Maximum Likelihood Analysis of Phylogenetic Trees

Benny Chor

School of Computer Science
Tel-Aviv University

Phylogenetic Reconstruction Methods

- Input: A set of n aligned sequences (genes, proteins) from n species,

Phylogenetic Reconstruction Methods

- Input: A set of n aligned sequences (genes, proteins) from n species,
- Goal: Reconstruct the tree which best explains the evolutionary history of this gene/protein.

Phylogenetic Reconstruction Methods

- Input: A set of n aligned sequences (genes, proteins) from n species,
- Goal: Reconstruct the tree which best explains the evolutionary history of this gene/protein.
- Tree reconstruction is still a challenge today.

Phylogenetic Reconstruction Methods

- Input: A set of n aligned sequences (genes, proteins) from n species,
- Goal: Reconstruct the tree which best explains the evolutionary history of this gene/protein.
- Tree reconstruction is still a challenge today.
- Many concrete questions are still unresolved (*e.g.* mammalian evolutionary tree).

Phylogenetic Reconstruction Methods

- Input: A set of n aligned sequences (genes, proteins) from n species,
- Goal: Reconstruct the tree which best explains the evolutionary history of this gene/protein.
- Tree reconstruction is still a challenge today.
- Many concrete questions are still unresolved (*e.g.* mammalian evolutionary tree).
- Most realistic formulations of the problem, which take errors into account, give rise to hard computational problems.

Popular Reconstruction Methods

- Distance based methods:

Popular Reconstruction Methods

- Distance based methods:
  - UPGMA

Popular Reconstruction Methods

- Distance based methods:
  - UPGMA
  - Neighbor Joining.

Popular Reconstruction Methods

- Distance based methods:
  - UPGMA
  - Neighbor Joining.
  - Buneman trees.

Popular Reconstruction Methods

- Distance based methods:
  - UPGMA
  - Neighbor Joining.
  - Buneman trees.
- Character Based Methods:

Popular Reconstruction Methods

- Distance based methods:
  - UPGMA
  - Neighbor Joining.
  - Buneman trees.
- Character Based Methods:
  - Maximum Parsimony.

Popular Reconstruction Methods

- Distance based methods:
  - UPGMA
  - Neighbor Joining.
  - Buneman trees.
- Character Based Methods:
  - Maximum Parsimony.
  - Maximum Likelihood.

Popular Reconstruction Methods

- Distance based methods:
  - UPGMA
  - Neighbor Joining.
  - Buneman trees.
- Character Based Methods:
  - Maximum Parsimony.
  - Maximum Likelihood.
- Additional Methods:

Popular Reconstruction Methods

- Distance based methods:
    - UPGMA
    - Neighbor Joining.
    - Buneman trees.
- Character Based Methods:
    - Maximum Parsimony.
    - Maximum Likelihood.
- Additional Methods:
    - Quartets Based.

Popular Reconstruction Methods

- Distance based methods:
  - UPGMA
  - Neighbor Joining.
  - Buneman trees.
- Character Based Methods:
  - Maximum Parsimony.
  - Maximum Likelihood.
- Additional Methods:
  - Quartets Based.
  - Disc Covering.

Popular Reconstruction Methods

- Distance based methods:
  - UPGMA
  - Neighbor Joining.
  - Buneman trees.
- Character Based Methods:
  - Maximum Parsimony.
  - Maximum Likelihood.
- Additional Methods:
  - Quartets Based.
  - Disc Covering.

Popular Reconstruction Methods

- Distance based methods:
  - UPGMA
  - Neighbor Joining.
  - Buneman trees.
- Character Based Methods:
  - Maximum Parsimony.
  - Maximum Likelihood.
- Additional Methods:
  - Quartets Based.
  - Disc Covering.

Talk Outline

- Maximum likelihood (ML).

Talk Outline

- Maximum likelihood (ML).
- The likelihood surface.

Talk Outline

- Maximum likelihood (ML).
- The likelihood surface.
- Existence of multiple maxima.

Talk Outline

- Maximum likelihood (ML).
- The likelihood surface.
- Existence of multiple maxima.
- Computation complexity: Maximum likelihood vs. maximum parsimony (MP).

Talk Outline

- Maximum likelihood (ML).
- The likelihood surface.
- Existence of multiple maxima.
- Computation complexity: Maximum likelihood vs. maximum parsimony (MP).
- Ancestral maximum likelihood (AML) and its computational complexity.

