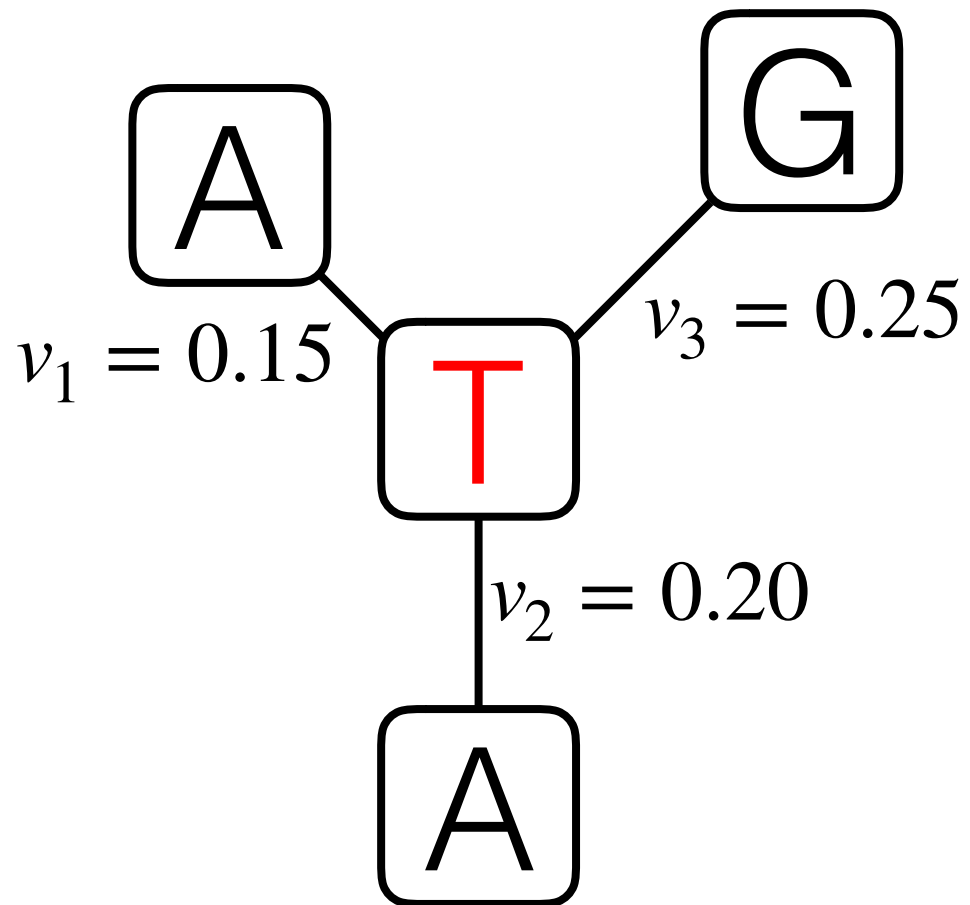
$+$

$)$

$$P_{AA}(0.15)P_{AG}(0.25)P_{AA}(0.20) + P_{CA}(0.15)P_{CG}(0.25)P_{CA}(0.20) +$$

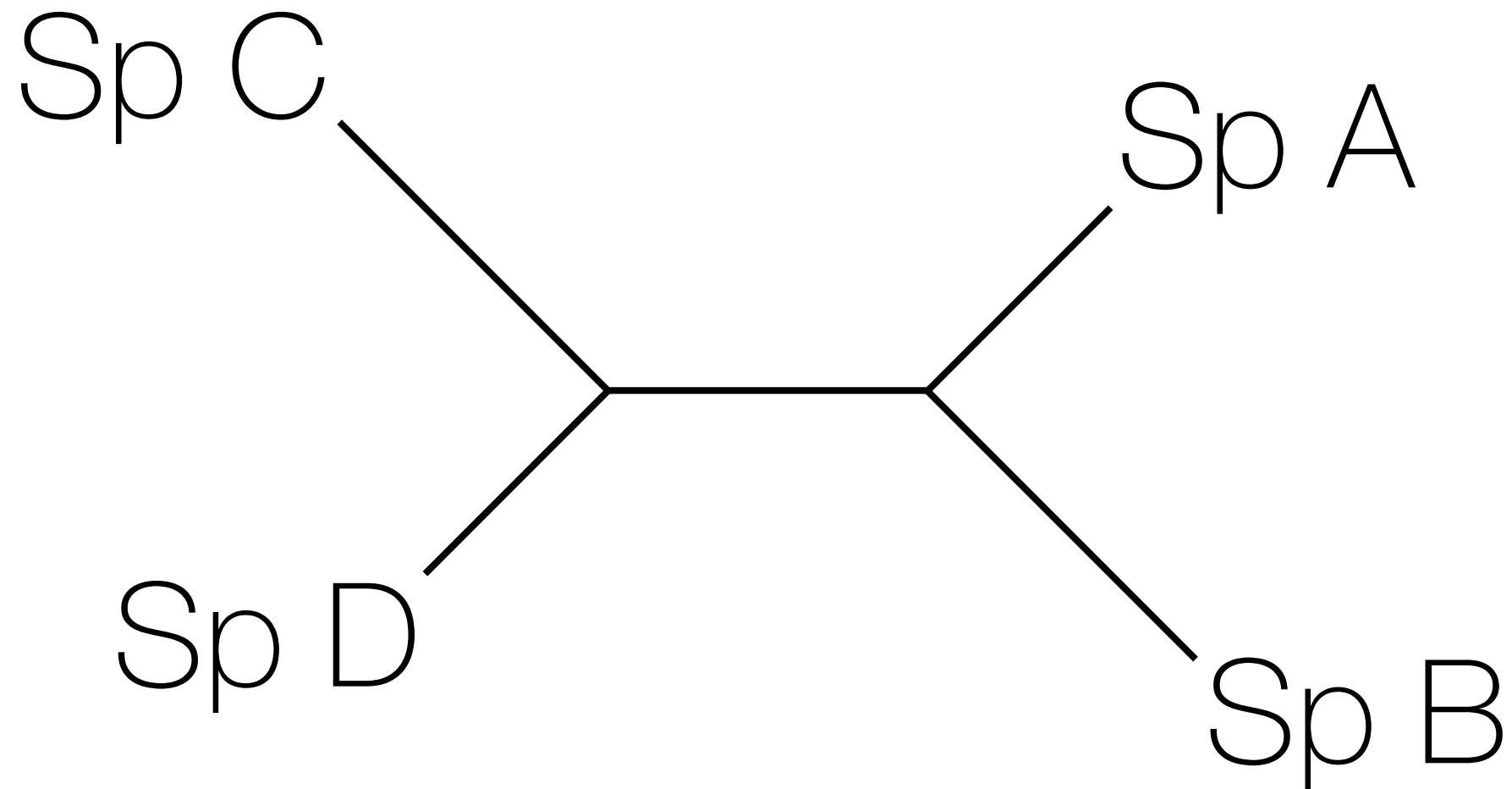
$$\mathcal{L} \left( \right.$$

$$P_{GA}(0.15)P_{GG}(0.25)P_{GA}(0.20) +$$

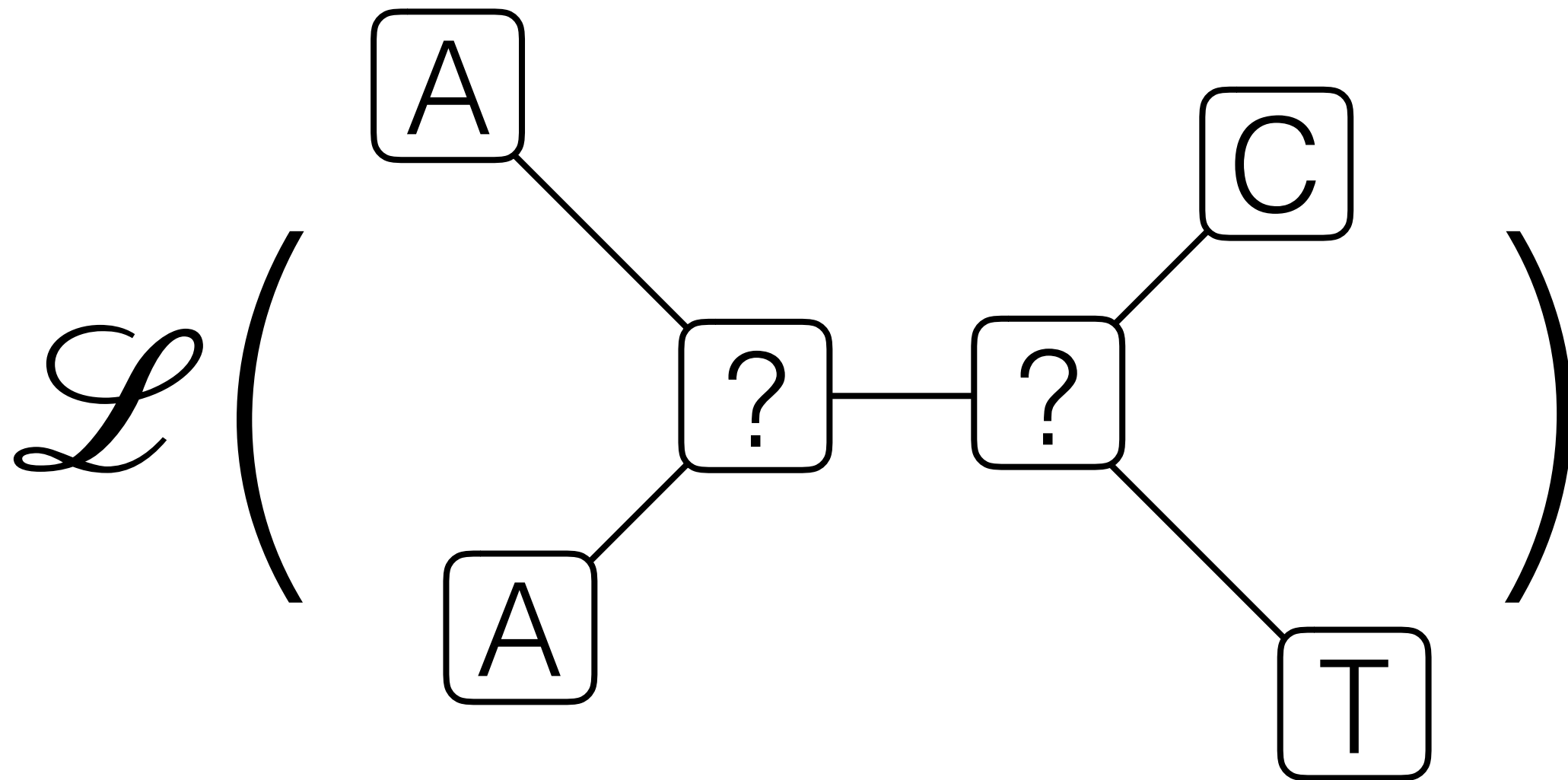


$$\begin{aligned}
 &P_{AA}(0.15)P_{AG}(0.25)P_{AA}(0.20) + P_{CA}(0.15)P_{CG}(0.25)P_{CA}(0.20) + \\
 &\mathcal{L}\left( \right. \\
 &\left. P_{GA}(0.15)P_{GG}(0.25)P_{GA}(0.20) + P_{TA}(0.15)P_{TG}(0.25)P_{TA}(0.20) \right)
 \end{aligned}$$

