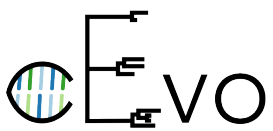
Computational Biology

Lecturers:
Tanja Stadler, Tim Vaughan & Carsten Magnus

Teaching Assistants:
Antoine Zwaans, Adrian Lison,
James Munday & Marcus Overwater

Computational Evolution
Department of Biosystems Science and Engineering

HS 2023



Why compute the likelihood?

- Maximum likelihood tree search;

Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);

Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);
- Why look for the ML tree?

Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);
- Why look for the ML tree?
  - Take higher order sequence correlations into account (unlike phenetic approaches);

Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);
- Why look for the ML tree?
  - Take higher order sequence correlations into account (unlike phenetic approaches);
  - Take hidden mutations into account (unlike cladistic approaches);

Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);
- Why look for the ML tree?
  - Take higher order sequence correlations into account (unlike phenetic approaches);
  - Take hidden mutations into account (unlike cladistic approaches);
- How to get the ML tree?

Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);
- Why look for the ML tree?
  - Take higher order sequence correlations into account (unlike phenetic approaches);
  - Take hidden mutations into account (unlike cladistic approaches);
- How to get the ML tree?
  - iterate over proposed parameter and tree space;

Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);
- Why look for the ML tree?
  - Take higher order sequence correlations into account (unlike phenetic approaches);
  - Take hidden mutations into account (unlike cladistic approaches);
- How to get the ML tree?
  - iterate over proposed parameter and tree space;
  - calculate the (log-)likelihood of each tree;

Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);
- Why look for the ML tree?
  - Take higher order sequence correlations into account (unlike phenetic approaches);
  - Take hidden mutations into account (unlike cladistic approaches);
- How to get the ML tree?
  - iterate over proposed parameter and tree space;
  - calculate the (log-)likelihood of each tree;
  - select the tree with the max (log-)likelihood;

Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);
- Why look for the ML tree?
  - Take higher order sequence correlations into account (unlike phenetic approaches);
  - Take hidden mutations into account (unlike cladistic approaches);
- How to get the ML tree?
  - iterate over proposed parameter and tree space;
  - calculate the (log-)likelihood of each tree;
  - select the tree with the max (log-)likelihood;
- How to compute the (log-)likelihood?

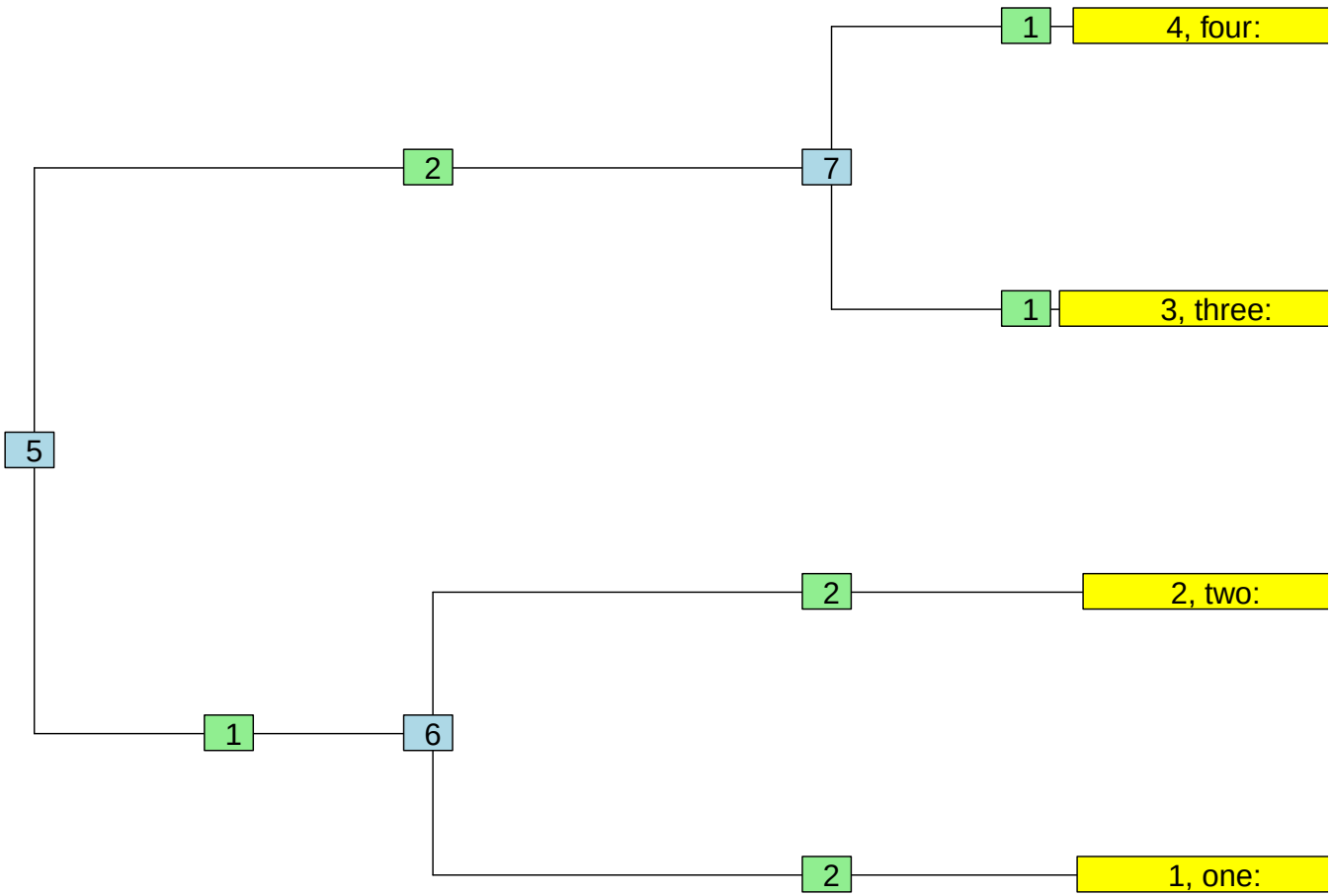
Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);
- Why look for the ML tree?
  - Take higher order sequence correlations into account (unlike phenetic approaches);
  - Take hidden mutations into account (unlike cladistic approaches);
- How to get the ML tree?
  - iterate over proposed parameter and tree space;
  - calculate the (log-)likelihood of each tree;
  - select the tree with the max (log-)likelihood;
- How to compute the (log-)likelihood?
  - Iterate through all possible combinations of nucleotides on the tree nodes ($O(4^{(N-1)})$);

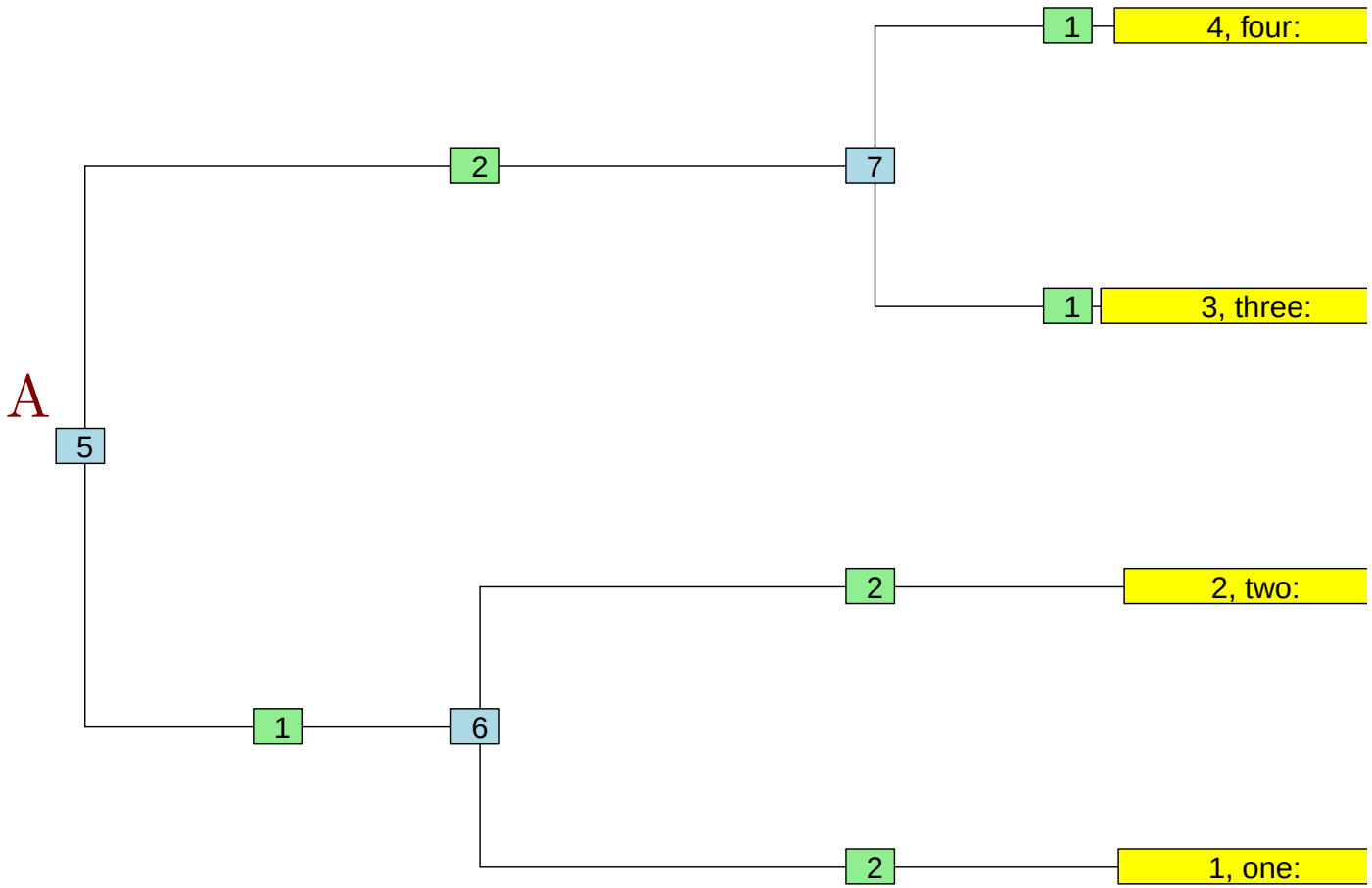
Why compute the likelihood?

- Maximum likelihood tree search;
- Any other likelihood-based tree search (e.g. Bayesian);
- Why look for the ML tree?
  - Take higher order sequence correlations into account (unlike phenetic approaches);
  - Take hidden mutations into account (unlike cladistic approaches);
- How to get the ML tree?
  - iterate over proposed parameter and tree space;
  - calculate the (log-)likelihood of each tree;
  - select the tree with the max (log-)likelihood;
- How to compute the (log-)likelihood?
  - Iterate through all possible combinations of nucleotides on the tree nodes ($O(4^{(N-1)})$);
  - Felsenstein's tree pruning algorithm.