Maximum Likelihood

- Input: A set of n observed sequences and an underlying substitution model.

Maximum Likelihood

- Input: A set of n observed sequences and an underlying substitution model.
- Desired Output: The weighted tree T that maximizes the likelihood of the data.

Maximum Likelihood

- Input: A set of n observed sequences and an underlying substitution model.
- Desired Output: The weighted tree T that maximizes the likelihood of the data.
- Likelihood of a data: The conditional probability of producing the data, given the model parameters.

Maximum Likelihood

- Input: A set of n observed sequences and an underlying substitution model.
- Desired Output: The weighted tree T that maximizes the likelihood of the data.
- Likelihood of a data: The conditional probability of producing the data, given the model parameters.
- Likelihood is a common optimization criteria in numerous settings, including phylogenetic (Felsenstein 1981).

Neyman 2–State Substitution Model

species	observed data
`1`	`XXXXXXXYYY XXY XY YX XY X`
`2`	`XXXXXXXYYY YYX YX YX YX X`
`3`	`XXXXXXXYYY YYX XY XY XY X`
`4`	`XXXXXXXYYY YYX XY XY YX Y`

- Just two characters states, X and Y.

Neyman 2–State Substitution Model

species	observed data
1	`XXXXXXXYYY XXY XY YX XY X`
2	`XXXXXXXYYY YYX YX YX YX X`
3	`XXXXXXXYYY YYX XY XY XY X`
4	`XXXXXXXYYY YYX XY XY YX Y`

- Just two characters states, X and Y.
- Transitions between states are symmetric.

Neyman 2–State Substitution Model

species	observed data
1	`XXXXXXXYYY XXY XY YX XY X`
2	`XXXXXXXYYY YYX YX YX YX X`
3	`XXXXXXXYYY YYX XY XY XY X`
4	`XXXXXXXYYY YYX XY XY YX Y`

- Just two characters states, X and Y.
- Transitions between states are symmetric.
- Equal rates across sites.

Neyman 2–State Substitution Model

species	observed data
1	`XXXXXXXYYY XXY XY YX XY X`
2	`XXXXXXXYYY YYX YX YX YX X`
3	`XXXXXXXYYY YYX XY XY XY X`
4	`XXXXXXXYYY YYX XY XY YX Y`

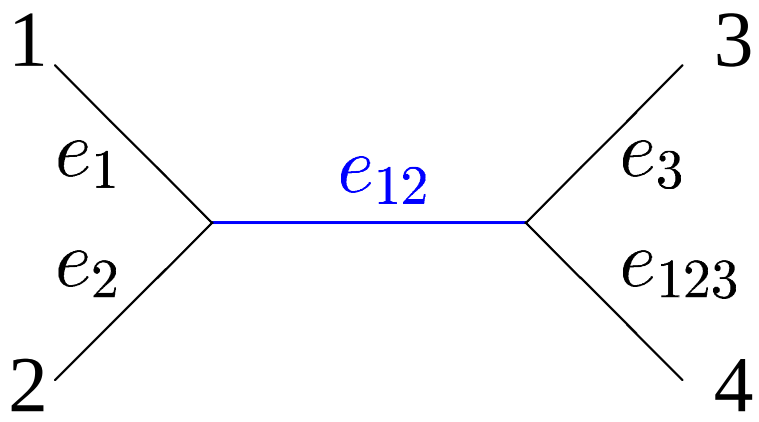
- Just two characters states, X and Y.
- Transitions between states are symmetric.
- Equal rates across sites.
- Every column induces a pattern.

Neyman 2–State Substitution Model

species	observed data
1	XXXXXXXYYY XXY XY YX XY X
2	XXXXXXXYYY YYX YX YX YX X
3	XXXXXXXYYY YYX XY XY XY X
4	XXXXXXXYYY YYX XY XY YX Y

- Just two characters states, X and Y.
- Transitions between states are symmetric.
- Equal rates across sites.
- Every column induces a pattern.
- Remark: A simple model, yet very powerful.