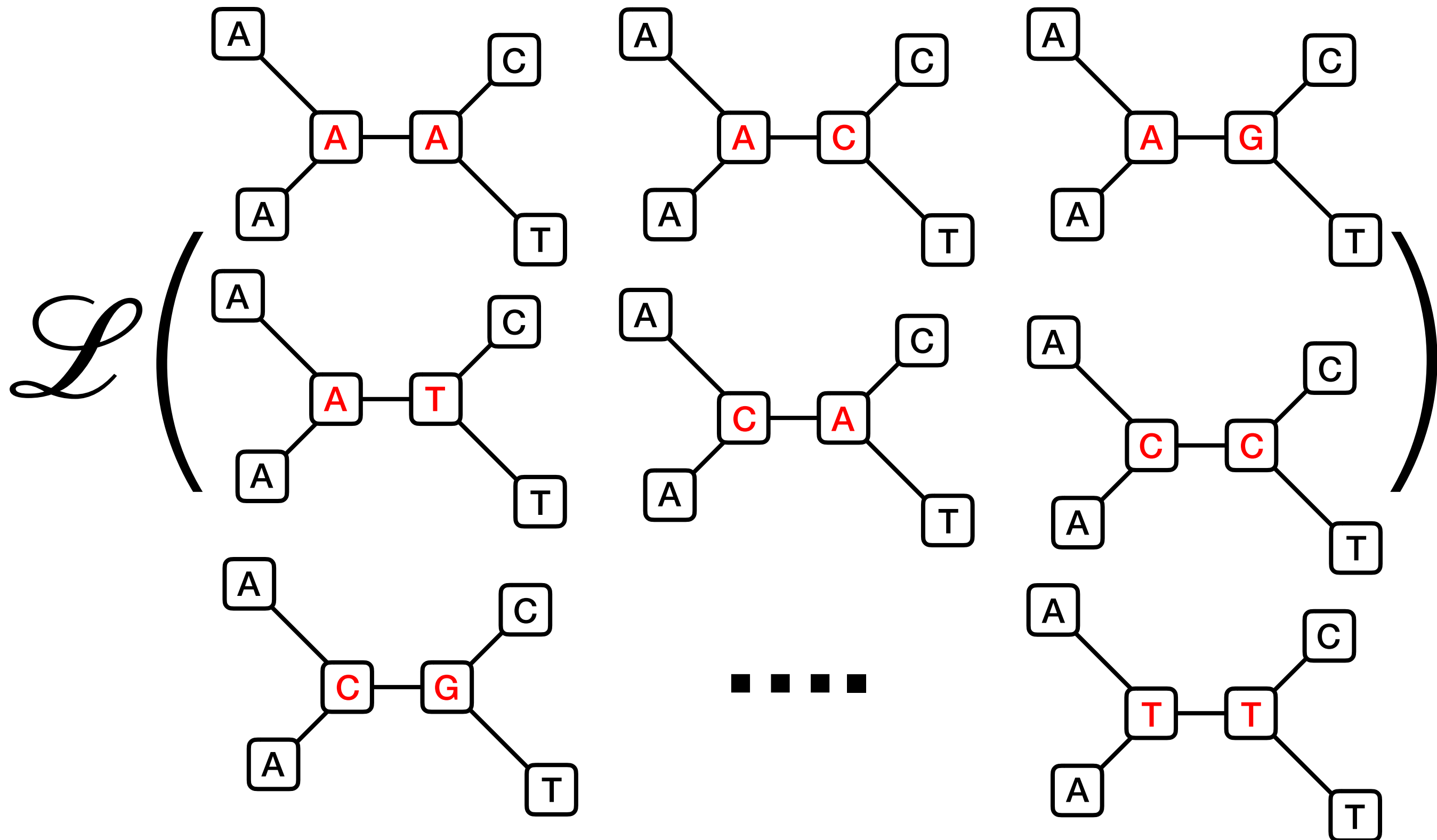
# Four-Taxon Unrooted Tree



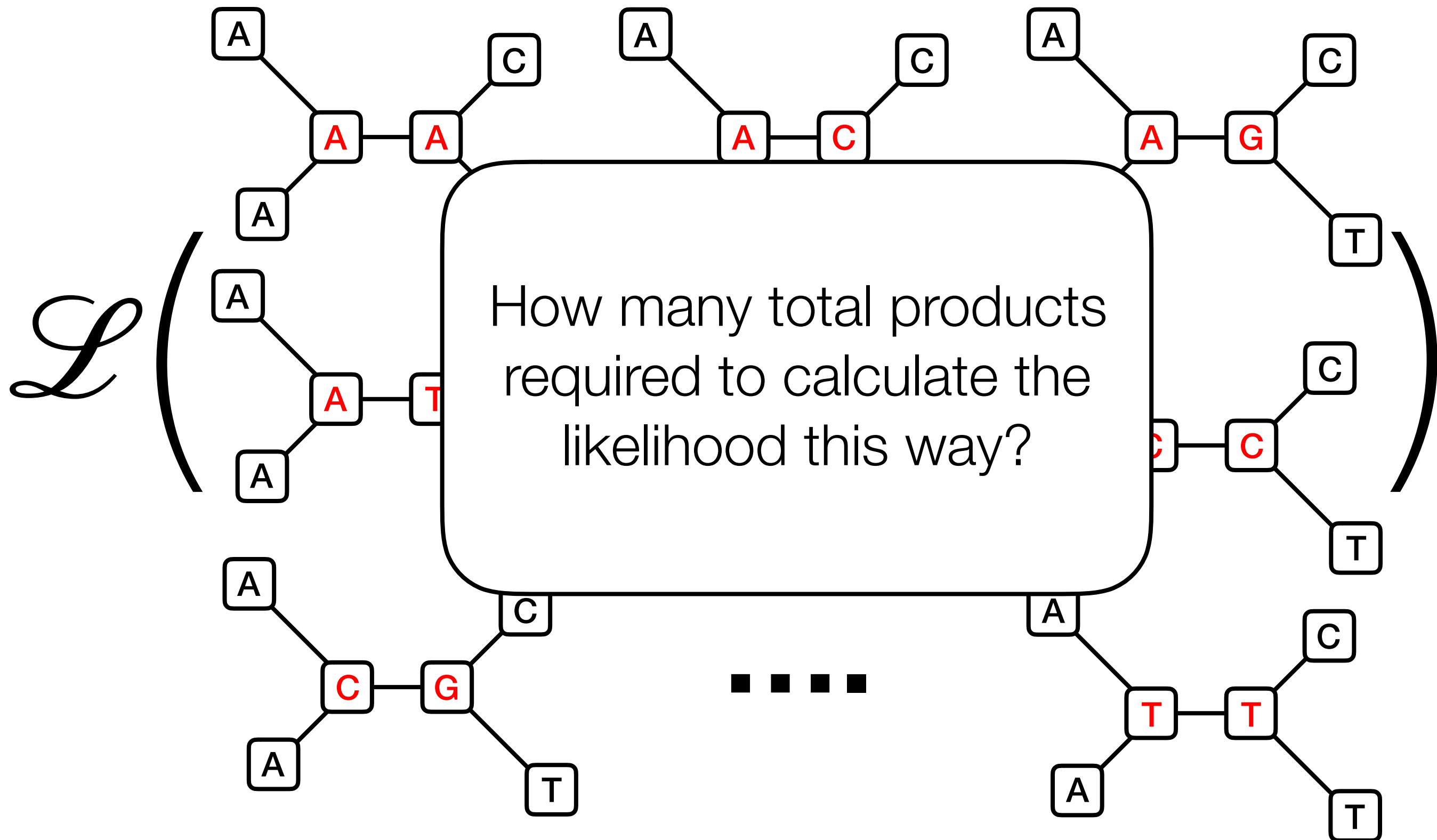
# Four-Taxon Unrooted Tree



# Four-Taxon Unrooted Tree

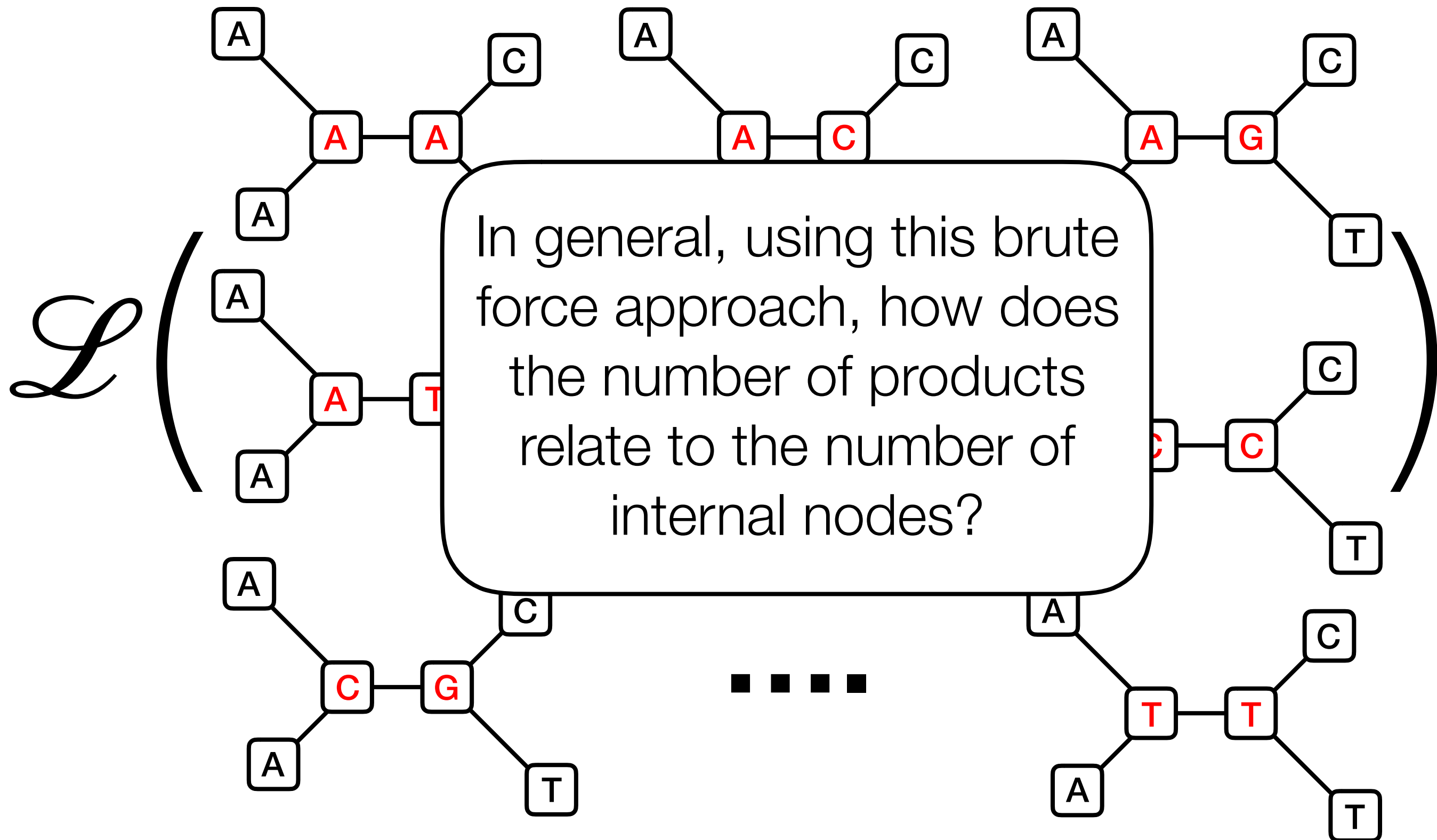


# Four-Taxon Unrooted Tree

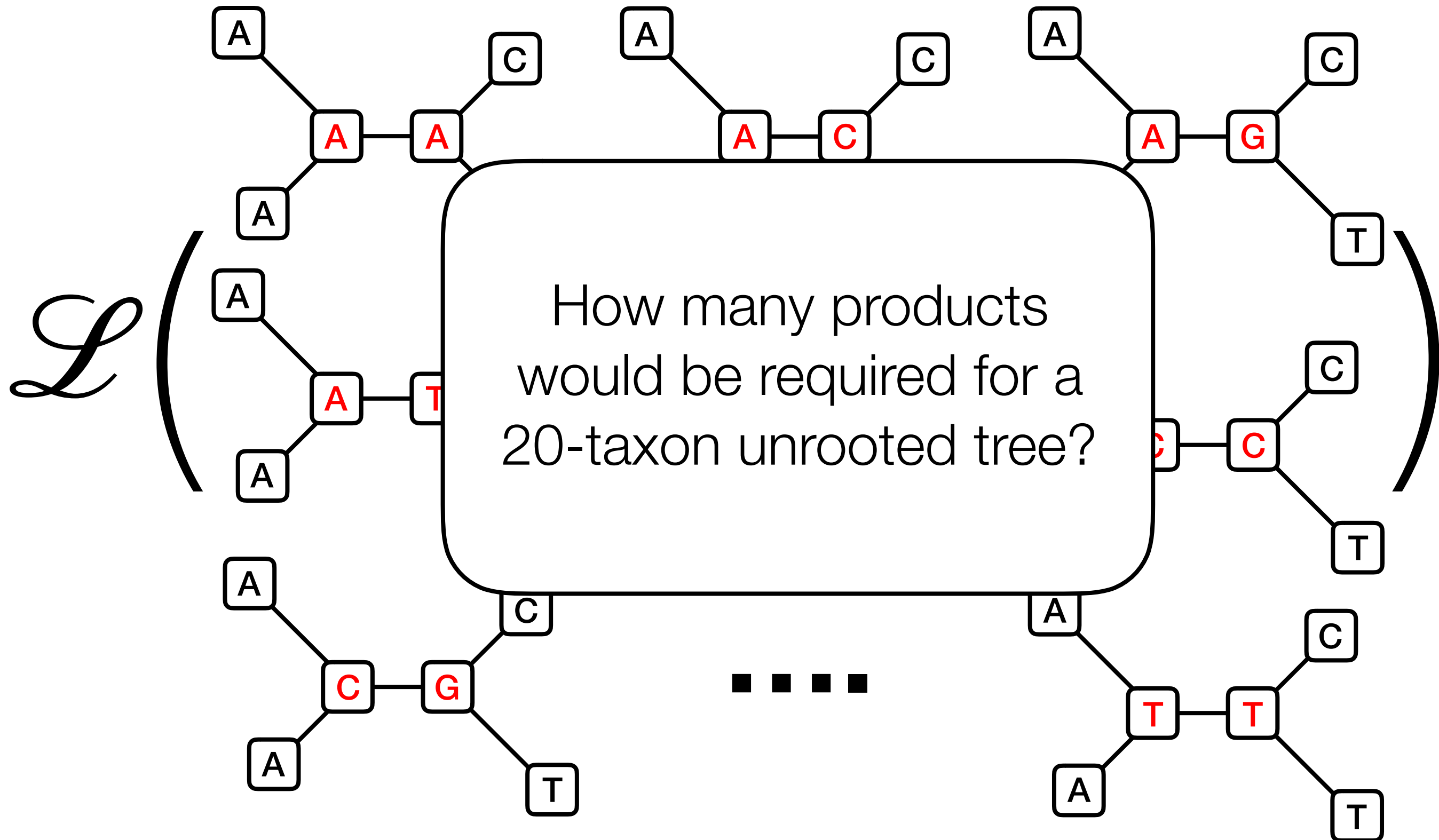




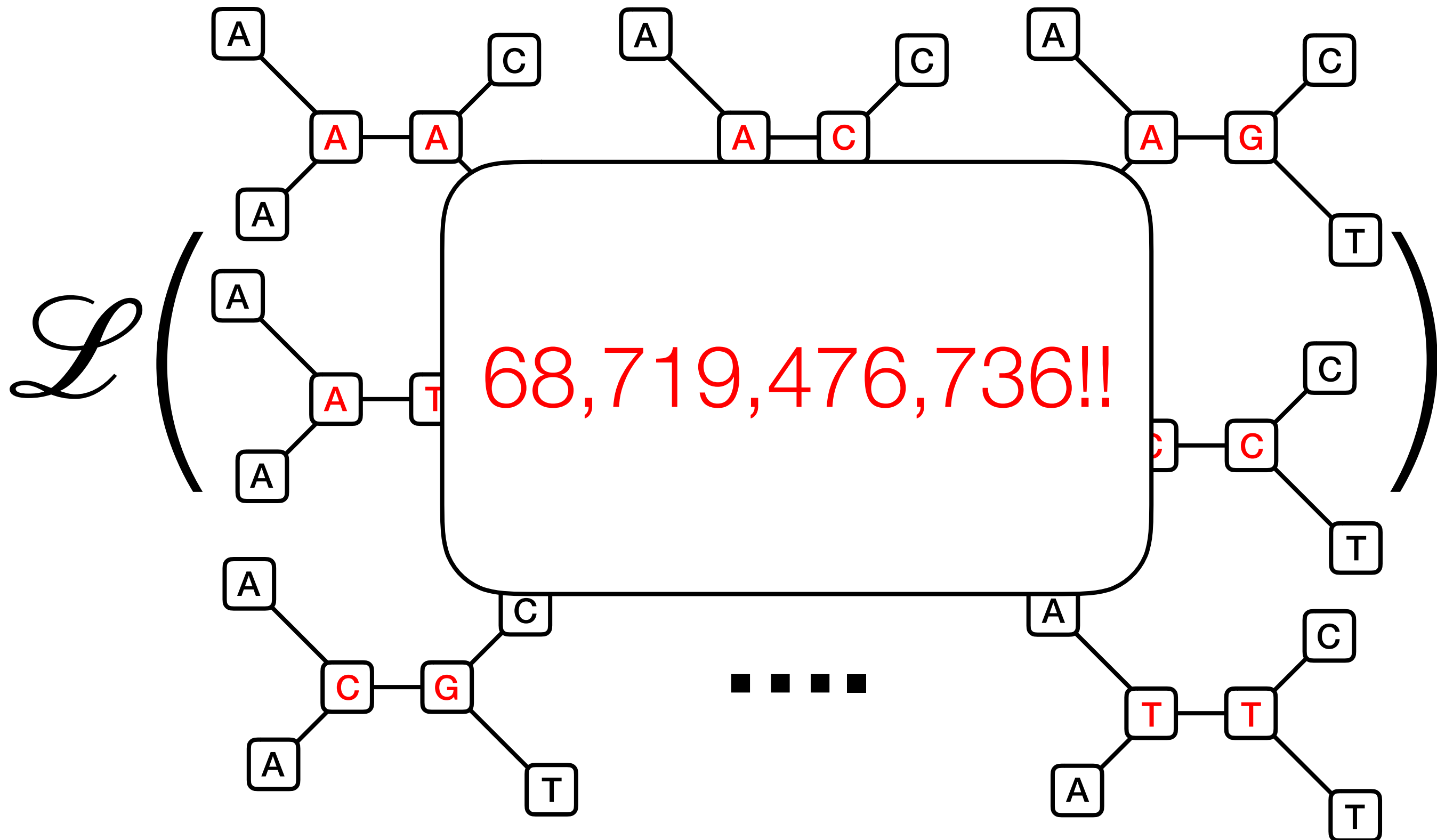
# Four-Taxon Unrooted Tree



# Four-Taxon Unrooted Tree



# Four-Taxon Unrooted Tree



# Felsenstein's Pruning Algorithm

J Mol Evol (1981) 17:368–376

---

Journal of  
Molecular Evolution  
© Springer-Verlag 1981

---

## Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach

Joseph Felsenstein

Department of Genetics, University of Washington, Seattle, Washington 98195, USA

**Summary.** The application of maximum likelihood techniques to the estimation of evolutionary trees from nucleic acid sequence data is discussed. A computationally feasible method for finding such maximum likelihood estimates is developed, and a computer program is available. This method has advantages over the traditional parsimony algorithms, which can give misleading results if rates of evolution differ in different lineages. It also allows the testing of hypotheses about the constancy of evolutionary rates by likelihood ratio tests, and gives rough indication of the error of the estimate of the tree.