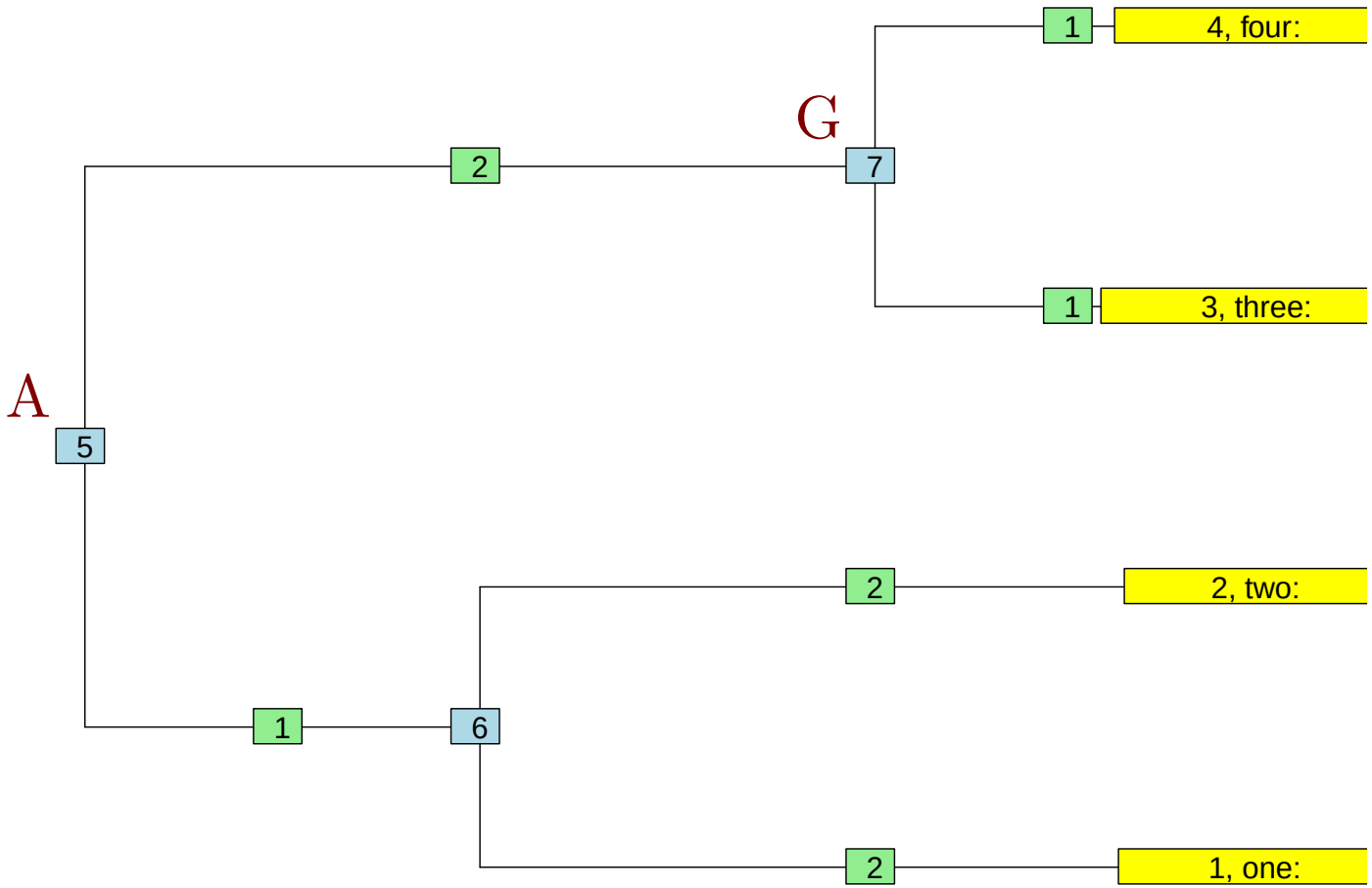
Tree traversal to simulate nucleotide evolution



Tree traversal to simulate nucleotide evolution

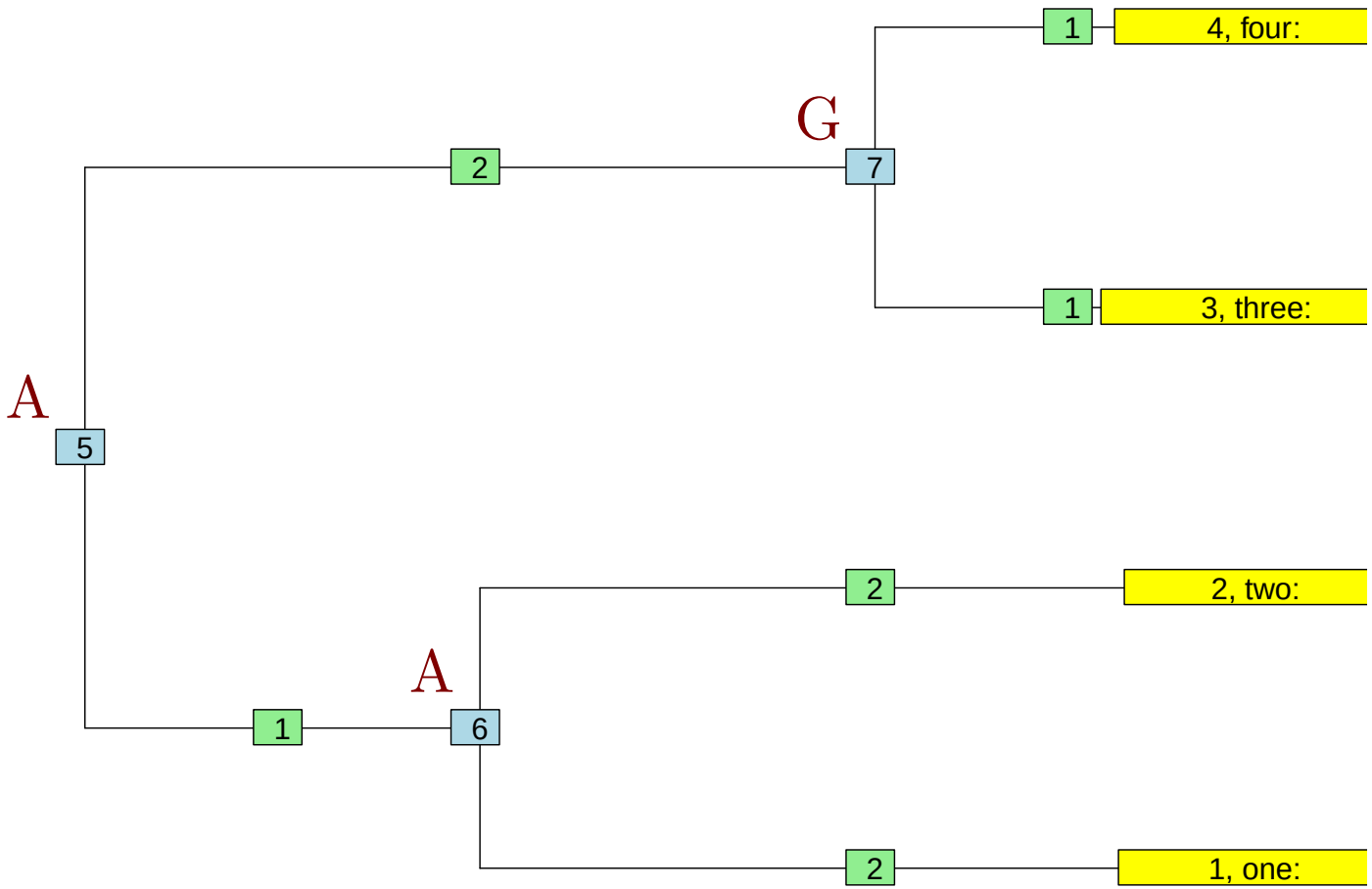


Tree traversal to simulate nucleotide evolution



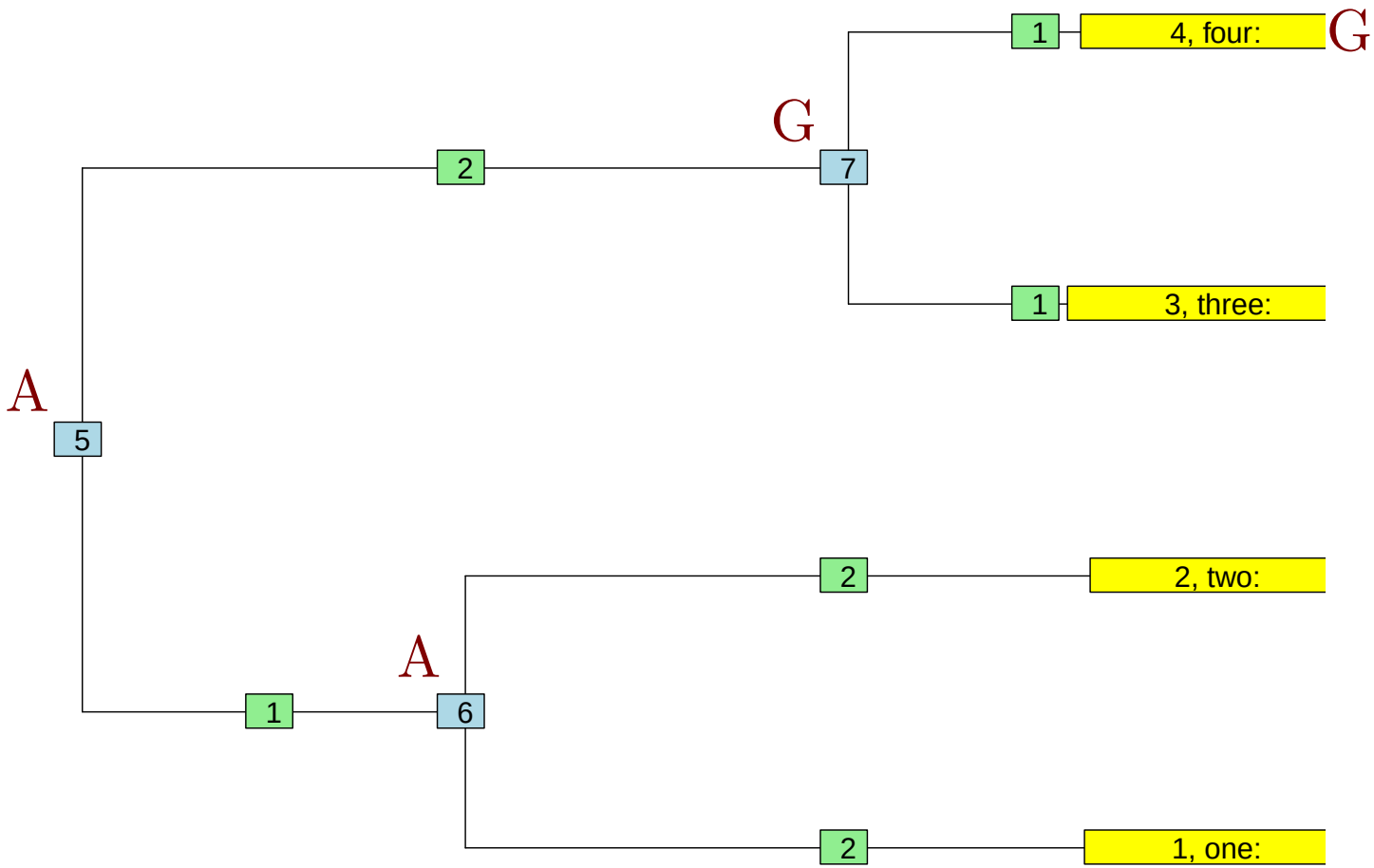
Tree traversal to simulate nucleotide evolution