Neyman 2–State Substitution Model



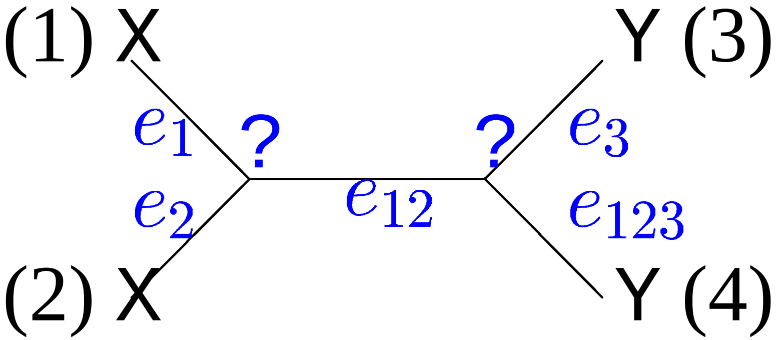
For each edge e of a tree T, the edge weight p_e represents the probability of having different states at the two ends of e.

A **Very** Simple Example

Four species ($n = 4$), just one site ($c = 1$)

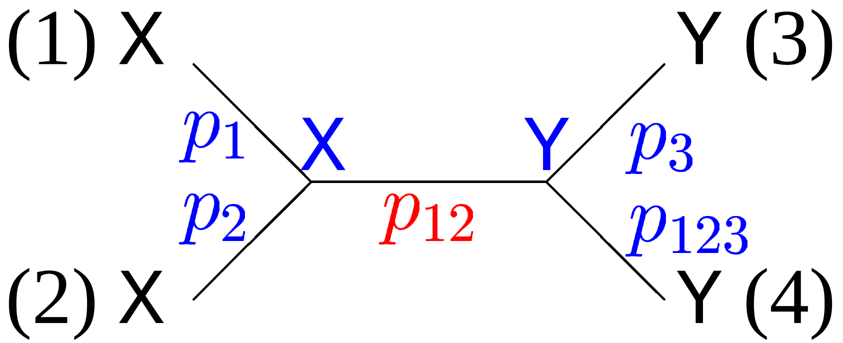
species	observed data
1	X
2	X
3	Y
4	Y

Analyze the *natural* tree $(12)(34)$



Computing the Likelihood

Each unknown state (?) can assume one of two possibilities, X or Y. For example, the assignment



contributes $(1-p_1)\cdot(1-p_2)\cdot p_{12}\cdot(1-p_3)\cdot(1-p_{123})$.

The likelihood is the sum of this

$+$ three similar expressions. . .

Computing the Likelihood

- Last expression has the form $\sum_{\text{internal assignments}} \prod_{\text{edges}} m_{\text{e,a,t}}$ where each $m_{\text{e,a,t}}$ is either p_{e} or $1 - p_{\text{e}}$, depending on the assignment a, and input pattern t at two ends of the edge.

Computing the Likelihood

- Last expression has the form $\sum_{\text{internal assignments}} \prod_{\text{edges}} m_{\text{e,a,t}}$ where each $m_{\text{e,a,t}}$ is either p_{e} or $1 - p_{\text{e}}$, depending on the assignment a, and input pattern t at two ends of the edge.
- When the data has more then one column, we multiply the expressions to get the likelihood of the data, given the model parameters, $L(\text{data}|\text{tree \& edge weights})$:

$$\prod_{\text{columns}} \sum_{\text{internal assignments}} \prod_{\text{edges}} m_{\text{e,a,t}} \, .$$

Three Likelihood Versions

- Big Likelihood: Given the sequence data, find a tree and edge weights that maximize $L(\text{data}|\text{tree \& edge weights})$.

Three Likelihood Versions

- Big Likelihood: Given the sequence data, find a tree and edge weights that maximize $L(\text{data}|\text{tree \& edge weights})$.
- Small Likelihood: Given observed data & a tree, but not the edge weights, find the edge weights that maximize the likelihood.

Three Likelihood Versions

- Big Likelihood: Given the sequence data, find a tree and edge weights that maximize $L(\text{data}|\text{tree \& edge weights})$.
- Small Likelihood: Given observed data & a tree, but not the edge weights, find the edge weights that maximize the likelihood.
- Tiny Likelihood: Given observed data & a tree & edge weights, find the likelihood.

Three Likelihood Versions

- Big Likelihood: Given the sequence data, find a tree and edge weights that maximize $L(\text{data}|\text{tree \& edge weights})$.
- Small Likelihood: Given observed data & a tree, but not the edge weights, find the edge weights that maximize the likelihood.
- Tiny Likelihood: Given observed data & a tree & edge weights, find the likelihood.
- Tiny likelihood can be efficiently computed using dynamic programming (Felsenstein, 1981).