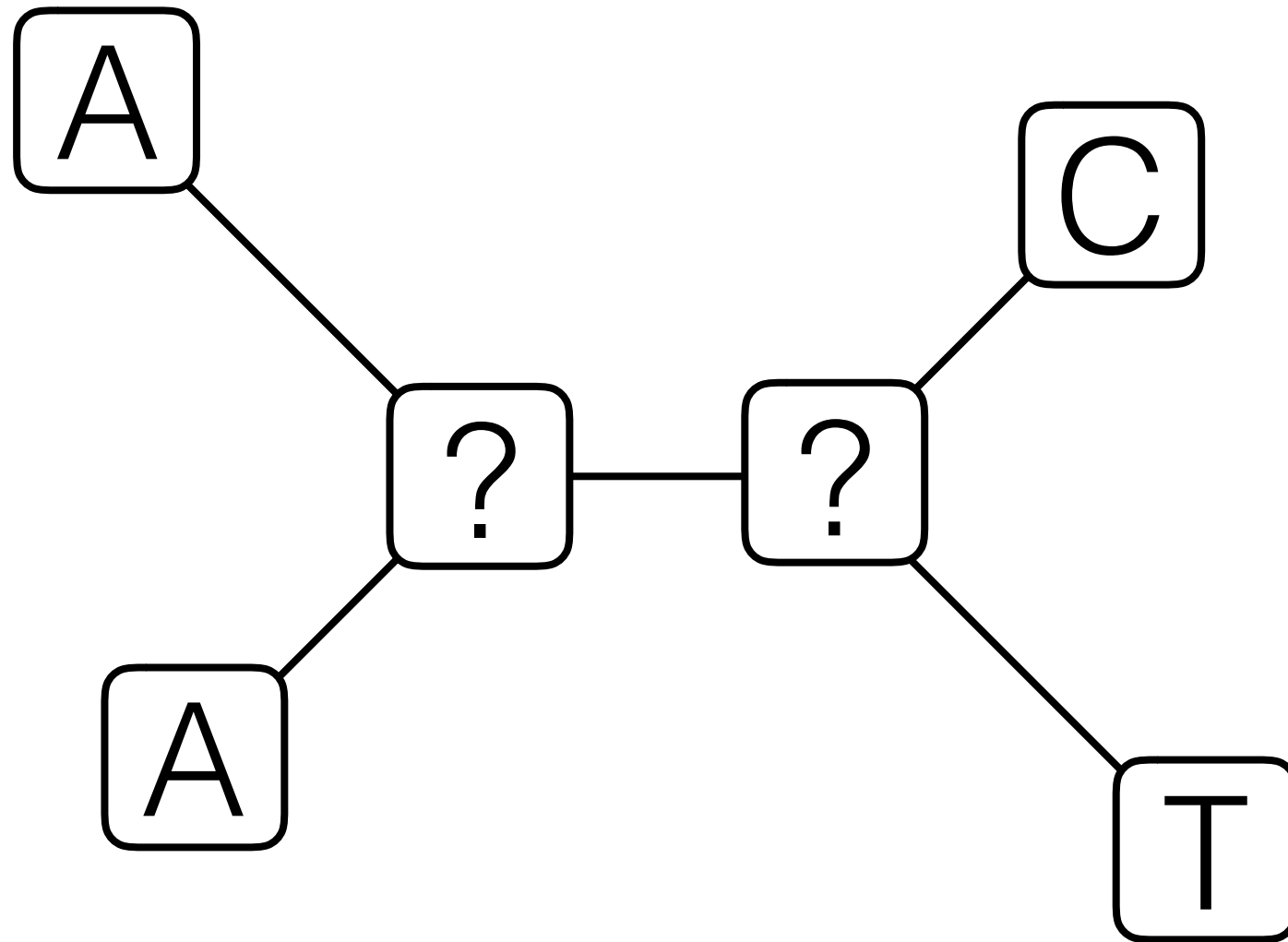
**Key words:** Evolution — Phylogeny — Maximum likelihood — Parsimony — Estimation — DNA sequences

---

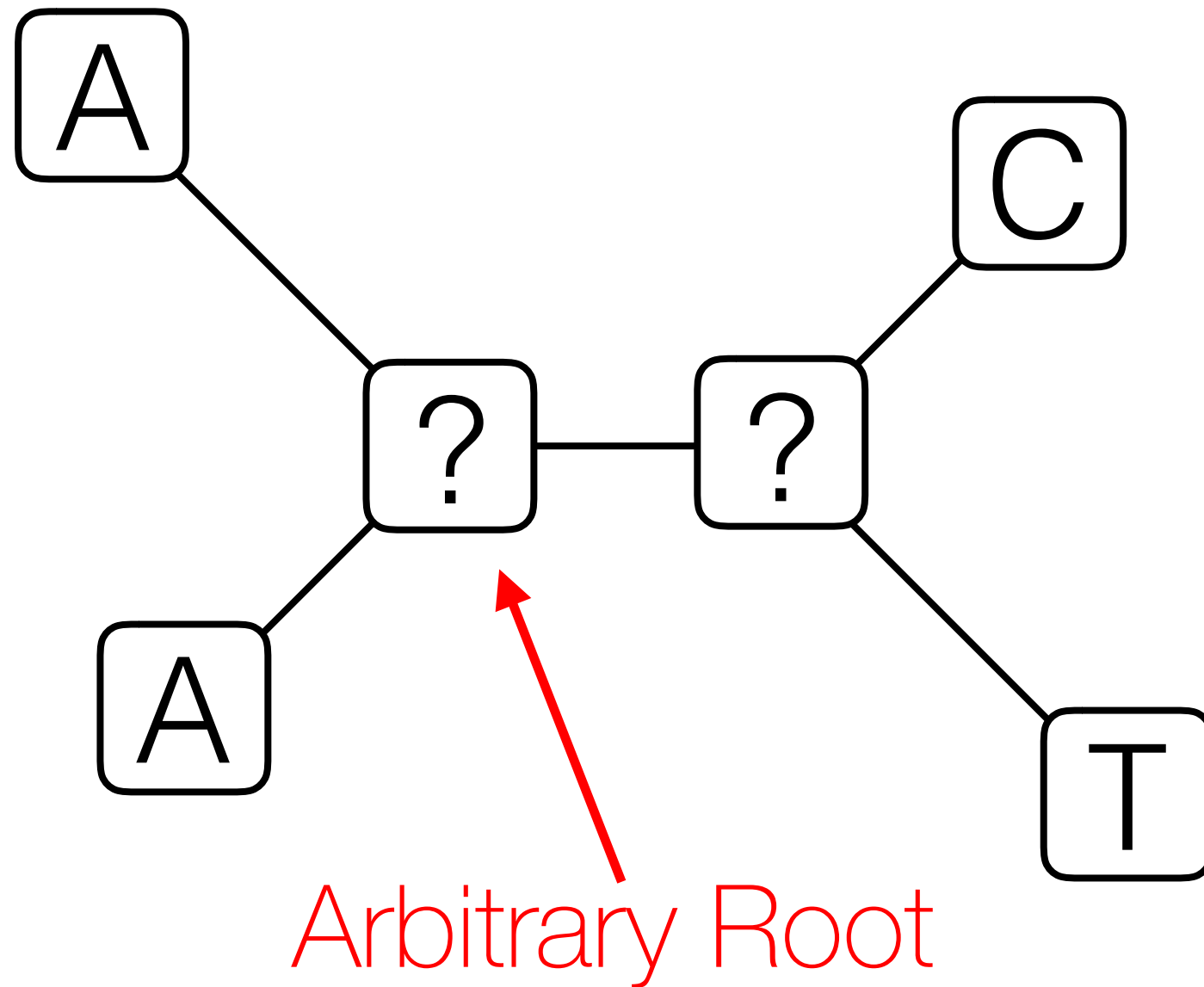
produced by parsimony methods (Edwards 1963; Edwards and Cavalli-Sforza 1964; Camin and Sokal 1965). These methods implicitly assume that change is improbable a priori (Felsenstein 1973, 1979). If the amount of change is small over the evolutionary times being considered, parsimony methods will be well-justified statistical methods.

Most data involve moderate to large amounts of change, and it is in such cases that parsimony methods can fail. When amounts of evolutionary change in different lineages are sufficiently unequal, it can be shown (Felsenstein 1978b) that parsimony methods make an inconsistent estimate of the evolutionary tree, converging to the wrong tree with increasing certainty as more sequences are considered for the same set of species. The compatibility approach to estimating evolutionary trees (Le Quesne 1969; Sneath et al. 1975;