Always visit the parent node before the two children



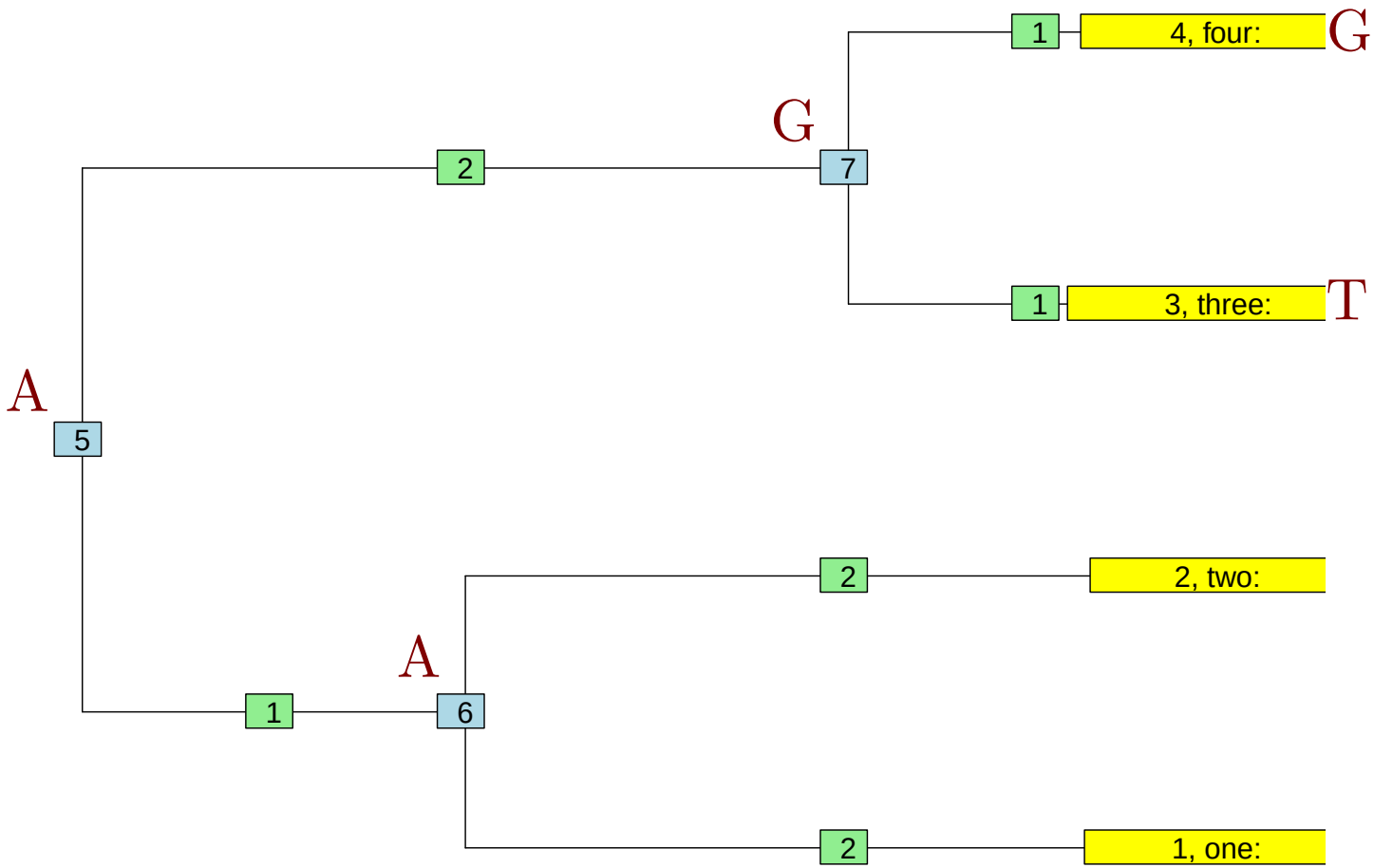
Tree traversal to simulate nucleotide evolution

Always visit the parent node before the two children



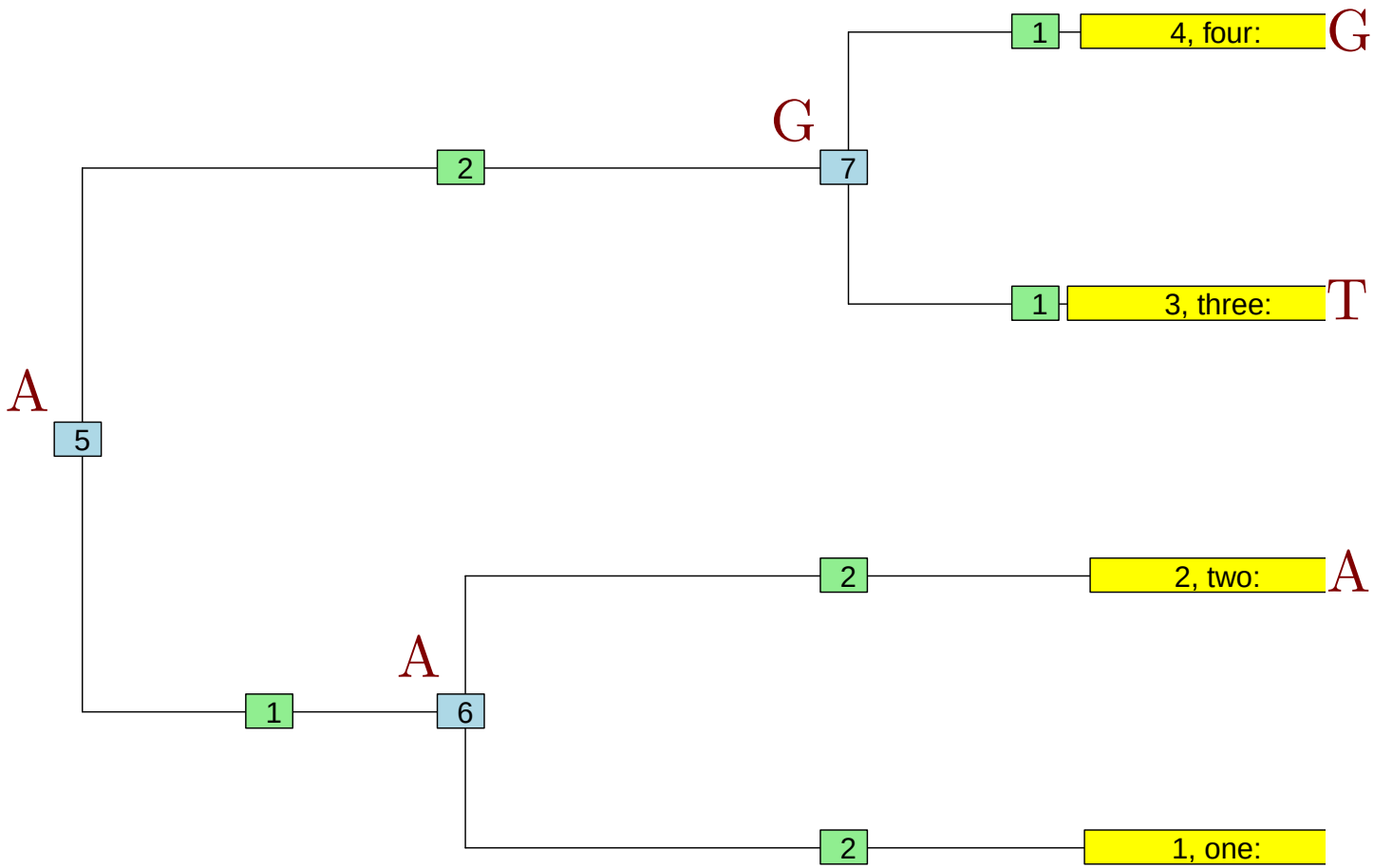
Tree traversal to simulate nucleotide evolution

Always visit the parent node before the two children



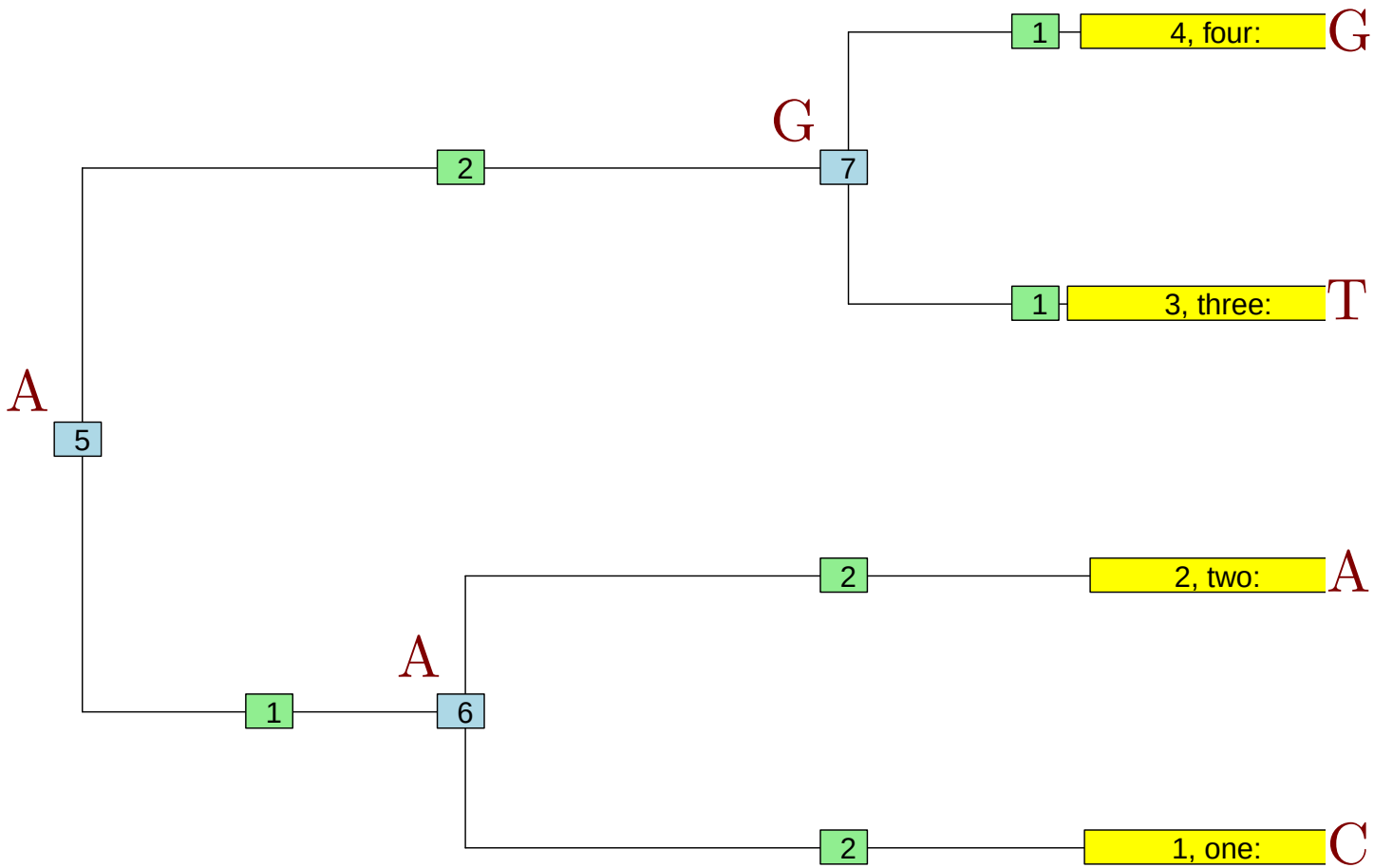
Tree traversal to simulate nucleotide evolution