Hill Climbing and Small Likelihood

- Typical approach to small likelihood, used in practice:

Hill Climbing and Small Likelihood

- Typical approach to small likelihood, used in practice:
- Start at some initial point with edge weights $\mathbf{p}$.

Hill Climbing and Small Likelihood

- Typical approach to small likelihood, used in practice:
- Start at some initial point with edge weights $\mathbf{p}$.
- Apply hill climbing on the likelihood function to reach a maximum.

The Likelihood Surface

- For hill climbing to be guaranteed to find the maximum, there must be a single *local and global* maximum in the parameter space.

The Likelihood Surface

- For hill climbing to be guaranteed to find the maximum, there must be a single *local and global* maximum in the parameter space.
- Fukami and Tateno (89), Tillier (94): For any tree, the ML point will be unique.

The Likelihood Surface

- For hill climbing to be guaranteed to find the maximum, there must be a single *local and global* maximum in the parameter space.
- Fukami and Tateno (89), Tillier (94): For any tree, the ML point will be unique.
- Steel (94): Proofs are erroneous - A simple but pathological counter example (multiple maxima on the wrong tree).

The Likelihood Surface

- For hill climbing to be guaranteed to find the maximum, there must be a single *local and global* maximum in the parameter space.
- Fukami and Tateno (89), Tillier (94): For any tree, the ML point will be unique.
- Steel (94): Proofs are erroneous - A simple but pathological counter example (multiple maxima on the wrong tree).
- (94–present): Hill climbing techniques still used. Steel's counter example is considered too "biologically unrealistic" to warrant concern.

The Likelihood Surface (cont.)

- Rogers and Swofford (99): Simulation Study

The Likelihood Surface (cont.)

- Rogers and Swofford (99): Simulation Study
    - Data is simulated on a tree.

The Likelihood Surface (cont.)

- Rogers and Swofford (99): Simulation Study
  - Data is simulated on a tree.
  - Multiple optima are rare...

The Likelihood Surface (cont.)

- Rogers and Swofford (99): Simulation Study
  - Data is simulated on a tree.
  - Multiple optima are rare...
  - ...especially on the *correct* tree.

The Likelihood Surface (cont.)

- Rogers and Swofford (99): Simulation Study
  - Data is simulated on a tree.
  - Multiple optima are rare...
  - ...especially on the *correct* tree.
- Goal here: Investigate the problem analytically (joint work with Hendy, Holland, Penny).

Maximizing Likelihood on Trees

Tools used

- Hadamard conjugation (Hendy and Penny 93).

Maximizing Likelihood on Trees

Tools used

- Hadamard conjugation (Hendy and Penny 93).
- Splits and sequence spectra (change of variables)

Maximizing Likelihood on Trees

Tools used

- Hadamard conjugation (Hendy and Penny 93).
- Splits and sequence spectra (change of variables)
- Constrained optimization.

Maximizing Likelihood on Trees

Tools used

- Hadamard conjugation (Hendy and Penny 93).
- Splits and sequence spectra (change of variables)
- Constrained optimization.
- Systems of polynomial equations.

Maximizing Likelihood on Trees

Tools used

- Hadamard conjugation (Hendy and Penny 93).
- Splits and sequence spectra (change of variables)
- Constrained optimization.
- Systems of polynomial equations.
- Analytical solution: very hard in general, even for four taxa.