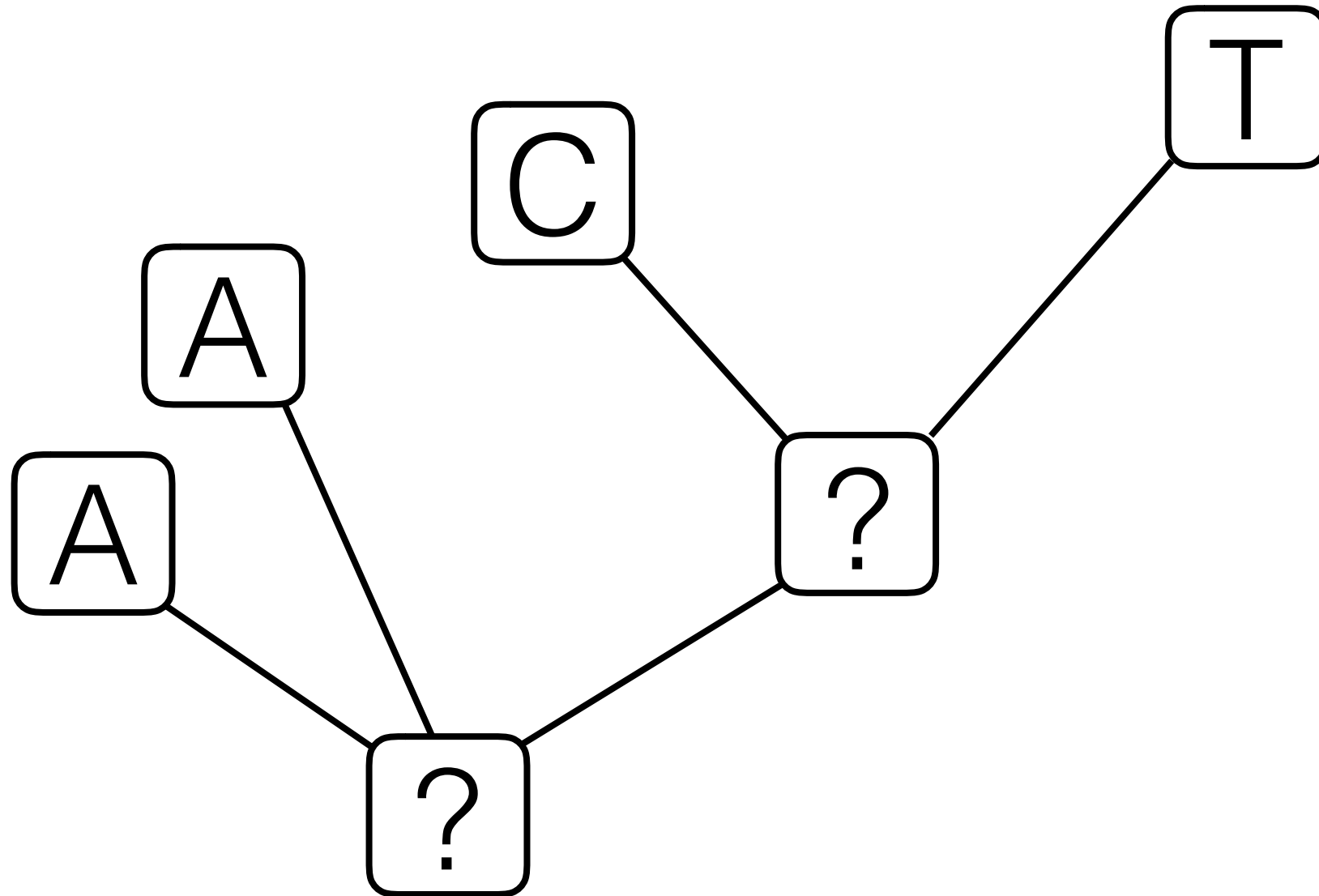
# Felsenstein's Pruning Algorithm



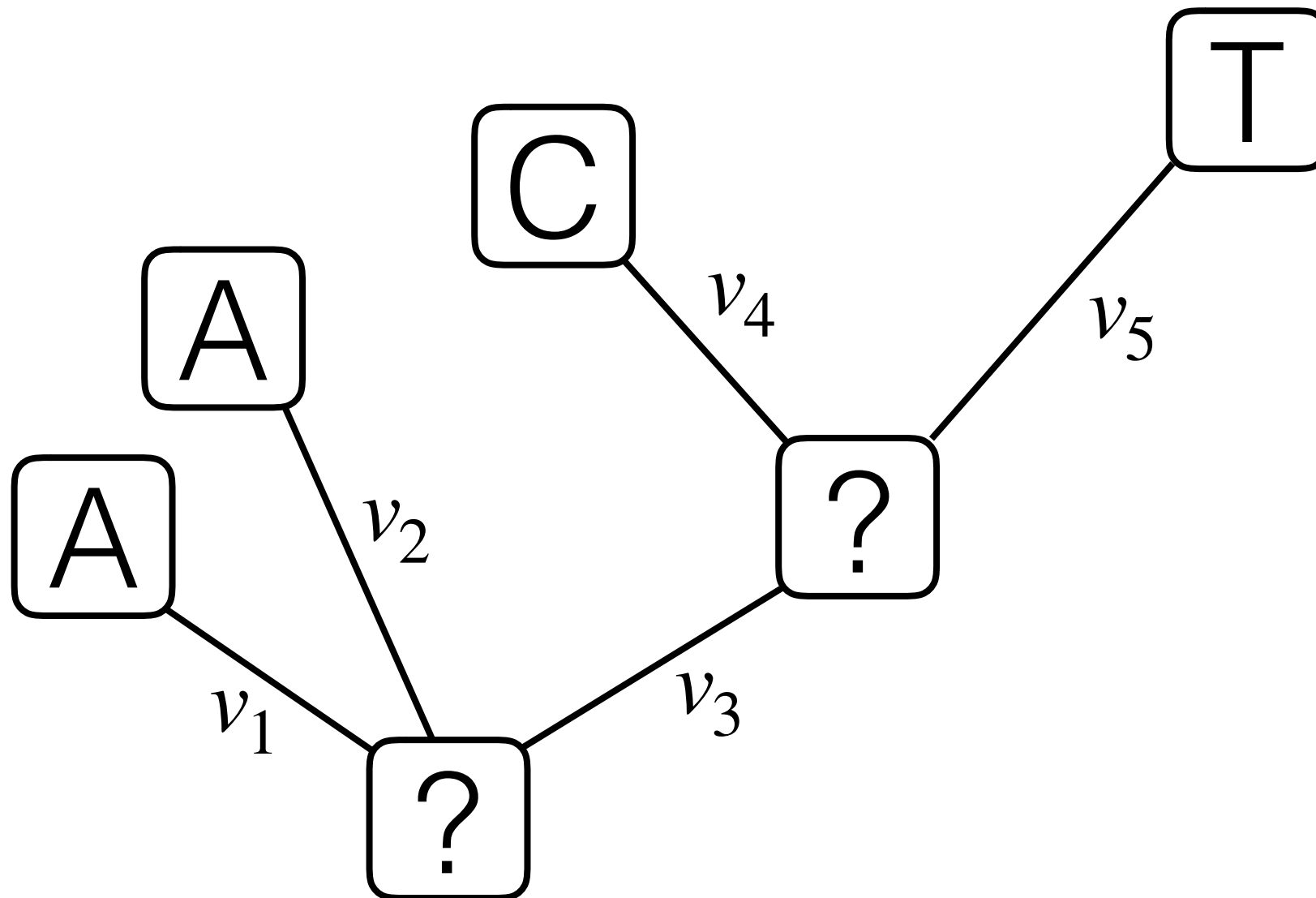
# Felsenstein's Pruning Algorithm



# Felsenstein's Pruning Algorithm

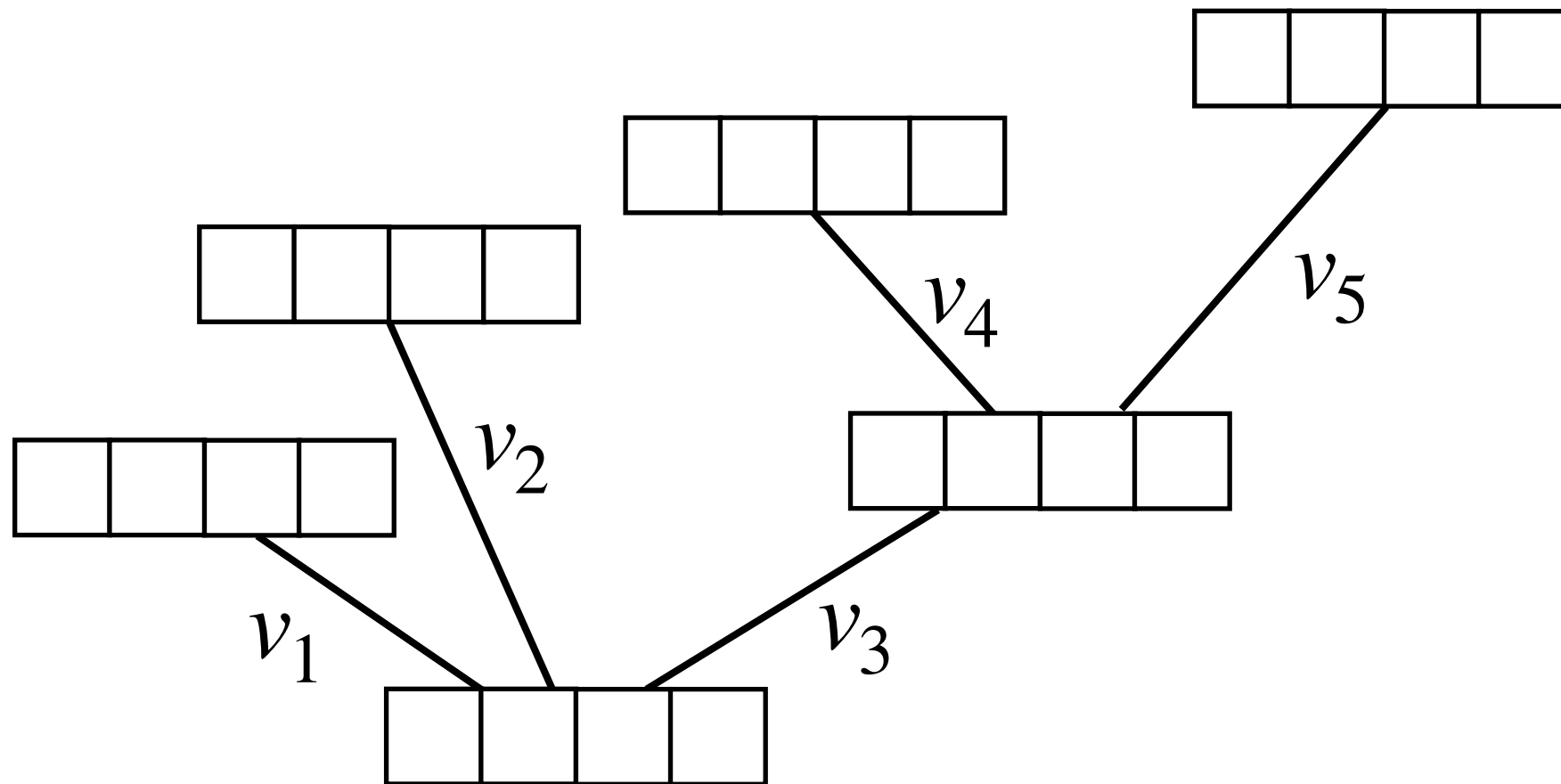


# Felsenstein's Pruning Algorithm

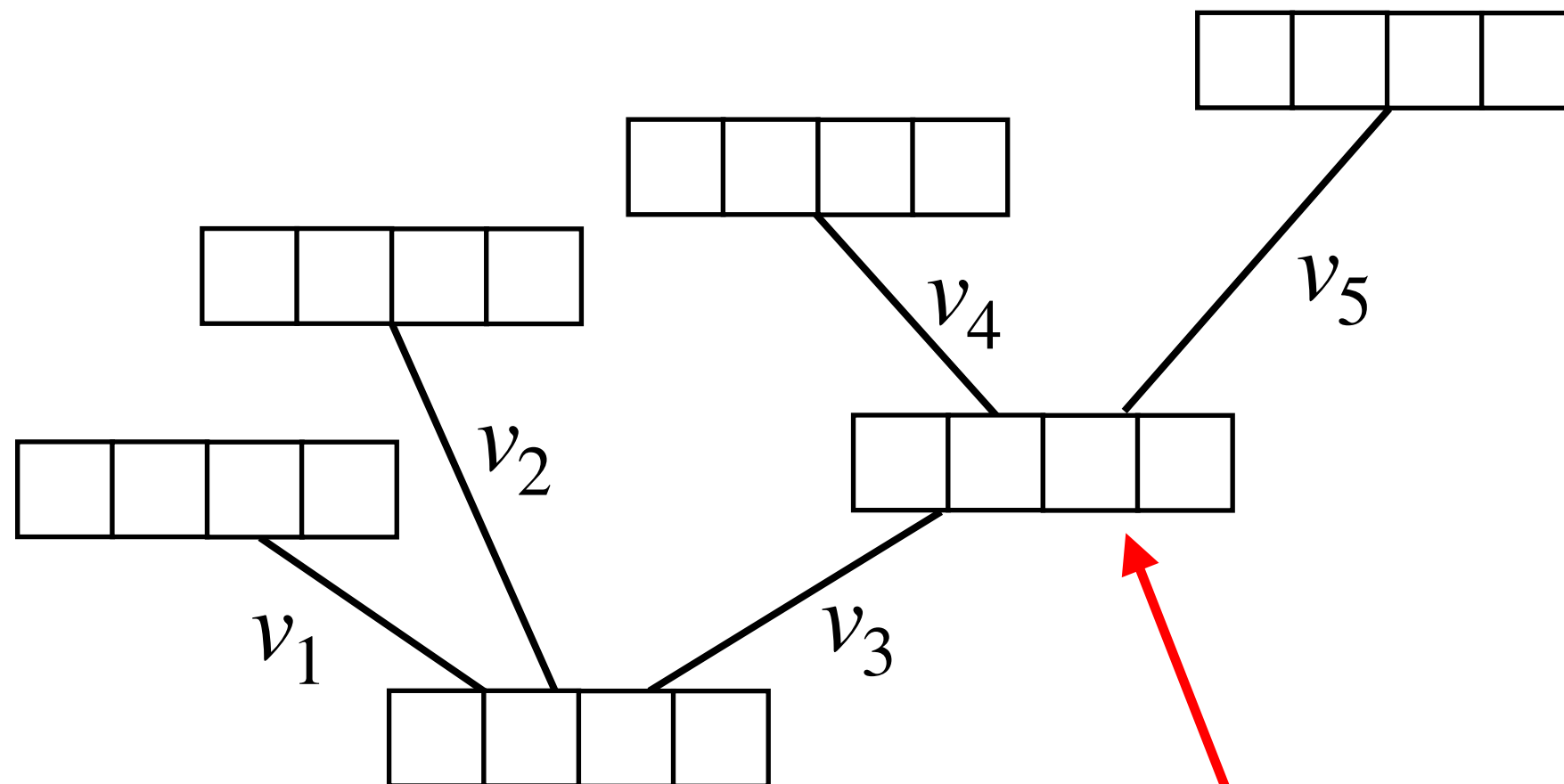




# Felsenstein's Pruning Algorithm

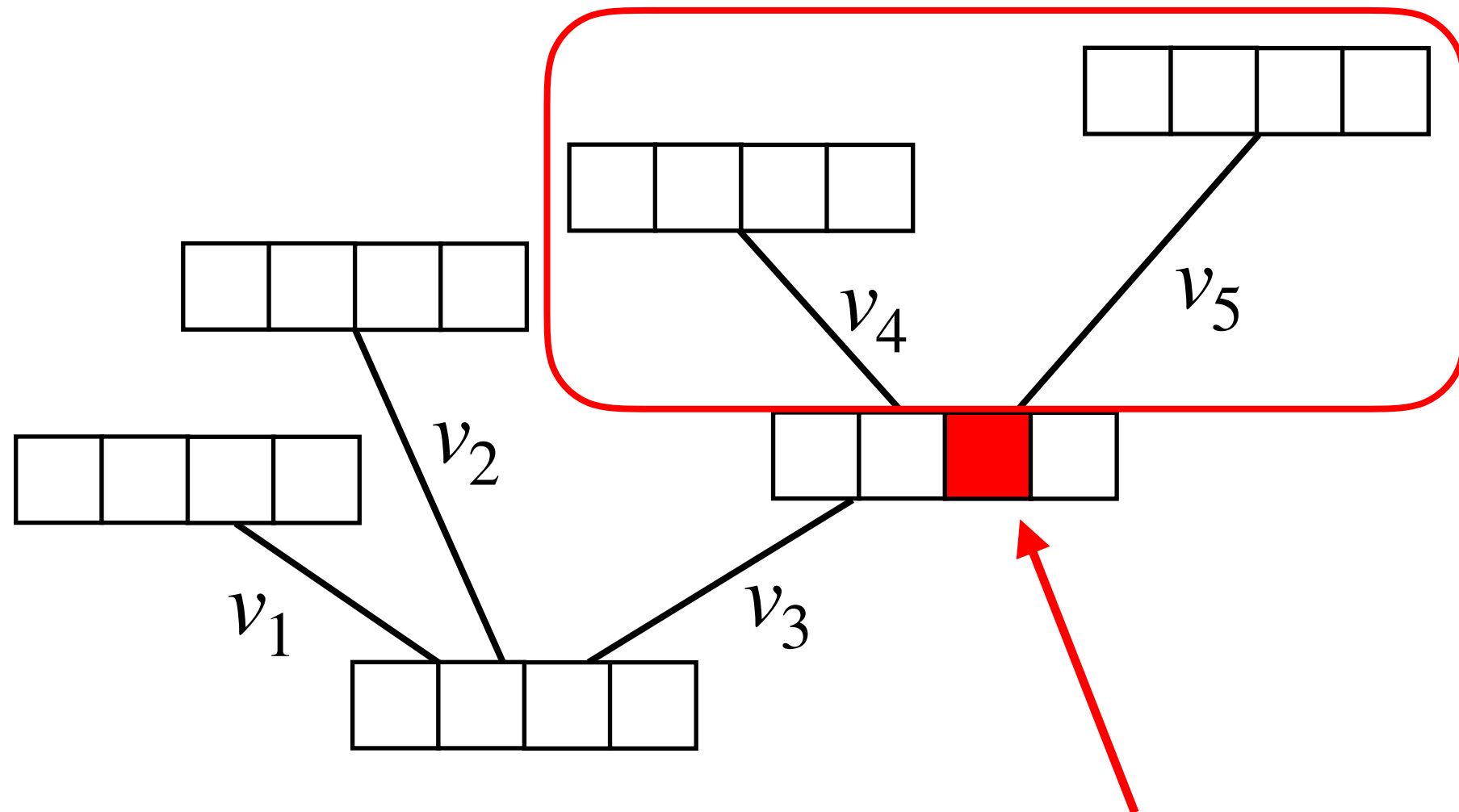


# Felsenstein's Pruning Algorithm



Conditional Likelihoods

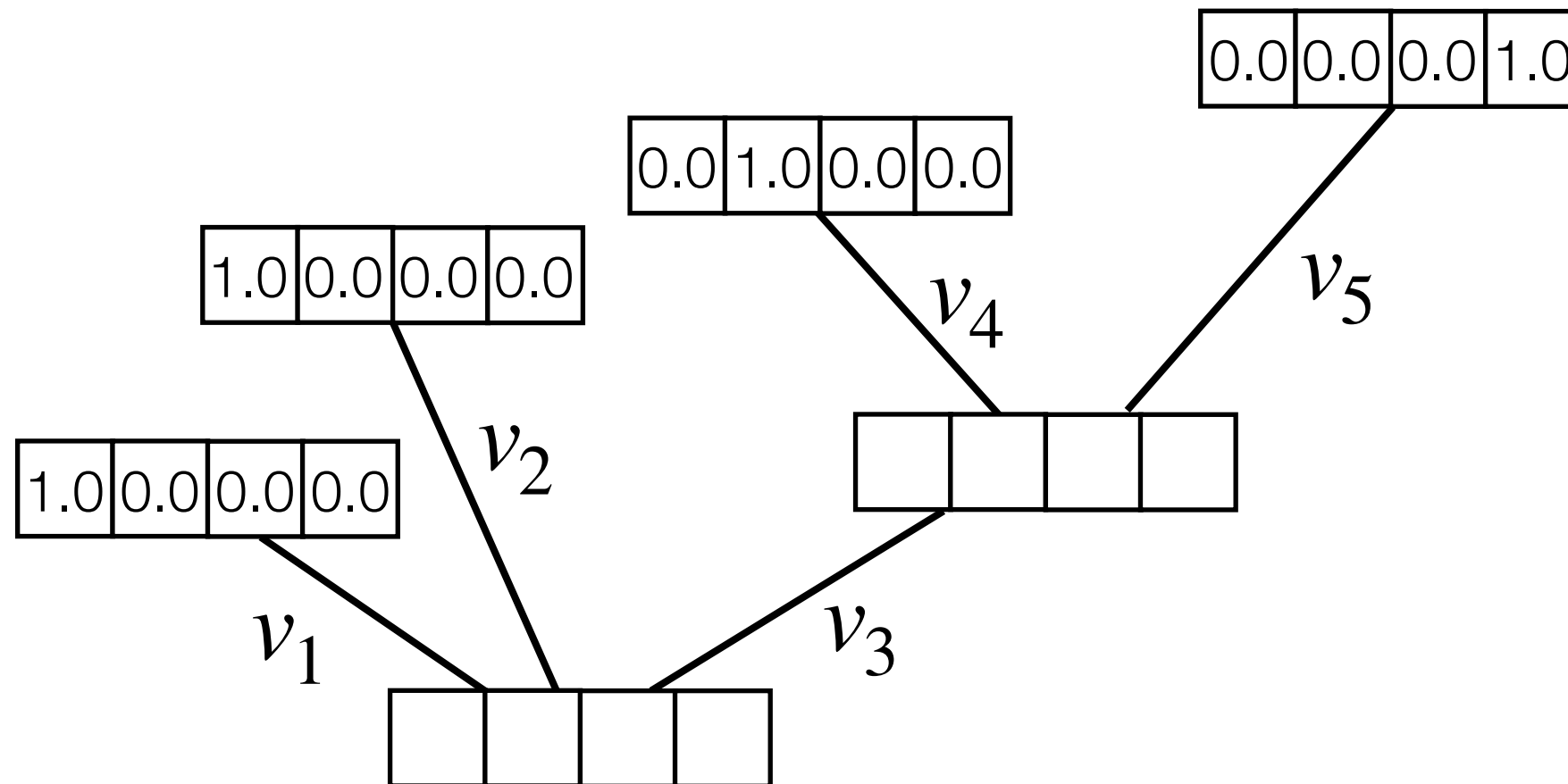