Always visit the parent node before the two children



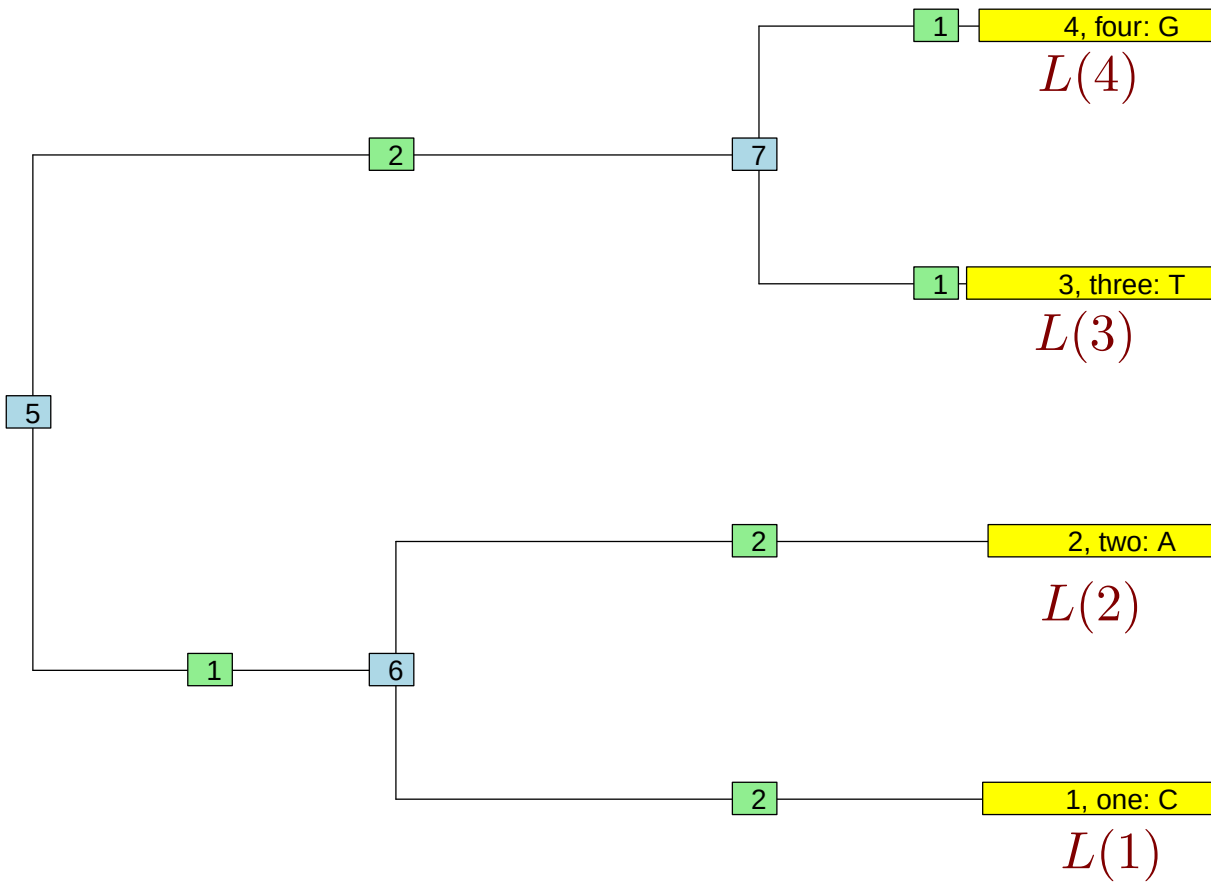
Tree traversal to simulate nucleotide evolution

Always visit the parent node before the two children



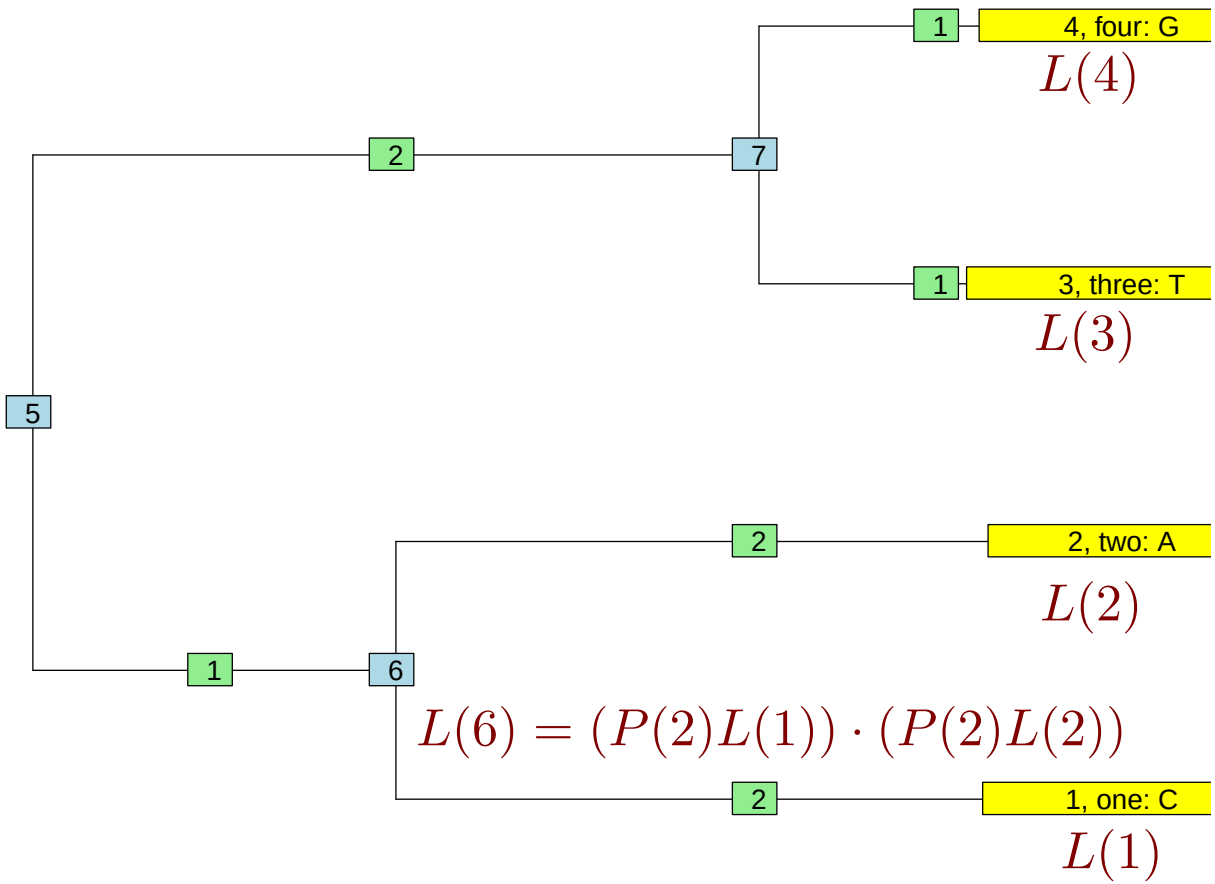
Tree traversal for likelihood computation

We look for the probability of the data below by each node, given its nucleotide state.



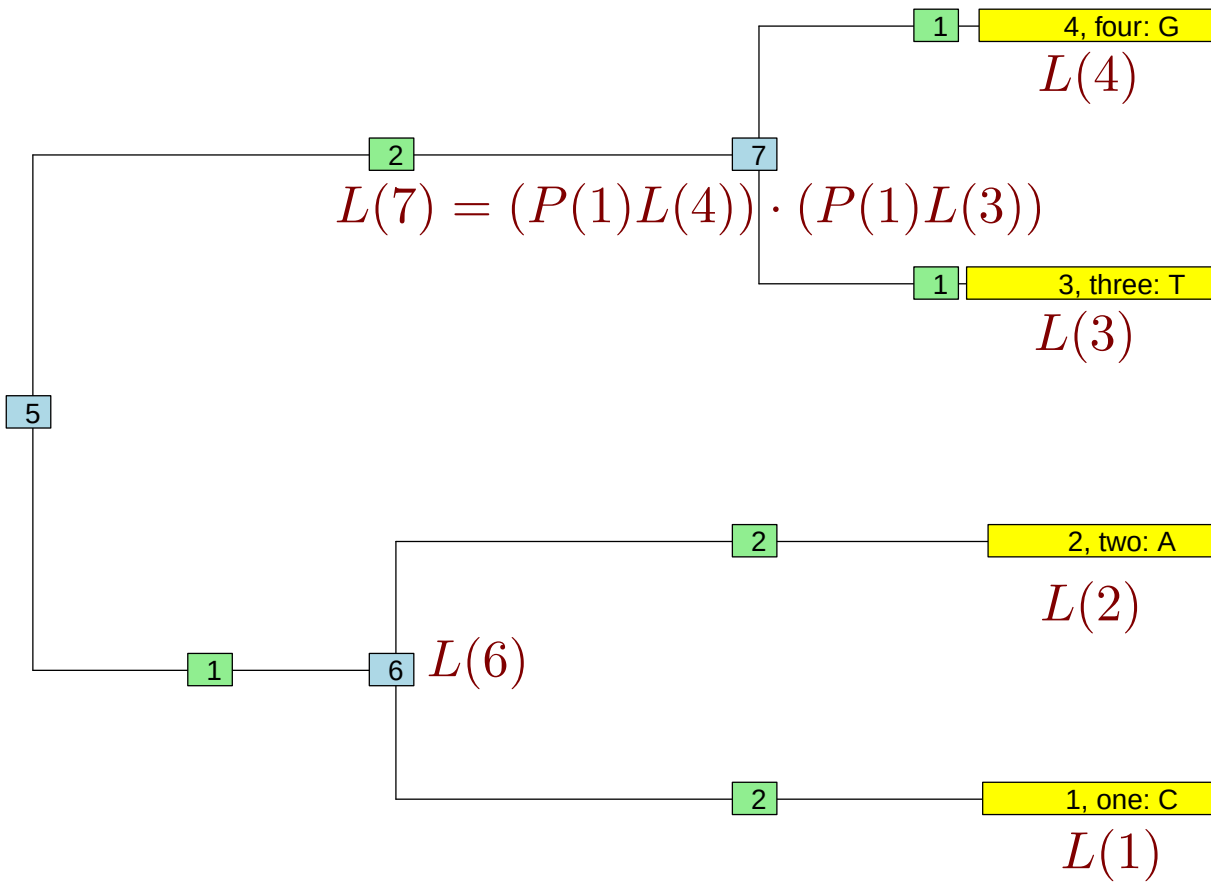
Tree traversal for likelihood computation

We look for the probability of the data below by each node, given its nucleotide state.