Maximizing Likelihood on Trees

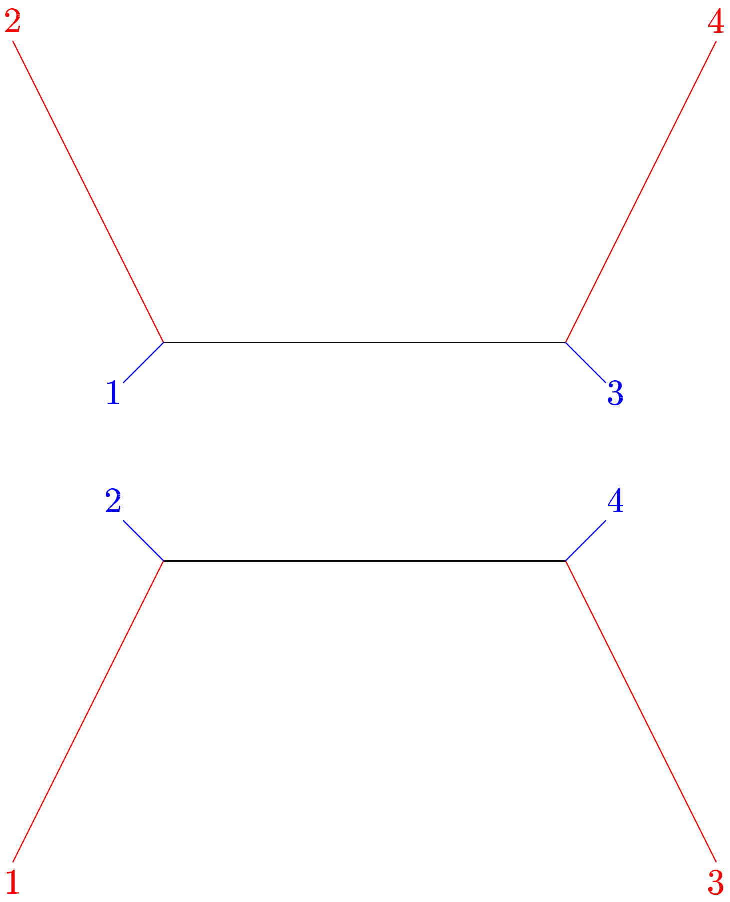
Tools used

- Hadamard conjugation (Hendy and Penny 93).
- Splits and sequence spectra (change of variables)
- Constrained optimization.
- Systems of polynomial equations.
- Analytical solution: very hard in general, even for four taxa.
- Employing computer algebra and algebraic geometry tools.

Example: Conservative Data, Two Very Different ML Trees

species	observed data
1	XXXXXXXYYY XXY XY YX XY X
2	XXXXXXXYYY YYX YX YX YX X
3	XXXXXXXYYY YYX XY XY XY X
4	XXXXXXXYYY YYX XY XY YX Y

Example: Conservative Data, Two Very Different ML Trees



Molecular Clock Model

- Phylogenetic trees under a molecular clock (MC):

3 taxa

one rooted topology

MC-triplet

4 taxa

two rooted topologies

MC-fork

MC-comb

Molecular Clock Model

- Phylogenetic trees under a molecular clock (MC):
  - Rooted trees.

3 taxa

one rooted topology

MC-triplet

4 taxa

two rooted topologies

MC-fork

MC-comb

Molecular Clock Model

- Phylogenetic trees under a molecular clock (MC):
  - Rooted trees.
  - Equal distance from root to all leaves.

3 taxa

one rooted topology

MC-triplet

4 taxa

two rooted topologies

MC-fork

MC-comb

Molecular Clock Model

- Phylogenetic trees under a molecular clock (MC):
  - Rooted trees.
  - Equal distance from root to all leaves.

3 taxa

one rooted topology

MC-triplet

4 taxa

two rooted topologies

MC-fork

MC-comb

Molecular Clock Model

- Phylogenetic trees under a molecular clock (MC):

Molecular Clock Model

- Phylogenetic trees under a molecular clock (MC):
  - Rooted trees.

Molecular Clock Model

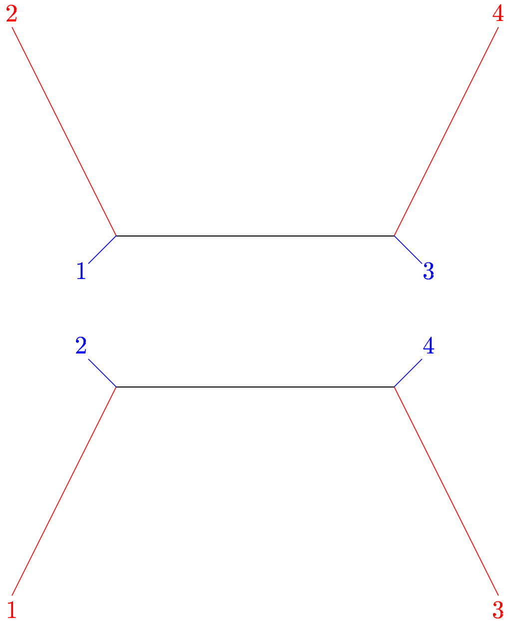
- Phylogenetic trees under a molecular clock (MC):
    - Rooted trees.
    - Equal distance from root to all leaves.