# Felsenstein's Pruning Algorithm



Conditional Likelihoods

What's probability of everything  
above this node if this was a G?

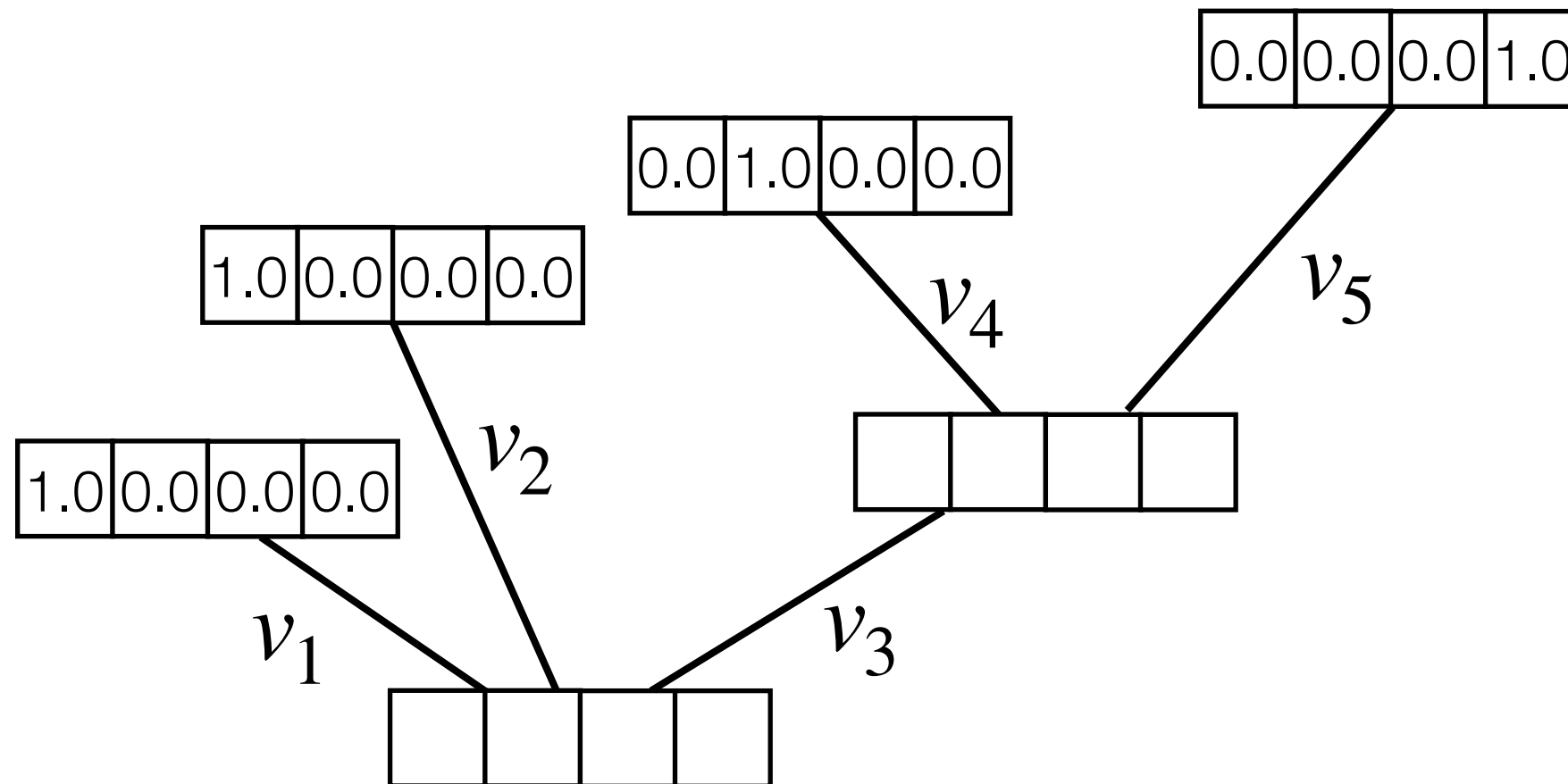
# Felsenstein's Pruning Algorithm



Step 1. Fill in Observations

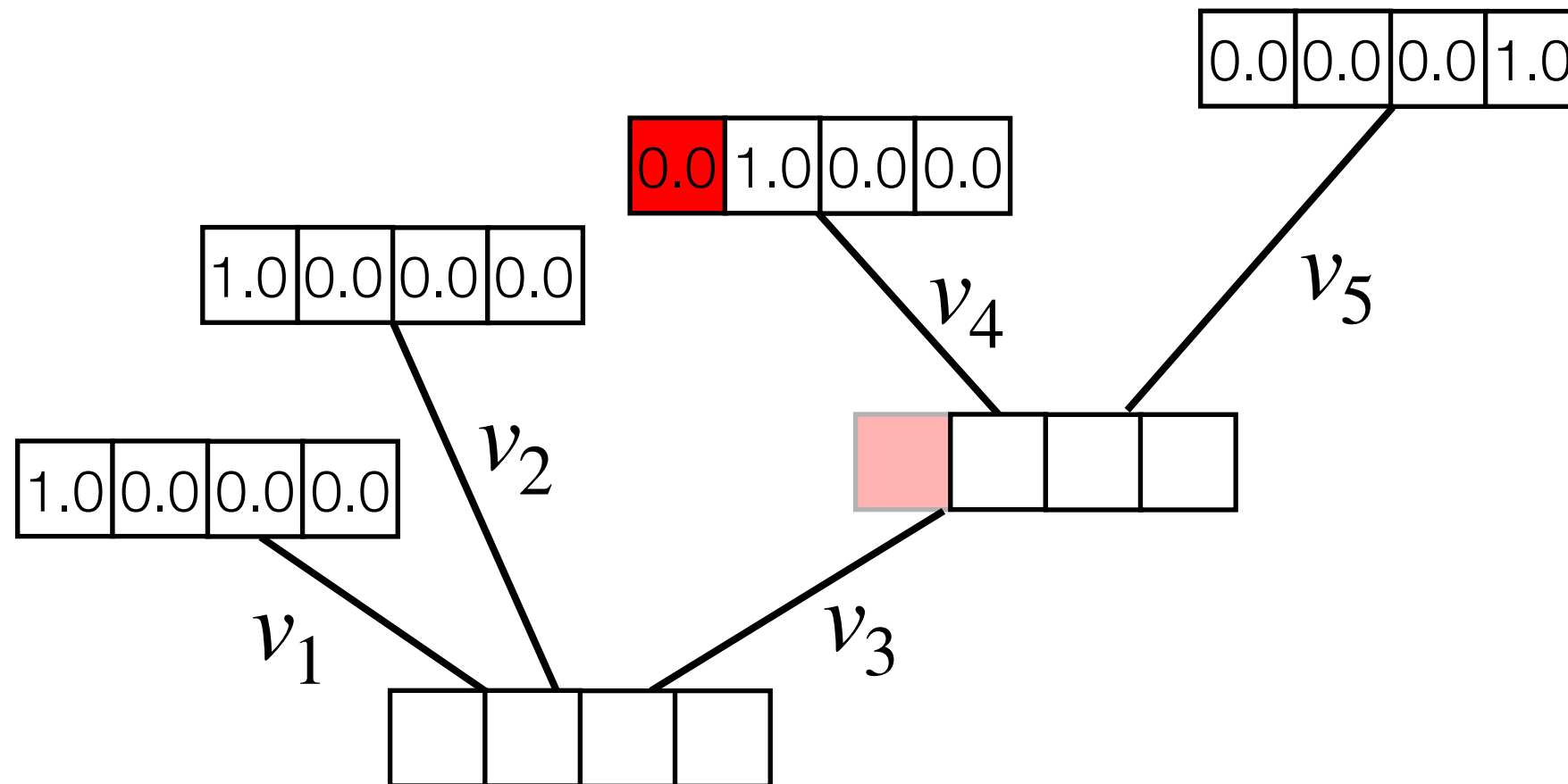
What bases are consistent with your observation?

# Felsenstein's Pruning Algorithm



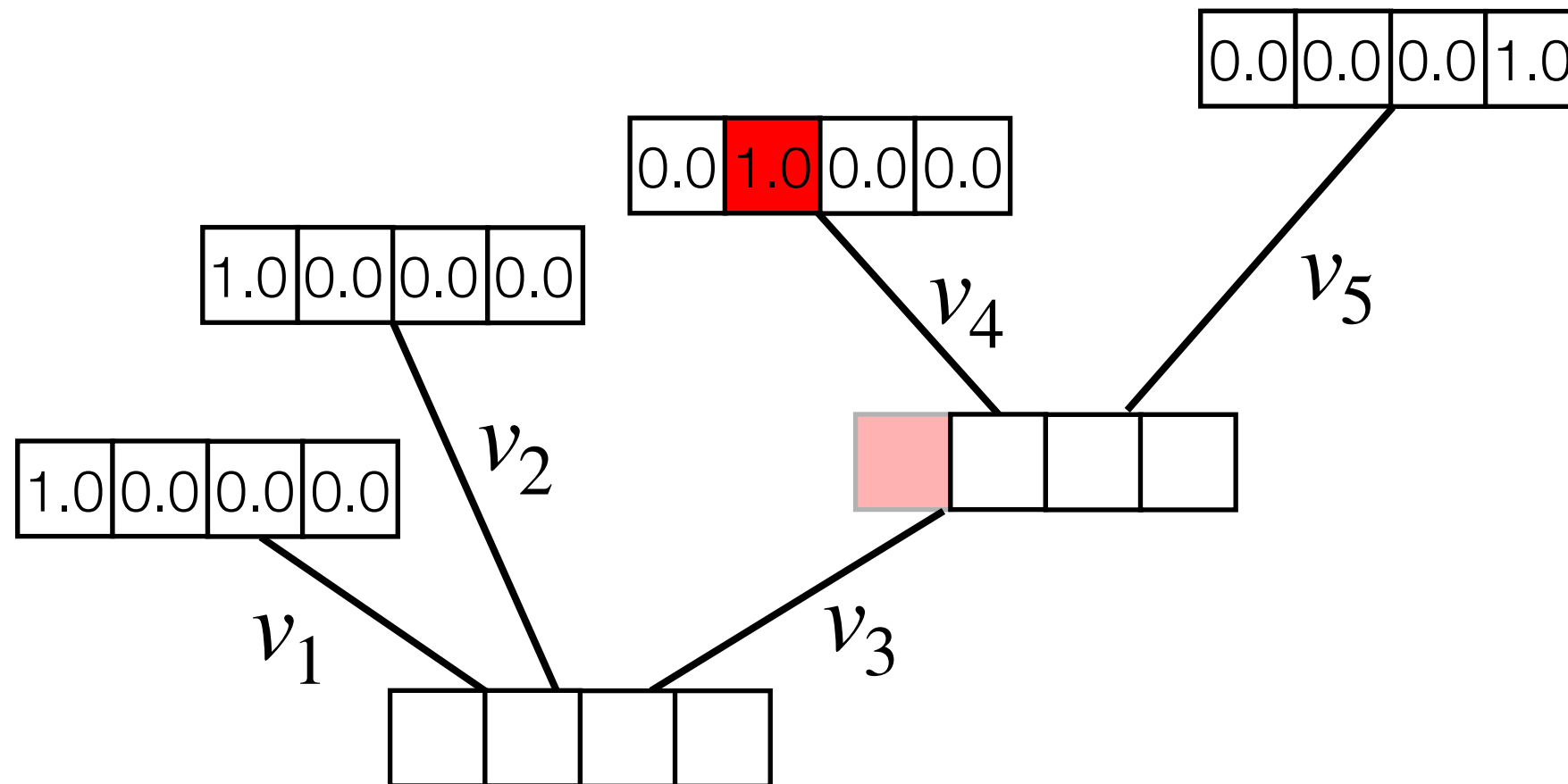
Step 2. Work your way down the tree  
Post-order traversal

