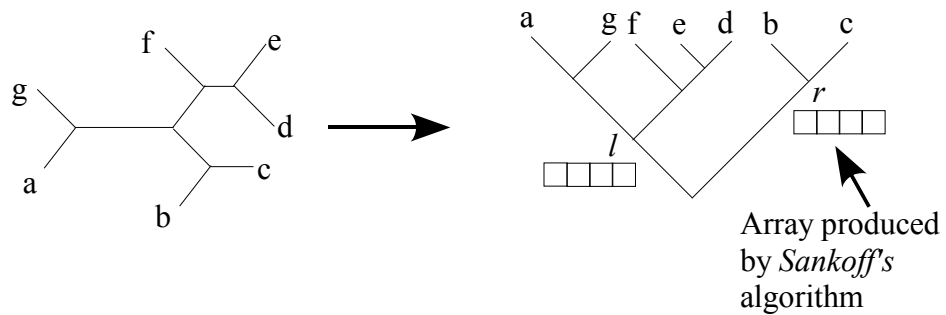


## View

Consider a branch in  $T$ :

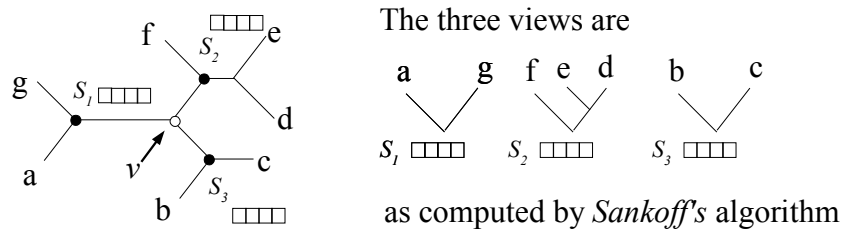
### Example



For every  $E$  of  $T$ , obtain two arrays  $S_l$  and  $S_r$  in this way.  $S_l$  and  $S_r$  are called **views** (for  $E$ ), they summarize the parsimony scores in the subtrees "defined" by  $E$ .

In a binary node, any interior node has three views.

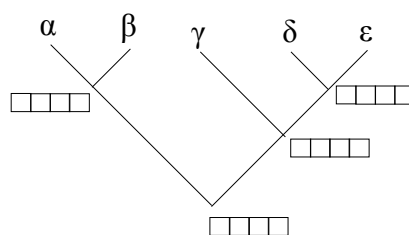
### Example

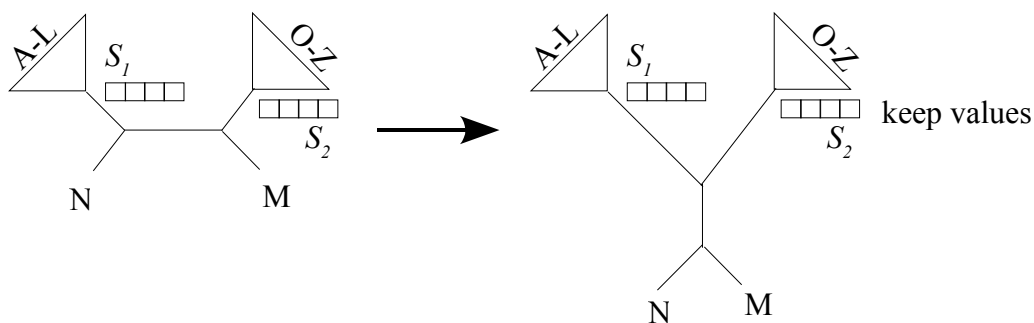


## Views and tree alterations

Sankoff uses one view per internal node.

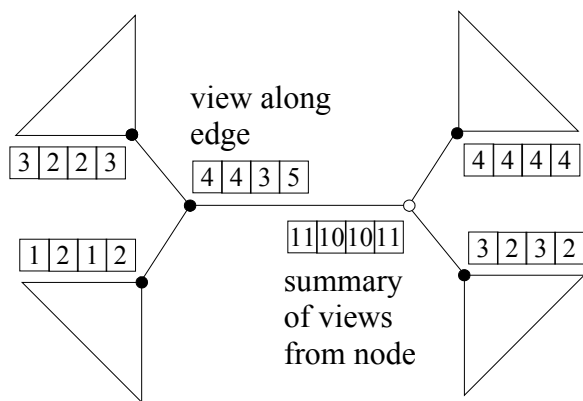
### Example





Views summarize the parsimony scores for those parts of the tree that don't change.

**Example**



Score Matrix

	A	C	G	T
A	0	1	1	1
C	1	0	1	1
G	1	1	0	1
T	1	1	1	0

Re-use views, don't recalculate.