Tree traversal for likelihood computation

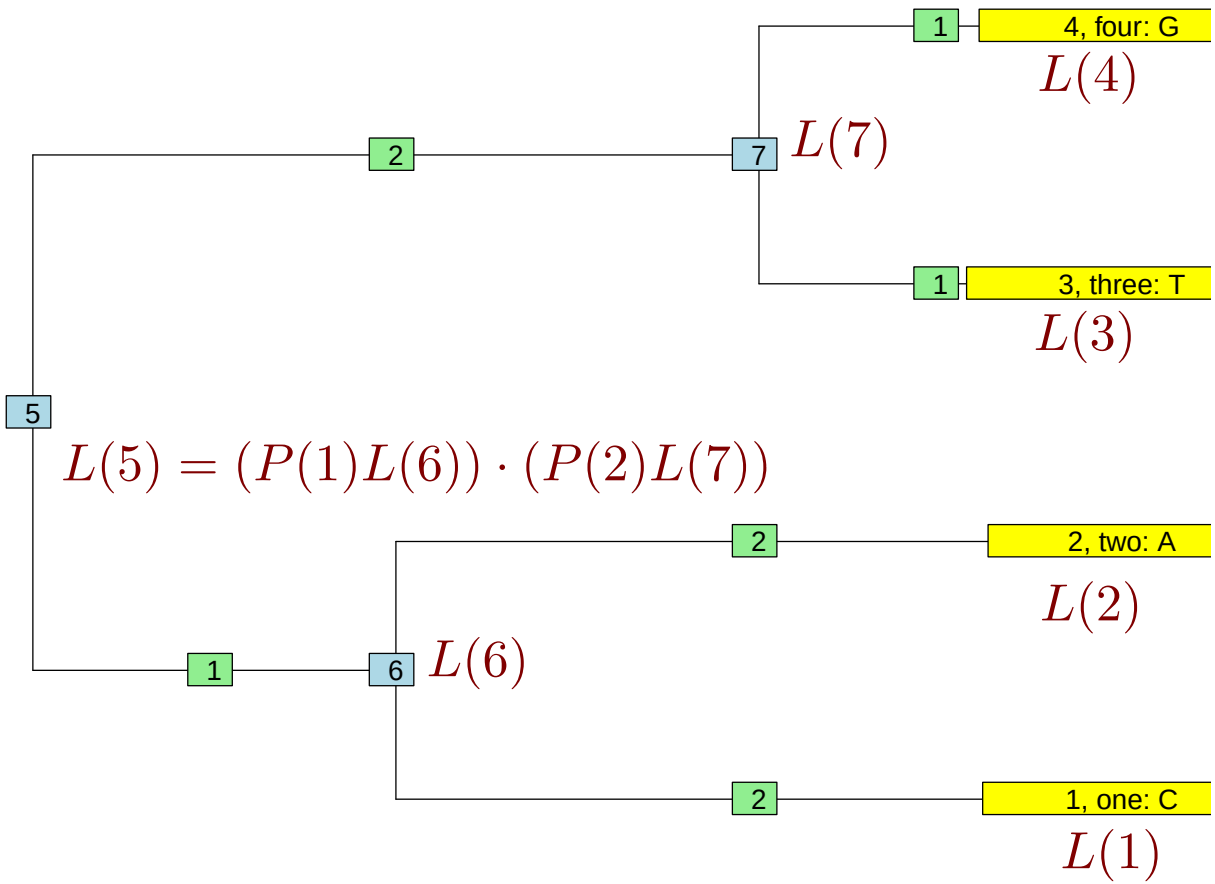
We look for the probability of the data below by each node, given its nucleotide state.



Tree traversal for likelihood computation

We look for the probability of the data below by each node, given its nucleotide state.

Always visit the two children before the parent

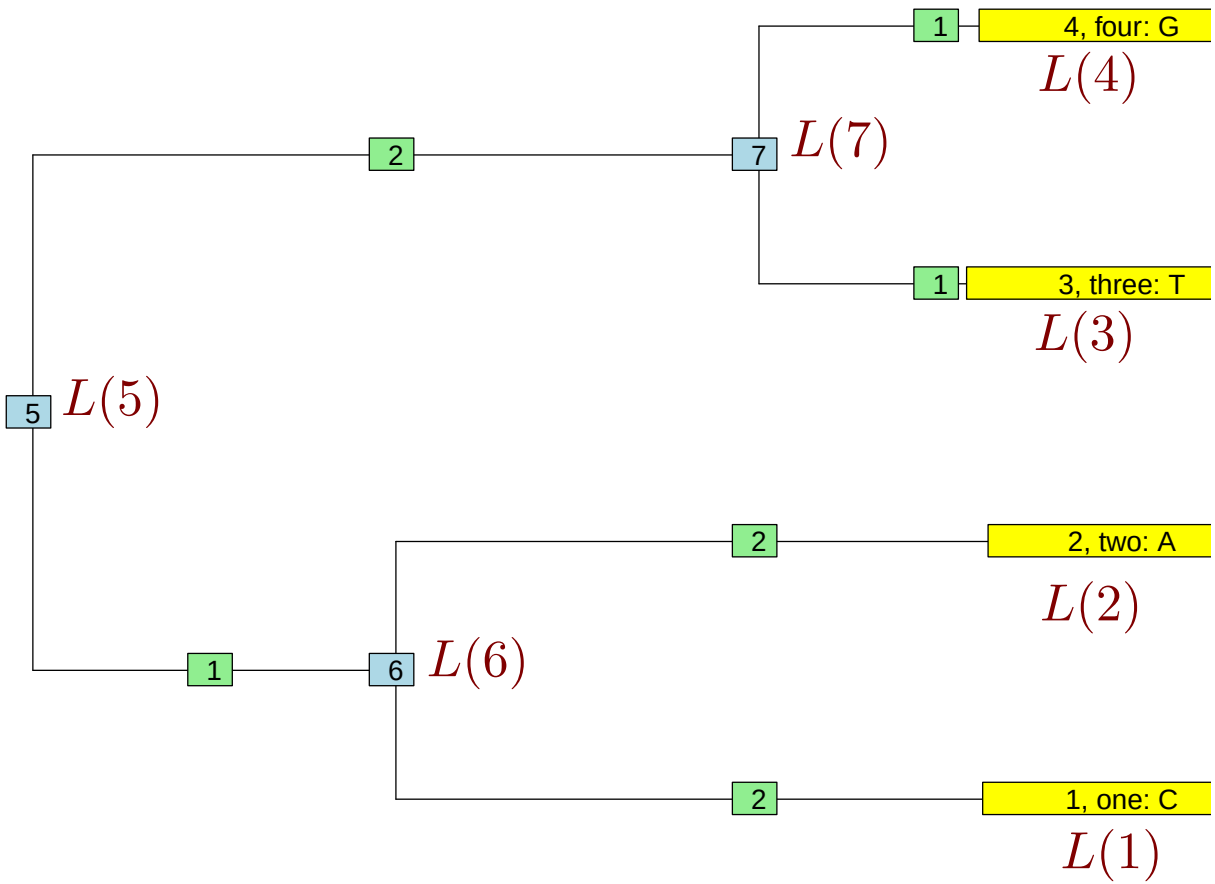


Tree traversal for likelihood computation

We look for the probability of the data below by each node, given its nucleotide state.

Always visit the two children before the parent

$L = \pi L(5)$



Tree node/nucleotide likelihood computation

For each tree node k, for each site i (1 to N), and for each nucleotide X, we need to compute:

$$L_X^{(i)}(k) := \text{Probability of observing the sequence site } i \text{ among tips descending from node } k \text{ conditioned on node } k \text{ having nucleotide } X$$

Tree node/nucleotide likelihood computation

For each tree node k, for each site i (1 to N), and for each nucleotide X, we need to compute:

$L_X^{(i)}(k)$:= Probability of observing the sequence site i
among tips descending from node k
conditioned on node k having nucleotide X

	T	C	A	G
node **k**				
1	?	?	?	?
2	?	?	?	?
...				
N	?	?	?	?

Tree node/nucleotide likelihood computation

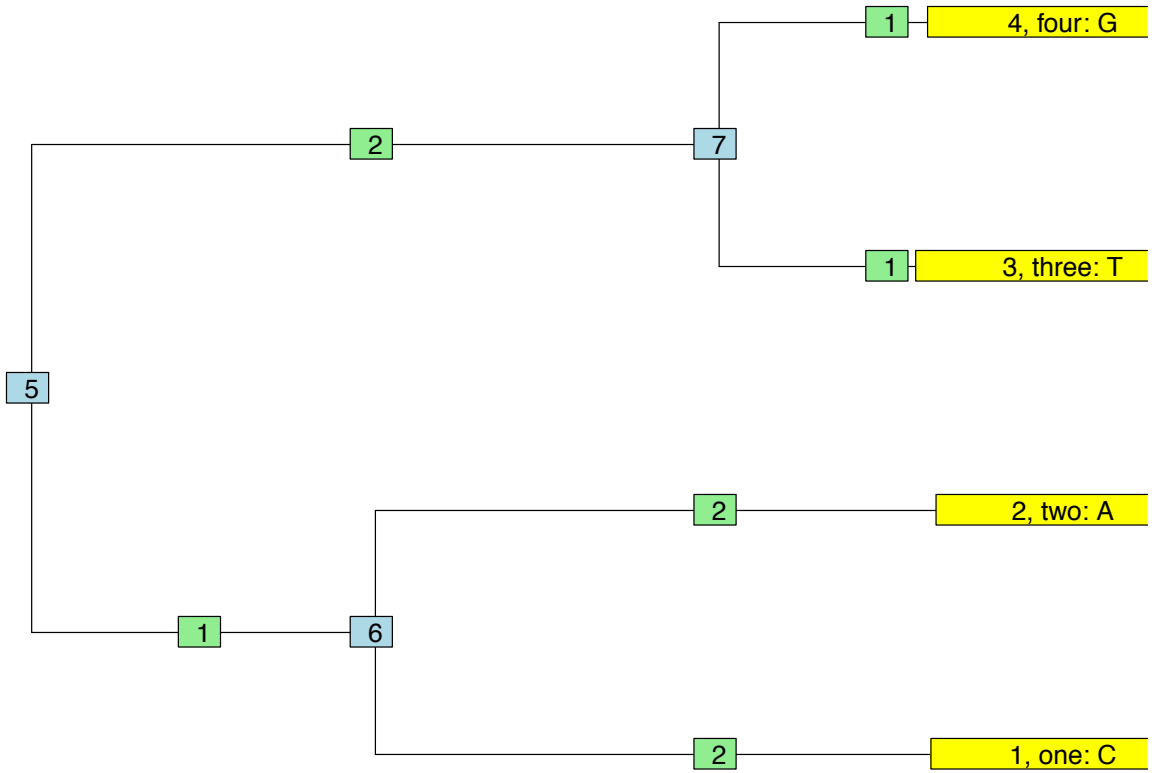
For each tree node k, for each site i (1 to N), and for each nucleotide X, we need to compute:

$$L_X^{(i)}(k) := \text{Probability of observing the sequence site } i \text{ among tips descending from node } k \text{ conditioned on node } k \text{ having nucleotide } X$$

	T	C	A	G
node **k**				
1	?	?	?	?
2	?	?	?	?
...				
N	?	?	?	?