# Felsenstein's Pruning Algorithm



Step 2. Work your way down the tree  
Post-order traversal

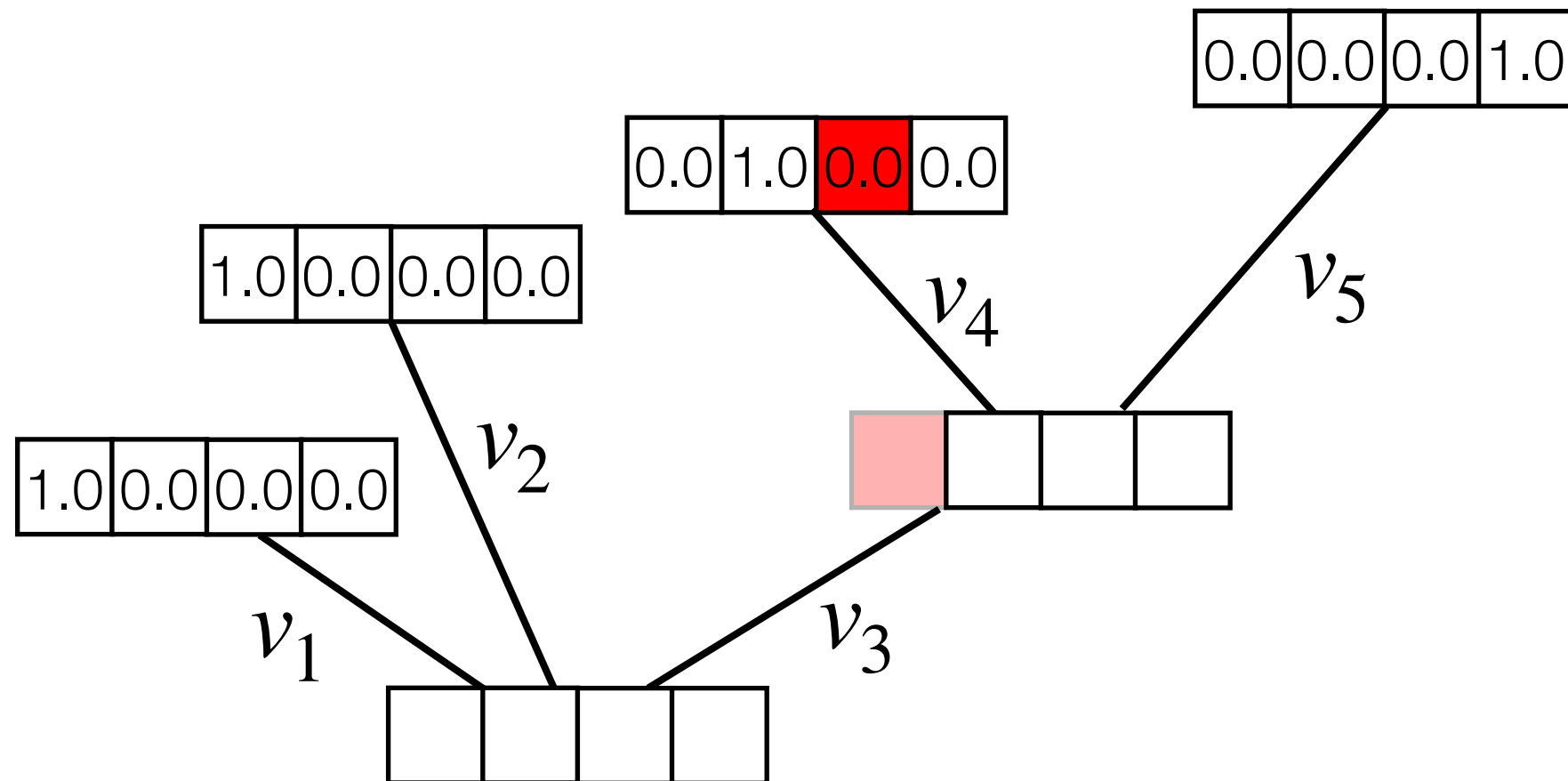
# Felsenstein's Pruning Algorithm



Step 2. Work your way down the tree

Post-order traversal

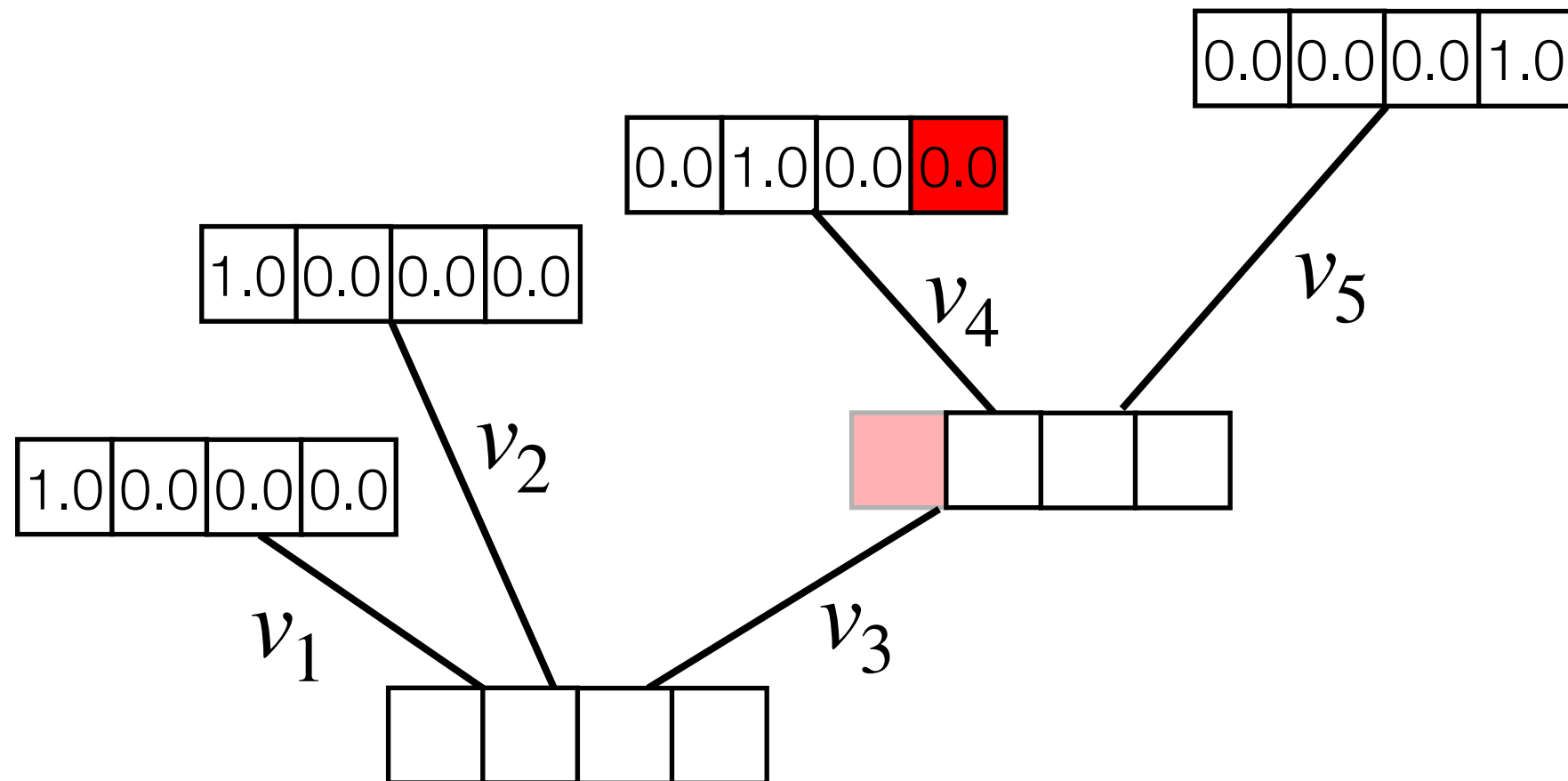
# Felsenstein's Pruning Algorithm



Step 2. Work your way down the tree  
Post-order traversal

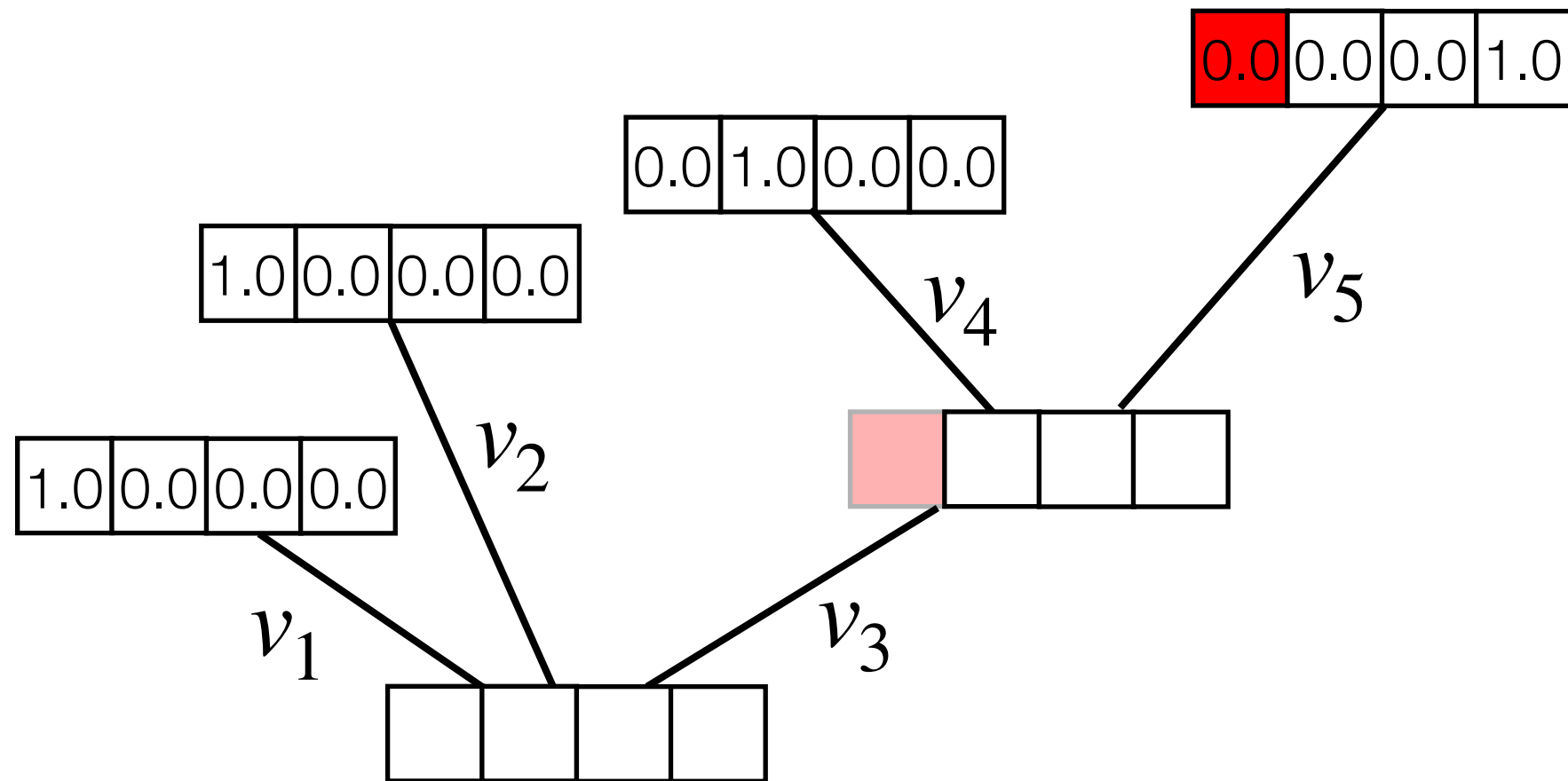


# Felsenstein's Pruning Algorithm



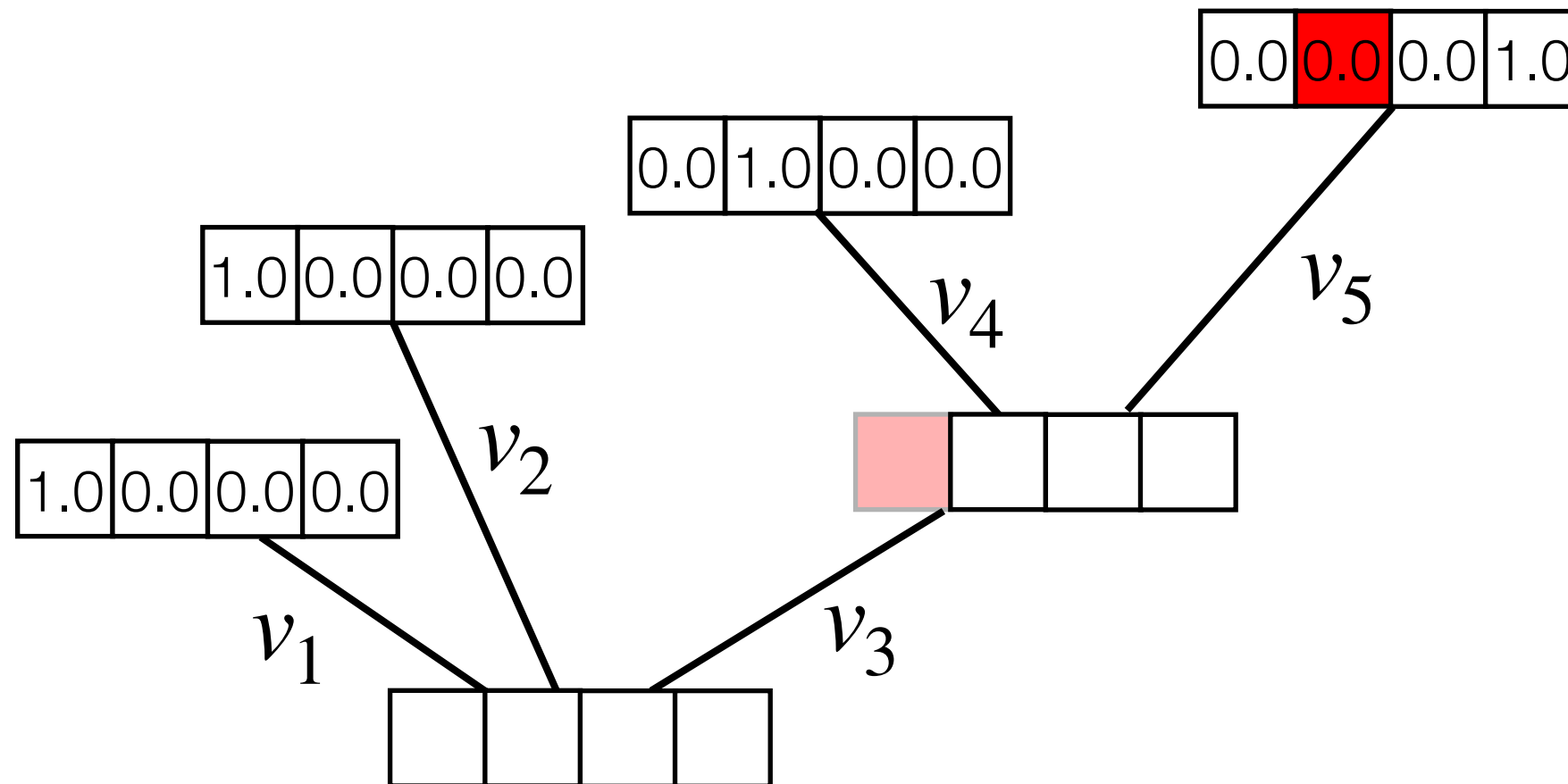