Molecular Clock Model

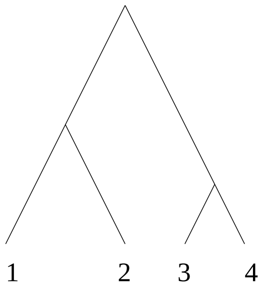
- Phylogenetic trees under a molecular clock (MC):
  - Rooted trees.
  - Equal distance from root to all leaves.
- Negative Examples:

Molecular Clock Model

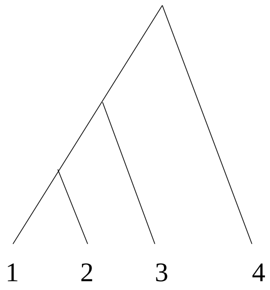
- Phylogenetic trees under a molecular clock (MC):
    - Rooted trees.
    - Equal distance from root to all leaves.
- Negative Examples:



MC Trees, 4 Taxa



MC-fork

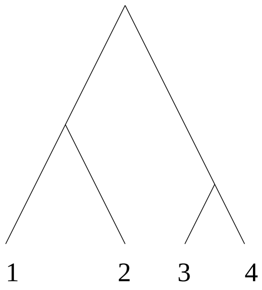


MC-comb

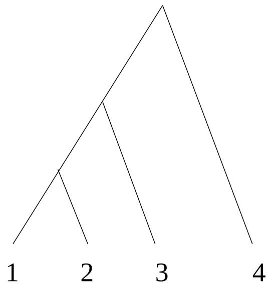
- Fork: Closed form ML solution.

Joint work with Snir and Khetan.

MC Trees, 4 Taxa



MC-fork



MC-comb

- Fork: Closed form ML solution.
- Comb: Analytical ML solution (root of 9-th degree poly).

Joint work with Snir and Khetan.

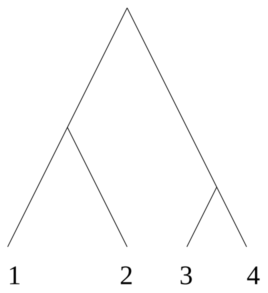
MC Trees, 4 Taxa

MC-fork MC-comb

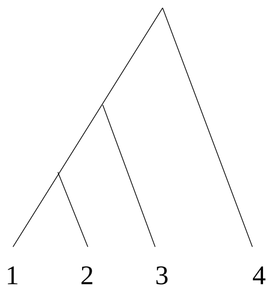
- Fork: Closed form ML solution.
- Comb: Analytical ML solution (root of 9-th degree poly).
- In both cases, ML solution is unique.

Joint work with Snir and Khetan.

MC Trees, 4 Taxa



MC-fork



MC-comb

- Fork: Closed form ML solution.
- Comb: Analytical ML solution (root of 9-th degree poly).
- In both cases, ML solution is unique.
- Attaining solutions requires fairly heavy math and computer algebra tools.

Joint work with Snir and Khetan.

Small Likelihood & Multiple Maxima

- Small Likelihood (reminder): Given observed data & a tree, but not the edge weights, find the edge weights that maximize the likelihood.

Small Likelihood & Multiple Maxima

- Small Likelihood (reminder): Given observed data & a tree, but not the edge weights, find the edge weights that maximize the likelihood.
- Multiple ML points for general case imply small likelihood cannot be solved by hill climbing.

Small Likelihood & Multiple Maxima

- Small Likelihood (reminder): Given observed data & a tree, but not the edge weights, find the edge weights that maximize the likelihood.
- Multiple ML points for general case imply small likelihood cannot be solved by hill climbing.
- Not clear if small likelihood has efficient (worst case) solutions.

Maximum Parsimony (MP)

- Big Parsimony: Given the sequence data, find a tree and assignment of sequences to internal nodes that minimizes the number of changes across all edges.

Maximum Parsimony (MP)

- Big Parsimony: Given the sequence data, find a tree and assignment of sequences to internal nodes that minimizes the number of changes across all edges.
- Small Parsimony: Given the sequence data and a tree, find internal assignment(s) that minimizes total number of changes.