For each node in the tree we need to know the conditional likelihoods of both child nodes.

Felsenstein's pruning algorithm: Example



Alignment:

```
sequences=list(one="C", two="A", three="T", four="G")
```

Felsenstein's pruning algorithm: Tips

4, four: G

3, three: T

2, two: A

1, one: C

if n *is a tip* **then**
  for $i \leftarrow 1$ **to** N **do**
    $L^{(i)}(n) \leftarrow [0, 0, 0, 0]$;
    $L^{(i)}_{A[n,i]}(n) \leftarrow 1$;
  end
else
  ...;
end

Felsenstein's pruning algorithm: Tips

4, four: G

3, three: T

2, two: A

1, one: C

if n *is a tip* **then**
  for $i \leftarrow 1$ **to** N **do**
    $L^{(i)}(n) \leftarrow [0, 0, 0, 0]$;
    $L^{(i)}_{A[n,i]}(n) \leftarrow 1$;
  end
else
  ...;
end

	T	C	A	G
node **1**				
1				
node **2**				
1				
node **3**				
1				
node **4**				
1				

Felsenstein's pruning algorithm: Tips

4, four: G

3, three: T

2, two: A

1, one: C

if n *is a tip* **then**
  for $i \leftarrow 1$ **to** N **do**
    $L^{(i)}(n) \leftarrow [0, 0, 0, 0]$;
    $L^{(i)}_{A[n,i]}(n) \leftarrow 1$;
  end
else
  ...;
end

	T	C	A	G
node **1**				
1	0	0	0	0
node **2**				
1				
node **3**				
1				
node **4**				
1				

Felsenstein's pruning algorithm: Tips

4, four: G

3, three: T

2, two: A

1, one: C

if n *is a tip* **then**
  for $i \leftarrow 1$ **to** N **do**
    $L^{(i)}(n) \leftarrow [0, 0, 0, 0]$;
    $L^{(i)}_{A[n,i]}(n) \leftarrow 1$;
  end
else
  ...;
end

	T	C	A	G
node **1**				
1	0	**1**	0	0
node **2**				
1				
node **3**				
1				
node **4**				
1				

Felsenstein's pruning algorithm: Tips

4, four: G

3, three: T

2, two: A

1, one: C

if n *is a tip* **then**
  for $i \leftarrow 1$ **to** N **do**
    $L^{(i)}(n) \leftarrow [0, 0, 0, 0]$;
    $L^{(i)}_{A[n,i]}(n) \leftarrow 1$;
  end
else
  ...;
end

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	**1**	0
node **3**				
1				
node **4**				
1				

Felsenstein's pruning algorithm: Tips

4, four: G

3, three: T

2, two: A

1, one: C

if n *is a tip* **then**
  for $i \leftarrow 1$ **to** N **do**
    $L^{(i)}(n) \leftarrow [0, 0, 0, 0]$;
    $L^{(i)}_{A[n,i]}(n) \leftarrow 1$;
  end
else
  ...;
end

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **3**				
1	**1**	0	0	0
node **4**				
1				

Felsenstein's pruning algorithm: Tips

4, four: G

3, three: T

2, two: A

1, one: C

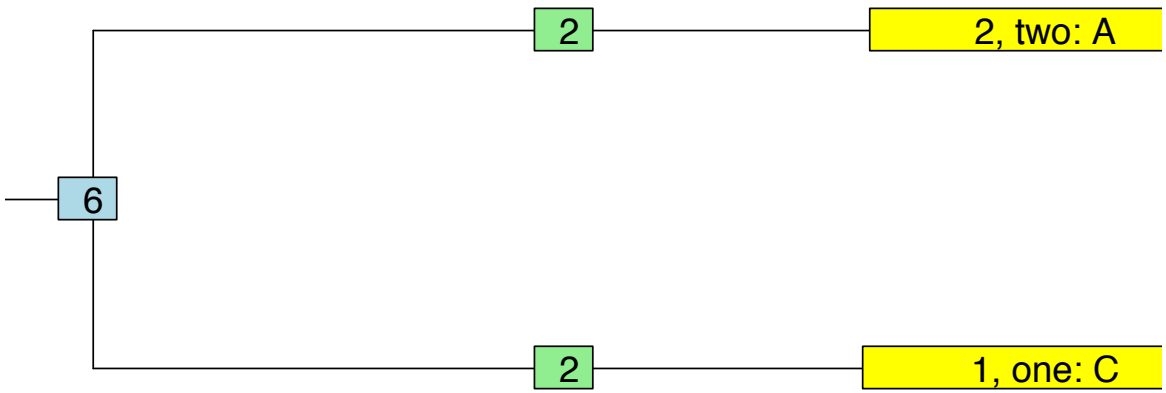
if n *is a tip* **then**
  for $i \leftarrow 1$ **to** N **do**
    $L^{(i)}(n) \leftarrow [0, 0, 0, 0]$;
    $L^{(i)}_{A[n,i]}(n) \leftarrow 1$;
  end
else
  ...;
end

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **3**				
1	1	0	0	0
node **4**				
1	0	0	0	**1**

Felsenstein's pruning algorithm: Internal nodes

if n *is a tip* **then**
  ...;
else
  for $i \leftarrow 1$ **to** N **do**
    for X *in* {T, C, A, G} **do**
      $L1 \leftarrow 0$;
      $L2 \leftarrow 0$;
      for Y *in* {T, C, A, G} **do**
        $L1 \leftarrow L1 + P_{XY}(|n, child1|) \cdot L_Y^{(i)}(child1)$;
        $L2 \leftarrow L2 + P_{XY}(|n, child2|) \cdot L_Y^{(i)}(child2)$;
      end
      $L_X^{(i)}(node) = L1 \cdot L2$;
    end
  end
end

Internal node 6



	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1				

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1				

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1				

$$L_T^{(1)}(6) = \left(\sum_{X \in \{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(1) \right) \times \left(\sum_{X \in \{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(2) \right)$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1				

$$
\begin{aligned}
L_T^{(1)}(6) &= \left(\sum_{X \in \{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(1) \right) \times \left(\sum_{X \in \{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(2) \right) \\
&= (P_{TT}(2) \times 0 + P_{TC}(2) \times 1 + P_{TA}(2) \times 0 + P_{TG}(2) \times 0) \\
&\quad \times (P_{TT}(2) \times 0 + P_{TC}(2) \times 0 + P_{TA}(2) \times 1 + P_{TG}(2) \times 0)
\end{aligned}
$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1				

$$
\begin{aligned}
L_T^{(1)}(6) &= \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(1)\right) \times \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(2)\right) \\
&= (P_{TT}(2)\times 0 + P_{TC}(2)\times 1 + P_{TA}(2)\times 0 + P_{TG}(2)\times 0) \\
&\quad \times (P_{TT}(2)\times 0 + P_{TC}(2)\times 0 + P_{TA}(2)\times 1 + P_{TG}(2)\times 0) \\
&= (0.76\times 0 + 0.24\times 1 + 0.00089\times 0 + 0.00051\times 0) \\
&\quad \times (0.76\times 0 + 0.24\times 0 + 0.00089\times 1 + 0.00051\times 0)
\end{aligned}
$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1				

$$
\begin{aligned}
L_T^{(1)}(6) &= \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(2)L_X^{(1)}(1)\right) \times \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(2)L_X^{(1)}(2)\right) \\
&= (P_{TT}(2)\times 0 + P_{TC}(2)\times 1 + P_{TA}(2)\times 0 + P_{TG}(2)\times 0) \\
&\quad \times (P_{TT}(2)\times 0 + P_{TC}(2)\times 0 + P_{TA}(2)\times 1 + P_{TG}(2)\times 0) \\
&= (0.76\times 0 + 0.24\times 1 + 0.00089\times 0 + 0.00051\times 0) \\
&\quad \times (0.76\times 0 + 0.24\times 0 + 0.00089\times 1 + 0.00051\times 0) \\
&= 0.24\times 0.00089 \approx 0.00021
\end{aligned}
$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1	0.00021			

$$
\begin{aligned}
L_T^{(1)}(6) &= \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(2)L_X^{(1)}(1)\right) \times \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(2)L_X^{(1)}(2)\right) \\
&= (P_{TT}(2)\times 0 + P_{TC}(2)\times 1 + P_{TA}(2)\times 0 + P_{TG}(2)\times 0) \\
&\quad \times (P_{TT}(2)\times 0 + P_{TC}(2)\times 0 + P_{TA}(2)\times 1 + P_{TG}(2)\times 0) \\
&= (0.76\times 0 + 0.24\times 1 + 0.00089\times 0 + 0.00051\times 0) \\
&\quad \times (0.76\times 0 + 0.24\times 0 + 0.00089\times 1 + 0.00051\times 0) \\
&= 0.24\times 0.00089 \approx 0.00021
\end{aligned}
$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1	0.00021			

$$L_C^{(1)}(6) = \left(\sum_{X \in \{T,C,A,G\}} P_{CX}(2) L_X^{(1)}(1) \right) \times \left(\sum_{X \in \{T,C,A,G\}} P_{CX}(2) L_X^{(1)}(2) \right)$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1	0.00021			

$$\mathrm{L}_{\mathrm{C}}^{(1)}(6) = \left(\sum_{X \in \{\mathrm{T,C,A,G}\}} \mathrm{P}_{\mathrm{C}X}(2)\mathrm{L}_{X}^{(1)}(1)\right) \times \left(\sum_{X \in \{\mathrm{T,C,A,G}\}} \mathrm{P}_{\mathrm{C}X}(2)\mathrm{L}_{X}^{(1)}(2)\right)$$
$$= (\mathrm{P}_{\mathrm{CC}}(2) \times 1) \times (\mathrm{P}_{\mathrm{CA}}(2) \times 1) = 0.8 \times 0.00089 \approx 0.00071$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1	0.00021	0.00071		

$$\mathrm{L}_{\mathrm{C}}^{(1)}(6) = \left(\sum_{\mathrm{X} \in \{\mathrm{T},\mathrm{C},\mathrm{A},\mathrm{G}\}} \mathrm{P}_{\mathrm{C}\mathrm{X}}(2) \mathrm{L}_{\mathrm{X}}^{(1)}(1) \right) \times \left(\sum_{\mathrm{X} \in \{\mathrm{T},\mathrm{C},\mathrm{A},\mathrm{G}\}} \mathrm{P}_{\mathrm{C}\mathrm{X}}(2) \mathrm{L}_{\mathrm{X}}^{(1)}(2) \right)$$
$$= (\mathrm{P}_{\mathrm{C}\mathrm{C}}(2) \times 1) \times (\mathrm{P}_{\mathrm{C}\mathrm{A}}(2) \times 1) = 0.8 \times 0.00089 \approx 0.00071$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1	0.00021	0.00071		