Step 2. Work your way down the tree  
Post-order traversal

# Felsenstein's Pruning Algorithm



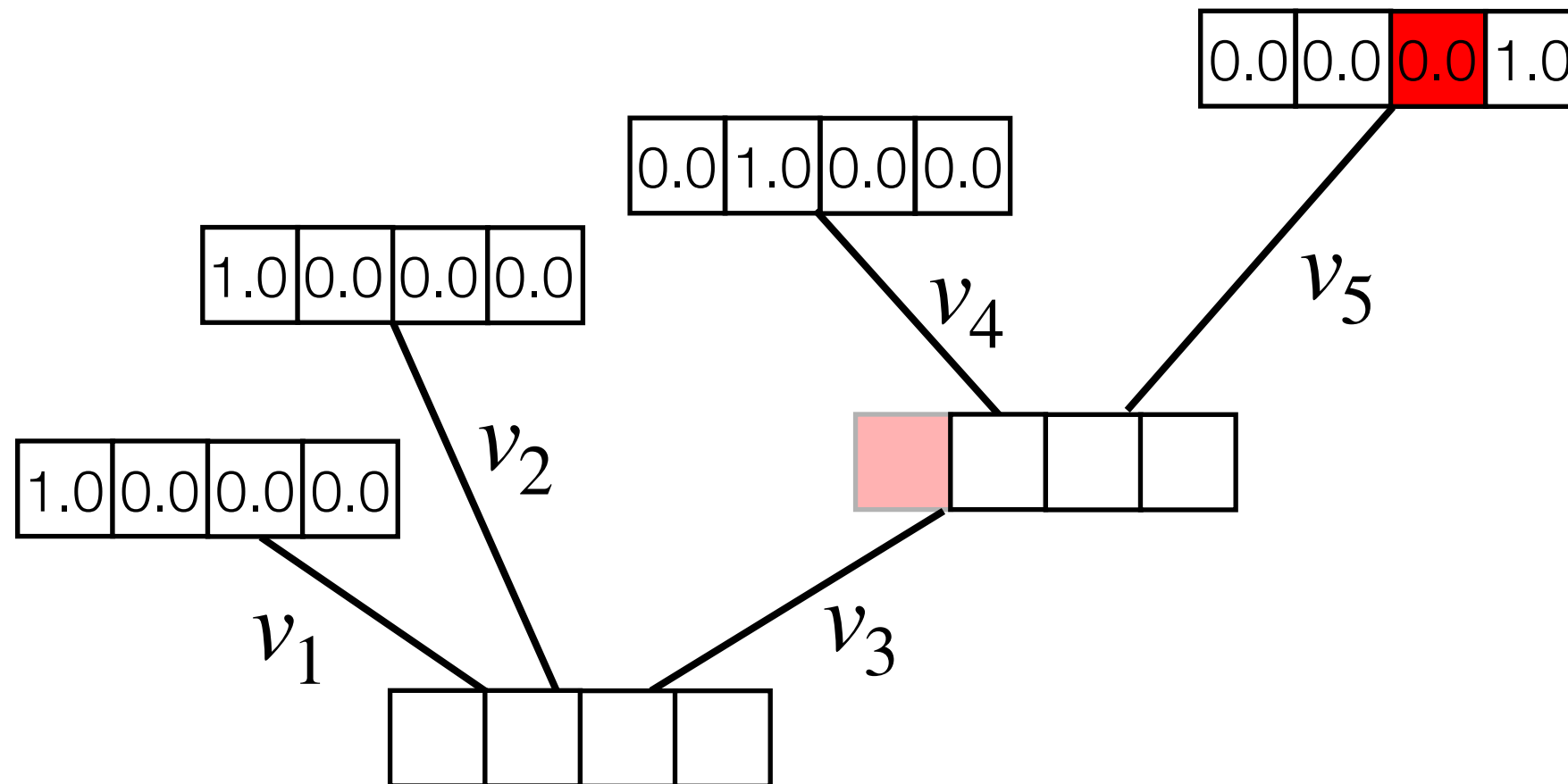
Step 2. Work your way down the tree  
Post-order traversal

# Felsenstein's Pruning Algorithm



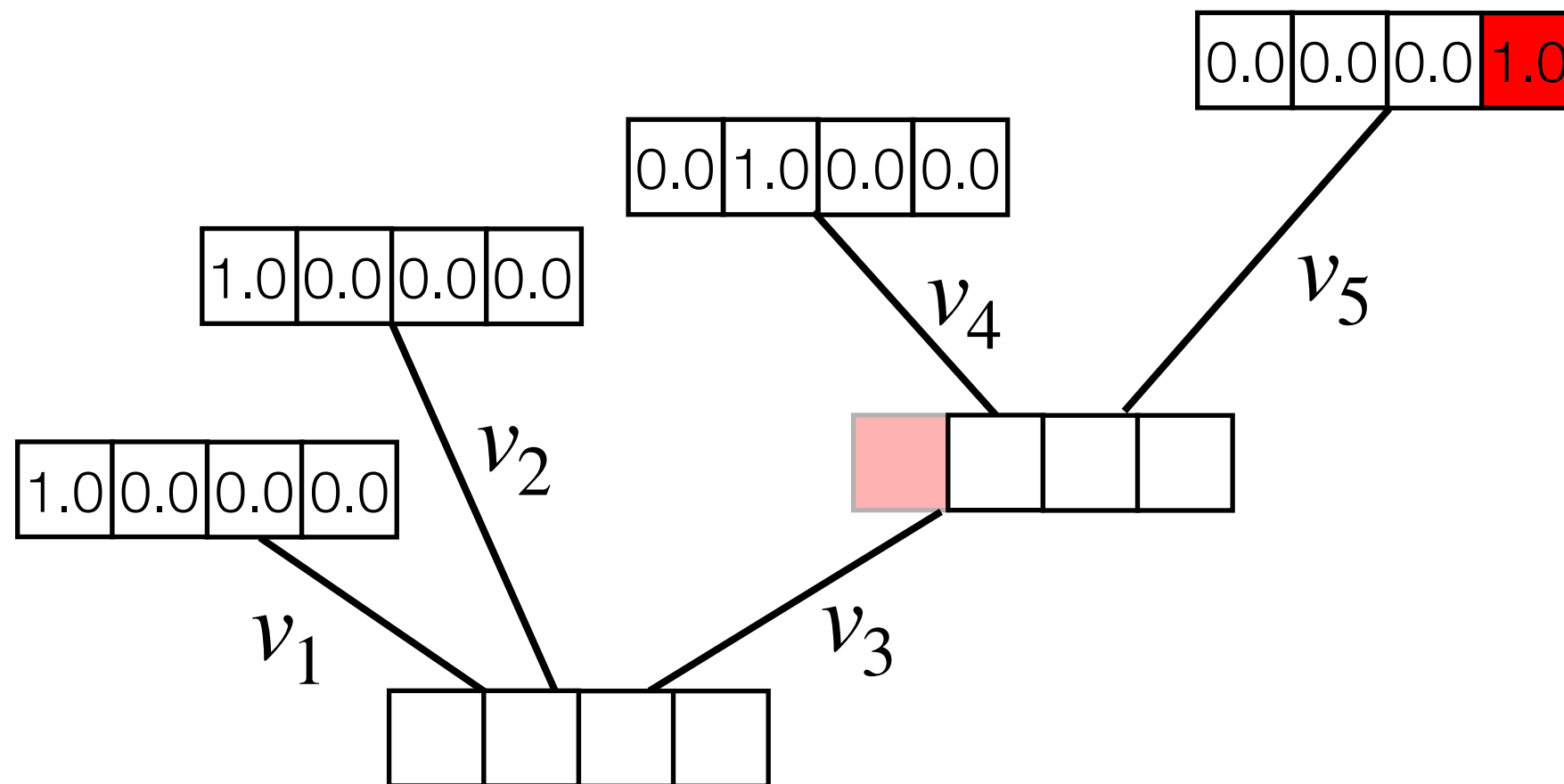
Step 2. Work your way down the tree  
Post-order traversal

# Felsenstein's Pruning Algorithm



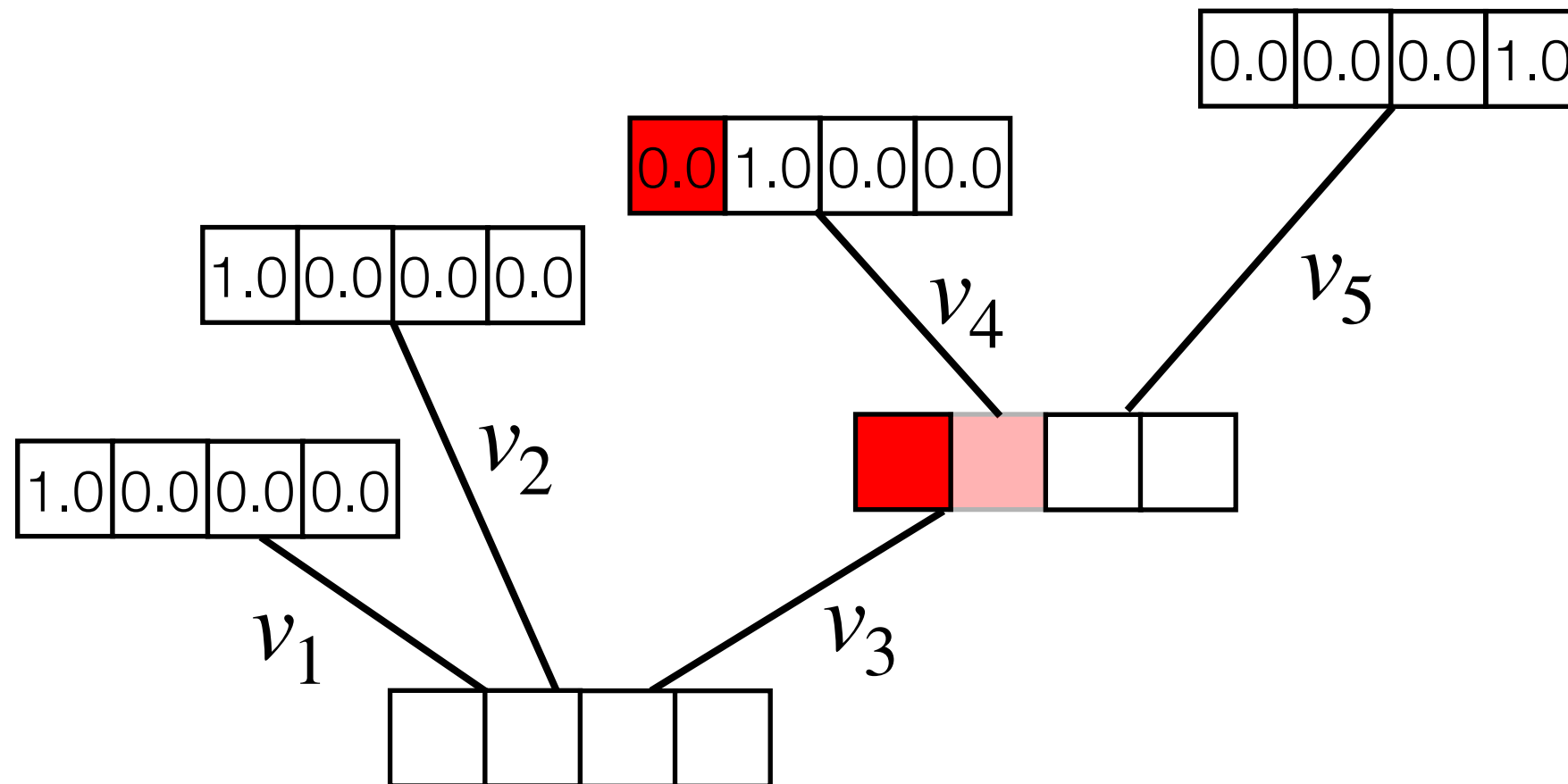
Step 2. Work your way down the tree  
Post-order traversal