### Chapter 3: How many trees?

General problem: Given a set # of species,  $X = \{x_1, \dots, x_n\}$

$A = \{a_1, \dots, a_n\}$  with  $a_i = a_{i_1}, a_{i_2}, \dots, a_{i_m}$ , m „length“ of sequences.

Find the „best“ tree  $T$  on  $X$ .

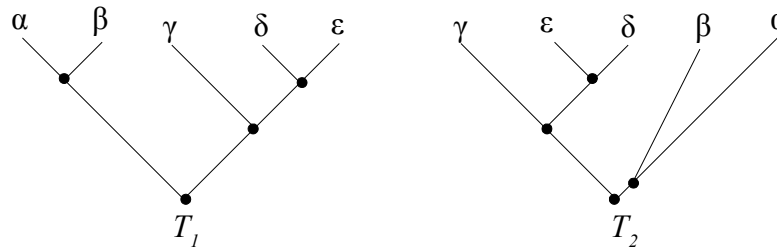
Naive approach:

- Consider every possible tree  $T$  on  $X$
- Compute score
- Return tree with best score

Is this feasible? (No, it is not.)

We will consider the number of “different“ trees on  $n$  species. What does different mean?

**Example**



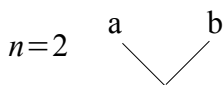
$T_1, T_2$  are of the same topology. For our purposes, we will consider two trees to be different if they have differently labeled topologies.

When counting trees we must distinguish three different cases:

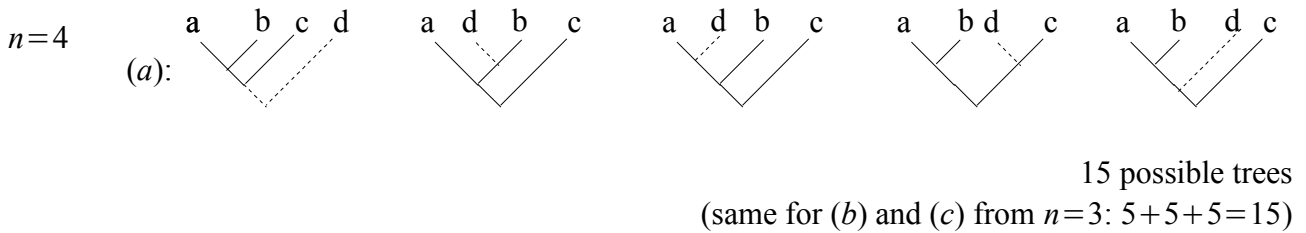
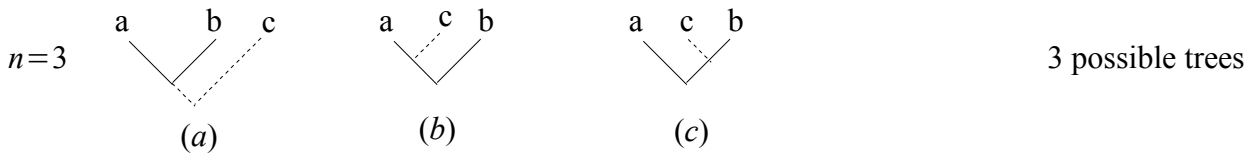
- Rooted or unrooted trees?
- Labeled or unlabeled trees?
- Bifurcating or multifurcating trees?

**Number of rooted, labeled, bifurcating trees:**

Obtainable as follows:



1 possible tree



Why does this work?

*Claim 1:* We get all the trees on  $n$  species in this way

*Claim 2:* All trees obtained are different

*Fact:* Operation of adding  $k$ -th species to tree on species  $1, \dots, k-1$  is inverse to deleting  $k$ -th species from tree on species  $1, \dots, k$ .

1. We get all trees on  $n$  species, because any such tree can be dismantled by first removing species  $n$ , then  $n-1, \dots, 1$ .
2. (There can't be two different ways of taking a tree apart.) Because deleting is inverse to adding.

How many?

- When we add a species, we can do so in each existing edge, including root edge
- Each addition produces two new edges

$n$	# of branches in a tree	# rooted bifurcated trees
2	3	1
3	5	3
4	7	15
5	9	105
6	11	945
7	13	10 395
8	15	134 135
20	39	$8 \cdot 10^{21}$
50	99	$2.7 \cdot 10^{76}$

$R(n)$  = # rooted bifurcating trees on  $n$  species

$$\begin{aligned} &= 3 \cdot 5 \cdot 7 \cdot 9 \cdot \dots \cdot (2n-3) \\ &= (2n-3)!! \\ &= \frac{(2n-3)!}{2^{n-2}(n-2)!} \end{aligned}$$

**Number of unrooted bifurcating trees**

To count these, call the first species “root”, then an unrooted tree on  $n$  species becomes a rooted tree on  $n-1$  species.

$$U(n) = R(n) = 3 \cdot 5 \cdot \dots \cdot (2n-5)$$

**Example**