$$\mathrm{L}_{\mathrm{A}}^{(1)}(6) = \left(\sum_{X \in \{\mathrm{T,C,A,G}\}} \mathrm{P}_{\mathrm{A}X}(2) \mathrm{L}_{X}^{(1)}(1) \right) \times \left(\sum_{X \in \{\mathrm{T,C,A,G}\}} \mathrm{P}_{\mathrm{A}X}(2) \mathrm{L}_{X}^{(1)}(2) \right)$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1	0.00021	0.00071		

$$\mathrm{L}_{\mathrm{A}}^{(1)}(6) = \left(\sum_{X \in \{\mathrm{T,C,A,G}\}} \mathrm{P}_{\mathrm{A}X}(2) \mathrm{L}_{X}^{(1)}(1) \right) \times \left(\sum_{X \in \{\mathrm{T,C,A,G}\}} \mathrm{P}_{\mathrm{A}X}(2) \mathrm{L}_{X}^{(1)}(2) \right)$$
$$= (\mathrm{P}_{\mathrm{AC}}(2) \times 1) \times (\mathrm{P}_{\mathrm{AA}}(2) \times 1) = 0.0007 \times 0.9 \approx 0.00063$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1	0.00021	0.00071	0.00063	

$$\mathrm{L}_{\mathrm{A}}^{(1)}(6) = \left(\sum_{X\in\{\mathrm{T,C,A,G}\}} \mathrm{P}_{\mathrm{A}X}(2)\mathrm{L}_{X}^{(1)}(1)\right) \times \left(\sum_{X\in\{\mathrm{T,C,A,G}\}} \mathrm{P}_{\mathrm{A}X}(2)\mathrm{L}_{X}^{(1)}(2)\right)$$
$$= (\mathrm{P}_{\mathrm{AC}}(2)\times 1)\times(\mathrm{P}_{\mathrm{AA}}(2)\times 1) = 0.0007\times 0.9 \approx 0.00063$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1	0.00021	0.00071	0.00063	

$$\mathrm{L}_{\mathrm{G}}^{(1)}(6) = \left(\sum_{\mathrm{X} \in \{\mathrm{T},\mathrm{C},\mathrm{A},\mathrm{G}\}} \mathrm{P}_{\mathrm{G}\mathrm{X}}(2) \mathrm{L}_{\mathrm{X}}^{(1)}(1) \right) \times \left(\sum_{\mathrm{X} \in \{\mathrm{T},\mathrm{C},\mathrm{A},\mathrm{G}\}} \mathrm{P}_{\mathrm{G}\mathrm{X}}(2) \mathrm{L}_{\mathrm{X}}^{(1)}(2) \right)$$

Internal node 6

	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1	0.00021	0.00071	0.00063	

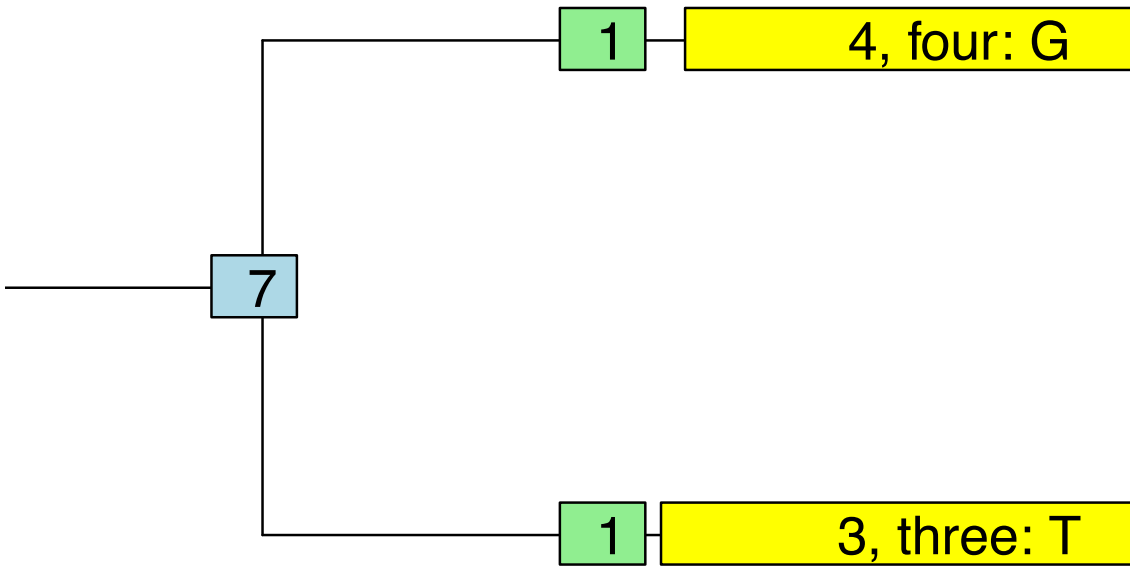
$$\mathrm{L}_{\mathrm{G}}^{(1)}(6) = \left(\sum_{X\in\{\mathrm{T,C,A,G}\}} \mathrm{P}_{\mathrm{G}X}(2)\mathrm{L}_{X}^{(1)}(1)\right) \times \left(\sum_{X\in\{\mathrm{T,C,A,G}\}} \mathrm{P}_{\mathrm{G}X}(2)\mathrm{L}_{X}^{(1)}(2)\right)$$
$$= (\mathrm{P}_{\mathrm{GC}}(2)\times 1)\times(\mathrm{P}_{\mathrm{GA}}(2)\times 1) = 0.0007\times 0.17 \approx 0.00012$$

Internal node 6

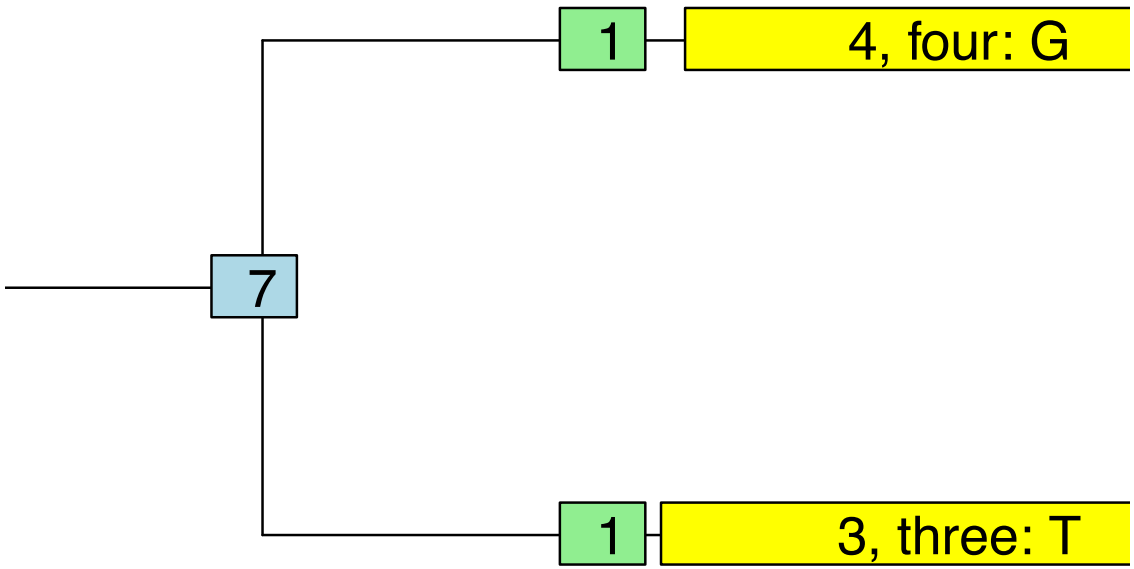
	T	C	A	G
node **1**				
1	0	1	0	0
node **2**				
1	0	0	1	0
node **6**				
1	0.00021	0.00071	0.00063	0.00012

$$L_G^{(1)}(6) = \left(\sum_{X \in \{T,C,A,G\}} P_{GX}(2) L_X^{(1)}(1) \right) \times \left(\sum_{X \in \{T,C,A,G\}} P_{GX}(2) L_X^{(1)}(2) \right)$$
$$= (P_{GC}(2) \times 1) \times (P_{GA}(2) \times 1) = 0.0007 \times 0.17 \approx 0.00012$$

Internal node 7



Internal node 7



	T	C	A	G
node **3**				
1	1	0	0	0
node **4**				
1	0	0	0	1
node **7**				
1	0.00022	0.000029	0.000015	0.00027

Internal node 5

	T	C	A	G
node **6**				
1	0.00021	0.00071	0.00063	0.00012
node **7**				
1	0.00022	0.000029	0.000015	0.00027
node **5**				
1				

Internal node 5

	T	C	A	G
node **6**				
1	0.00021	0.00071	0.00063	0.00012
node **7**				
1	0.00022	0.000029	0.000015	0.00027
node **5**				
1				

$$L_T^{(1)}(5) = \left(\sum_{X \in \{T,C,A,G\}} P_{TX}(1) L_X^{(1)}(6) \right) \times \left(\sum_{X \in \{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(7) \right)$$

Internal node 5

	T	C	A	G
node **6**				
1	0.00021	0.00071	0.00063	0.00012
node **7**				
1	0.00022	0.000029	0.000015	0.00027
node **5**				
1				

$$L_T^{(1)}(5) = \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(1) L_X^{(1)}(6)\right) \times \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(7)\right)$$
$$= (0.861 \times 0.00021 + 0.13 \times 0.00071 + 0.00045 \times 0.00063 + 0.00026 \times 0.00012)$$
$$\times (0.76 \times 0.00022 + 0.24 \times 0.000029 + 0.00089 \times 0.000015 + 0.00051 \times 0.00027)$$

Internal node 5

	T	C	A	G
node **6**				
1	0.00021	0.00071	0.00063	0.00012
node **7**				
1	0.00022	0.000029	0.000015	0.00027
node **5**				
1				

$$\begin{aligned} L_T^{(1)}(5) &= \left(\sum_{X \in \{T,C,A,G\}} P_{TX}(1) L_X^{(1)}(6) \right) \times \left(\sum_{X \in \{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(7) \right) \\ &= (0.861 \times 0.00021 + 0.13 \times 0.00071 + 0.00045 \times 0.00063 + 0.00026 \times 0.00012) \\ &\quad \times (0.76 \times 0.00022 + 0.24 \times 0.000029 + 0.00089 \times 0.000015 + 0.00051 \times 0.00027) \\ &\approx 4.89e{-}08 \end{aligned}$$

Internal node 5

	T	C	A	G
node **6**				
1	0.00021	0.00071	0.00063	0.00012
node **7**				
1	0.00022	0.000029	0.000015	0.00027
node **5**				
1	4.89e-08			

$$\begin{aligned} L_T^{(1)}(5) &= \left(\sum_{X \in \{T,C,A,G\}} P_{TX}(1) L_X^{(1)}(6) \right) \times \left(\sum_{X \in \{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(7) \right) \\ &= (0.861 \times 0.00021 + 0.13 \times 0.00071 + 0.00045 \times 0.00063 + 0.00026 \times 0.00012) \\ &\quad \times (0.76 \times 0.00022 + 0.24 \times 0.000029 + 0.00089 \times 0.000015 + 0.00051 \times 0.00027) \\ &\approx 4.89e{-}08 \end{aligned}$$