# Felsenstein's Pruning Algorithm



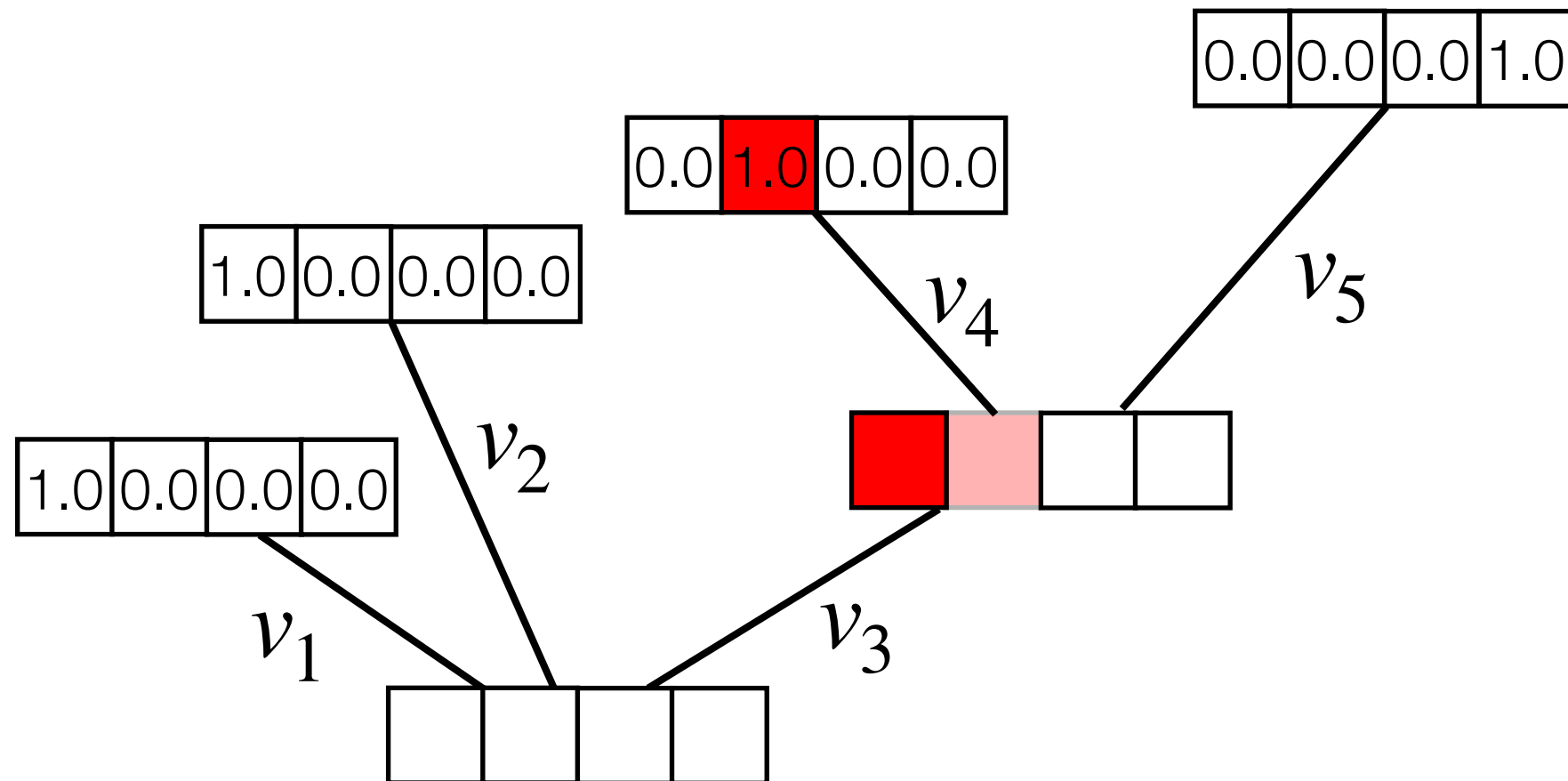
Step 2. Work your way down the tree  
Post-order traversal

# Felsenstein's Pruning Algorithm



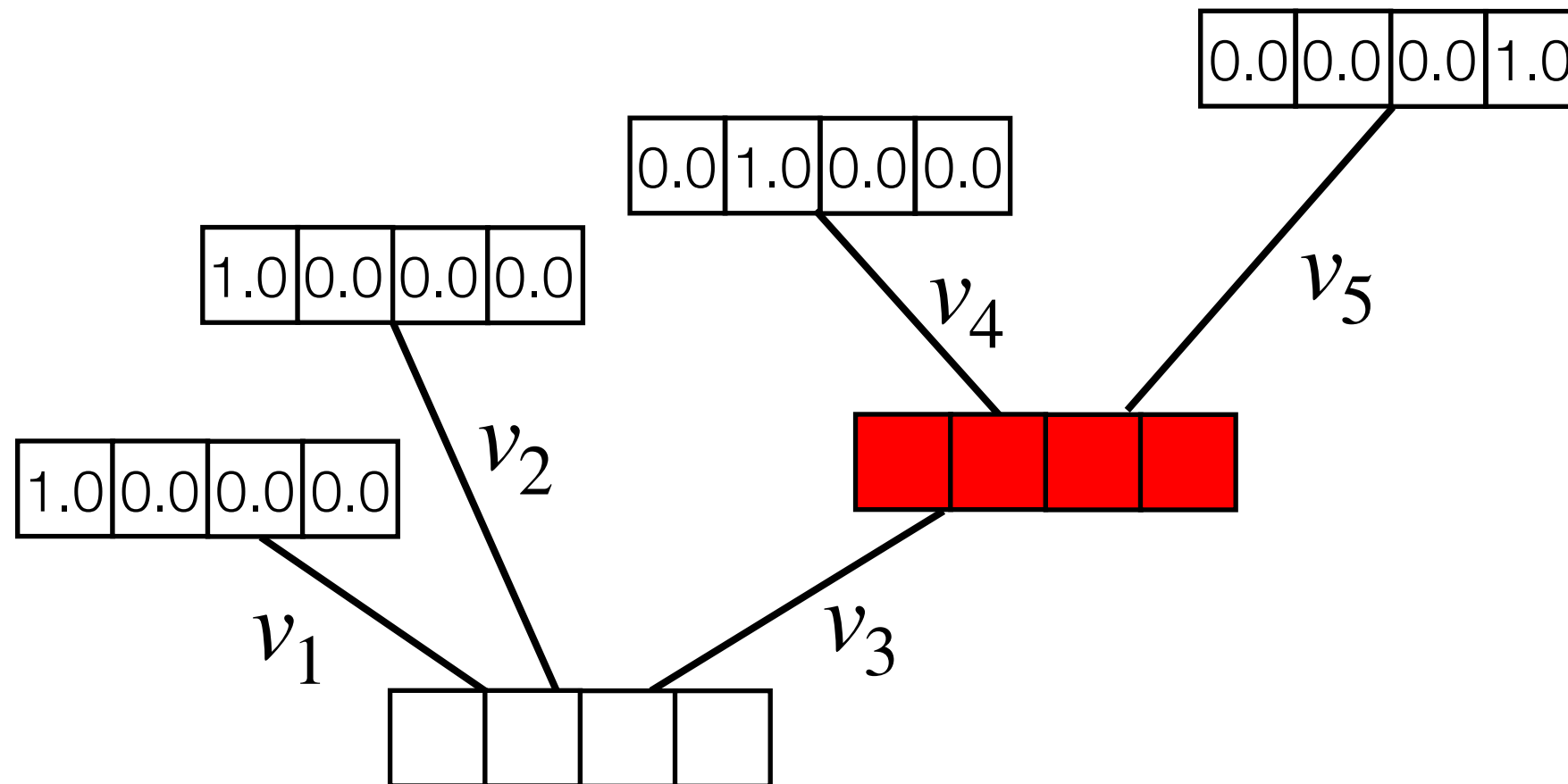
Step 2. Work your way down the tree  
Post-order traversal

# Felsenstein's Pruning Algorithm



Step 2. Work your way down the tree  
Post-order traversal

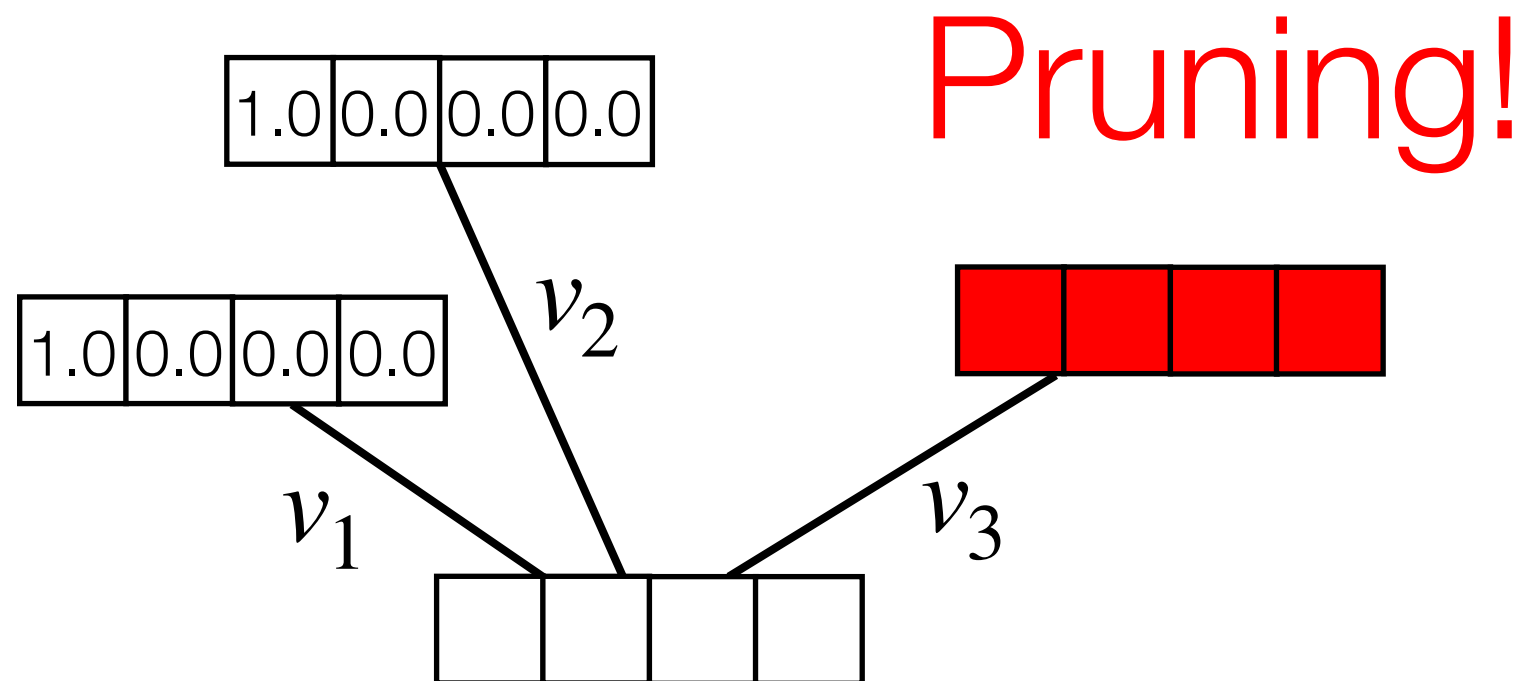
# Felsenstein's Pruning Algorithm



Step 2. Work your way down the tree  
Post-order traversal

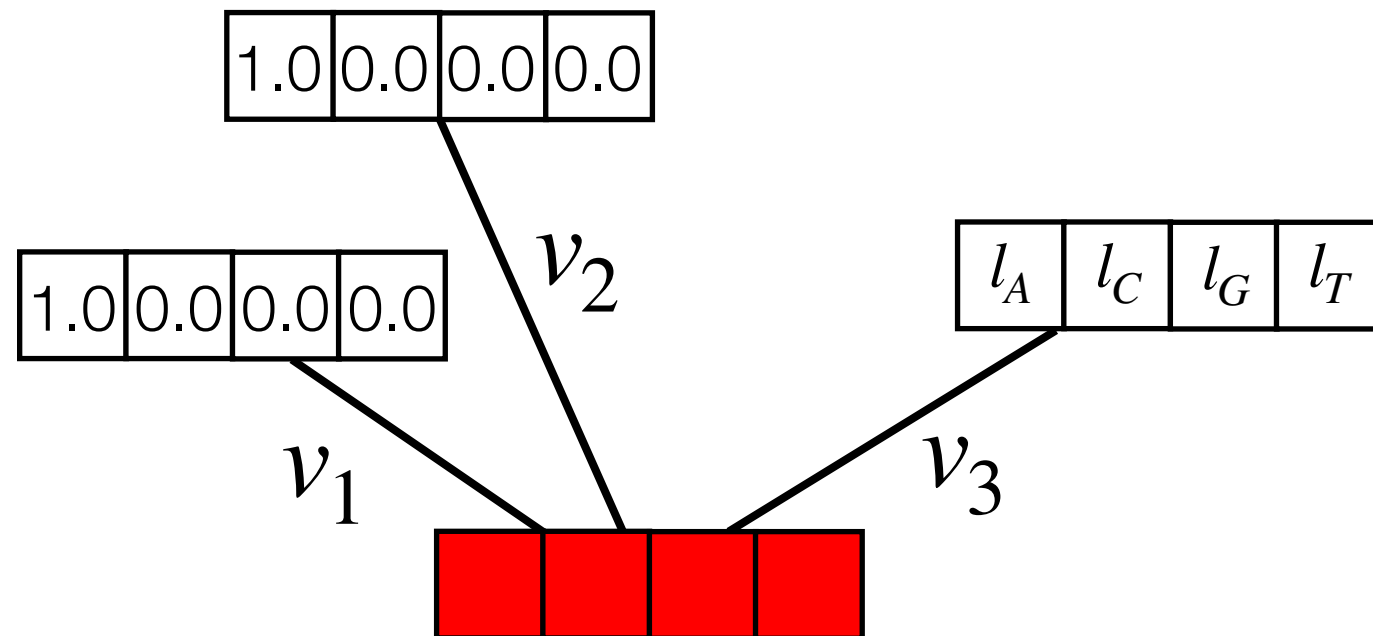


# Felsenstein's Pruning Algorithm



Step 2. Work your way down the tree  
Post-order traversal

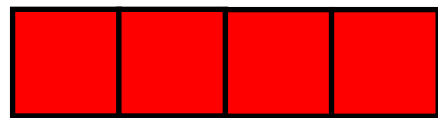
# Felsenstein's Pruning Algorithm



Step 2. Work your way down the tree  
Post-order traversal

# Felsenstein's Pruning Algorithm

$$\mathcal{L} = (\pi_A l_A^{root}) + (\pi_C l_C^{root}) + (\pi_G l_G^{root}) + (\pi_T l_T^{root})$$



Step 3. Weight partial likelihoods at root  
by base frequencies.