Internal node 5

	T	C	A	G
node **6**				
1	0.00021	0.00071	0.00063	0.00012
node **7**				
1	0.00022	0.000029	0.000015	0.00027
node **5**				
1	4.89e-08	4.43e-08		

$$\begin{aligned} L_T^{(1)}(5) &= \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(1) L_X^{(1)}(6)\right) \times \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(7)\right) \\ &= (0.861 \times 0.00021 + 0.13 \times 0.00071 + 0.00045 \times 0.00063 + 0.00026 \times 0.00012) \\ &\quad \times (0.76 \times 0.00022 + 0.24 \times 0.000029 + 0.00089 \times 0.000015 + 0.00051 \times 0.00027) \\ &\approx 4.89e{-}08 \end{aligned}$$

Internal node 5

	T	C	A	G
node **6**				
1	0.00021	0.00071	0.00063	0.00012
node **7**				
1	0.00022	0.000029	0.000015	0.00027
node **5**				
1	4.89e-08	4.43e-08	2.42e-08	

$$L_T^{(1)}(5) = \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(1)L_X^{(1)}(6)\right) \times \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(2)L_X^{(1)}(7)\right)$$
$$= (0.861 \times 0.00021 + 0.13 \times 0.00071 + 0.00045 \times 0.00063 + 0.00026 \times 0.00012)$$
$$\times (0.76 \times 0.00022 + 0.24 \times 0.000029 + 0.00089 \times 0.000015 + 0.00051 \times 0.00027)$$
$$\approx 4.89e{-}08$$

Internal node 5

	T	C	A	G
node **6**				
1	0.00021	0.00071	0.00063	0.00012
node **7**				
1	0.00022	0.000029	0.000015	0.00027
node **5**				
1	4.89e-08	4.43e-08	2.42e-08	3.72e-08

$$
\begin{aligned}
L_T^{(1)}(5) &= \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(1) L_X^{(1)}(6)\right) \times \left(\sum_{X\in\{T,C,A,G\}} P_{TX}(2) L_X^{(1)}(7)\right) \\
&= (0.861 \times 0.00021 + 0.13 \times 0.00071 + 0.00045 \times 0.00063 + 0.00026 \times 0.00012) \\
&\quad \times (0.76 \times 0.00022 + 0.24 \times 0.000029 + 0.00089 \times 0.000015 + 0.00051 \times 0.00027) \\
&\approx 4.89e{-}08
\end{aligned}
$$

From conditional likelihood to likelihood

	T	C	A	G
node **5**				
1	4.89e-08	4.43e-08	2.42e-08	3.72e-08

$$\pi = (0.22, 0.26, 0.33, 0.19)$$

From conditional likelihood to likelihood

	T	C	A	G
node **5**				
1	4.89e-08	4.43e-08	2.42e-08	3.72e-08

$$\pi = (0.22, 0.26, 0.33, 0.19)$$

$$\mathrm{logL} = \sum_{i=1}^{N} \log \left(\sum_{X \in \{T,C,A,G\}} \pi_X L_X^{(i)}(5) \right)$$

From conditional likelihood to likelihood

	T	C	A	G
node **5**				
1	4.89e-08	4.43e-08	2.42e-08	3.72e-08

$$\pi = (0.22, 0.26, 0.33, 0.19)$$

$$\begin{aligned} \mathrm{logL} &= \sum_{i=1}^{N} \log \left(\sum_{X \in \{T,C,A,G\}} \pi_X L_X^{(i)}(5) \right) \\ &= \log \left(\sum_{X \in \{T,C,A,G\}} \pi_X L_X^{(1)}(5) \right) \end{aligned}$$

From conditional likelihood to likelihood

	T	C	A	G
node **5**				
1	4.89e-08	4.43e-08	2.42e-08	3.72e-08

$$\pi = (0.22, 0.26, 0.33, 0.19)$$

$$\begin{aligned}
\mathrm{logL} &= \sum_{i=1}^{N} \log \left(\sum_{X \in \{T,C,A,G\}} \pi_X L_X^{(i)}(5) \right) \\
&= \log \left(\sum_{X \in \{T,C,A,G\}} \pi_X L_X^{(1)}(5) \right) \\
&= \log \left(\pi_T L_T^{(1)}(5) + \pi_C L_C^{(1)}(5) + \pi_A L_A^{(1)}(5) + \pi_G L_G^{(1)}(5) \right)
\end{aligned}$$

From conditional likelihood to likelihood

	T	C	A	G
node **5**				
1	4.89e-08	4.43e-08	2.42e-08	3.72e-08

$$\pi = (0.22, 0.26, 0.33, 0.19)$$

$$\begin{aligned}
\log L &= \sum_{i=1}^{N} \log \left(\sum_{X \in \{T,C,A,G\}} \pi_X L_X^{(i)}(5) \right) \\
&= \log \left(\sum_{X \in \{T,C,A,G\}} \pi_X L_X^{(1)}(5) \right) \\
&= \log \left(\pi_T L_T^{(1)}(5) + \pi_C L_C^{(1)}(5) + \pi_A L_A^{(1)}(5) + \pi_G L_G^{(1)}(5) \right) \\
&= \log(0.22 \times 4.89e{-}08 + 0.26 \times 4.43e{-}08 \\
&\quad + 0.33 \times 2.42e{-}08 + 0.19 \times 3.72e{-}08)
\end{aligned}$$

From conditional likelihood to likelihood

	T	C	A	G
node **5**				
1	4.89e-08	4.43e-08	2.42e-08	3.72e-08

$$\pi = (0.22, 0.26, 0.33, 0.19)$$

$$\begin{aligned}
\log L &= \sum_{i=1}^{N} \log \left(\sum_{X \in \{T,C,A,G\}} \pi_X L_X^{(i)}(5) \right) \\
&= \log \left(\sum_{X \in \{T,C,A,G\}} \pi_X L_X^{(1)}(5) \right) \\
&= \log \left(\pi_T L_T^{(1)}(5) + \pi_C L_C^{(1)}(5) + \pi_A L_A^{(1)}(5) + \pi_G L_G^{(1)}(5) \right) \\
&= \log(0.22 \times 4.89e{-}08 + 0.26 \times 4.43e{-}08 \\
&\quad + 0.33 \times 2.42e{-}08 + 0.19 \times 3.72e{-}08) \\
&= \log(3.73e{-}08) \approx -17.1
\end{aligned}$$

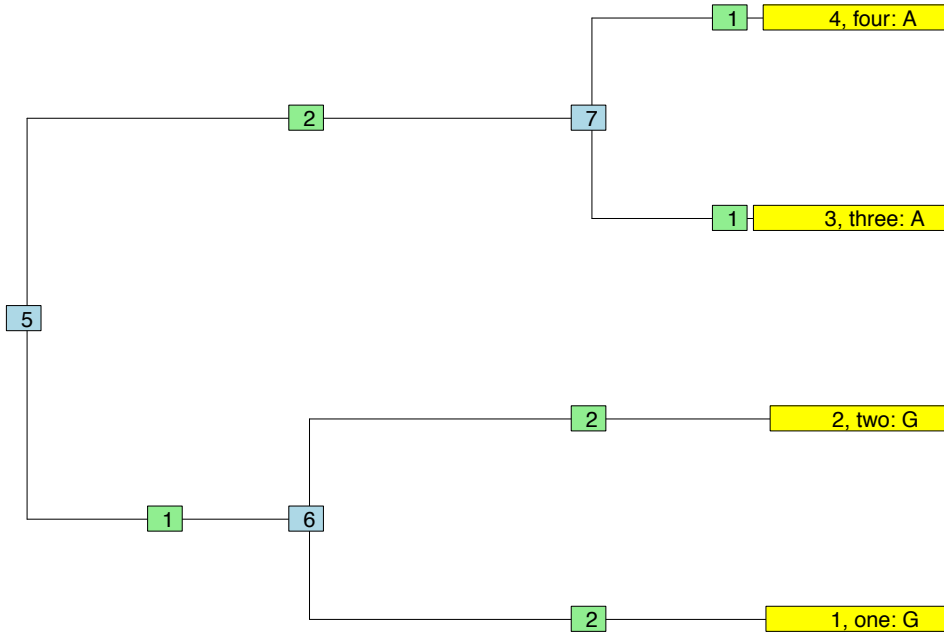
Full tree log-likelihood computation

Data: Sequence alignment A, tree τ
Result: $\log L$
$\log L \leftarrow 0$;
for $i \leftarrow 1$ **to** N **do**
  $L \leftarrow 0$;
  for X *in* $\{T, C, A, G\}$ **do**
    $L \leftarrow L + \pi_X L_X^{(i)}(\text{root})$;
  end
  $\log L \leftarrow \log L + \log(L)$;
end
return($\log L$);

Tree node/nucleotide likelihood computation

Data: Node n, sequence alignment A, tree τ, transition probability matrices $P_{TN93}(t)$ for each branch length t in τ
Result: $L_i(n)$
if n *is a tip* **then**
  for $i \leftarrow 1$ **to** N **do**
    $L^{(i)}(n) \leftarrow [0, 0, 0, 0]$;
    $L^{(i)}_{A[n,i]}(n) \leftarrow 1$;
  end
else
  for $i \leftarrow 1$ **to** N **do**
    for X *in* $\{T, C, A, G\}$ **do**
      $L1 \leftarrow 0$;
      $L2 \leftarrow 0$;
      for Y *in* $\{T, C, A, G\}$ **do**
        $L1 \leftarrow L1 + P_{XY}(|n, child1|) \cdot L^{(i)}_Y(child1)$;
        $L2 \leftarrow L2 + P_{XY}(|n, child2|) \cdot L^{(i)}_Y(child2)$;
      end
      $L^{(i)}_X(node) = L1 \times L2$;
    end
  end

Pen and paper exercise



$\pi = (0.22, 0.26, 0.33, 0.19)$

P(1):

	T	C	A	G
T	0.86	0.13	0.00045	0.00026
C	0.11	0.89	0.00045	0.00026
A	0.0003	0.00035	0.95	0.052
G	0.0003	0.00035	0.09	0.91

P(2):

	T	C	A	G
T	0.76	0.24	0.00089	0.00051
C	0.2	0.8	0.00089	0.00051
A	0.00059	0.0007	0.9	0.096
G	0.00059	0.0007	0.17	0.83

Pen and paper solution

	T	C	A	G
node 1				
1	0	0	0	1
node 2				
1	0	0	0	1
node 6				
1	2.601e-07	2.601e-07	9.216e-03	0.6889
node 3				
1	0	0	1	0
node 4				
1	0	0	1	0
node 7				
1	2.025e-07	2.025e-07	0.9025	8.1e-03
node 5				
1	1.482021e-07	1.448554e-07	0.036	0.099

$$\begin{aligned}\mathbf{LogL} &= \log(\pi_T \times L_1(5|T) + \pi_C \times L_1(5|C) + \pi_A \times L_1(5|A) + \pi_G \times L_1(5|G)) \\ &= \log(0.22 \times 1.482021e{-}07 + 0.26 \times 1.448554e{-}07 \\ &+ 0.33 \times 0.03624314450 + 0.19 \times 0.1005294544) \\ &= \log(0.0310609) \approx -3.471805\end{aligned}$$