Maximum Parsimony (MP)

- Big Parsimony: Given the sequence data, find a tree and assignment of sequences to internal nodes that minimizes the number of changes across all edges.
- Small Parsimony: Given the sequence data and a tree, find internal assignment(s) that minimizes total number of changes.
- MP considered by practitioners easier than ML. Indeed small parsimony has efficient algorithms (Fitch 1971, Sankoff and Cedergren 1983).

Complexity: MP vs. ML

- Small parsimony is in P.

Complexity: MP vs. ML

- Small parsimony is in P.
- Small likelihood – unknown.

Complexity: MP vs. ML

- Small parsimony is in P.
- Small likelihood – unknown.
- Big parsimony is NP hard (Day, Johnson and Sankoff, 1986).

Complexity: MP vs. ML

- Small parsimony is in P.
- Small likelihood – unknown.
- Big parsimony is NP hard (Day, Johnson and Sankoff, 1986).
- Big likelihood – unknown. Given the importance of ML, it would be nice to know more about its complexity than just “seems harder than MP”.

Ancestral Max. Likelihood (AML)

- A tree reconstruction method that is “in between” ML and MP.

Ancestral Max. Likelihood (AML)

- A tree reconstruction method that is “in between” ML and MP.
- The goal is to simultaneously find edge weights and assignment of sequences to internal nodes so that the likelihood of the data, given the tree parameters, is maximized.

Ancestral Max. Likelihood (AML)

- A tree reconstruction method that is “in between” ML and MP.
- The goal is to simultaneously find edge weights and assignment of sequences to internal nodes so that the likelihood of the data, given the tree parameters, is maximized.
- AML is widely used in evolutionary studies.

Ancestral Max. Likelihood (AML)

- A tree reconstruction method that is “in between” ML and MP.
- The goal is to simultaneously find edge weights and assignment of sequences to internal nodes so that the likelihood of the data, given the tree parameters, is maximized.
- AML is widely used in evolutionary studies.
- Also termed joint reconstruction of ancestral sequences.

Ancestral Max. Likelihood (AML)

- A tree reconstruction method that is “in between” ML and MP.
- The goal is to simultaneously find edge weights and assignment of sequences to internal nodes so that the likelihood of the data, given the tree parameters, is maximized.
- AML is widely used in evolutionary studies.
- Also termed joint reconstruction of ancestral sequences.
- AML computes the likelihood contribution resulting from best assignment to internal nodes, while “regular ML” sums up over all assignments.

Two AML Versions

- Big AML: Given the sequence data, find a tree, assignment to internal nodes, and edge weights that maximize the likelihood of the data.

Two AML Versions

- Big AML: Given the sequence data, find a tree, assignment to internal nodes, and edge weights that maximize the likelihood of the data.
- Small AML: Given observed data, a tree and edge weights, but not the internal assignment, find the assignment that maximize the likelihood.

Two AML Versions

- Big AML: Given the sequence data, find a tree, assignment to internal nodes, and edge weights that maximize the likelihood of the data.
- Small AML: Given observed data, a tree and edge weights, but not the internal assignment, find the assignment that maximize the likelihood.
- PPSG 2000: A poly time, dynamic programming algorithm for small AML.

Two AML Versions

- Big AML: Given the sequence data, find a tree, assignment to internal nodes, and edge weights that maximize the likelihood of the data.
- Small AML: Given observed data, a tree and edge weights, but not the internal assignment, find the assignment that maximize the likelihood.
- PPSG 2000: A poly time, dynamic programming algorithm for small AML.
- ACHLPW 2003: Big AML is NP-hard.

Useful AML Observation

- Given sequence data, a tree, and assignment to internal nodes.

Useful AML Observation

- Given sequence data, a tree, and assignment to internal nodes.
- The edge weights that maximize the likelihood of the data equal d_e/k.

Useful AML Observation

- Given sequence data, a tree, and assignment to internal nodes.
- The edge weights that maximize the likelihood of the data equal d_e/k.
- Where d_e equals the number of changes accross edge e, and k is the common sequence length.

AML, Reformulated

Previous observation implies

- **Input:** A set S of n binary sequences, each of length

AML, Reformulated

Previous observation implies

- **Input:** A set S of n binary sequences, each of length
- **Goal:** Find a tree T with n leaves, an assignment $p : E(T) \rightarrow [0, 1]$ of edge probabilities, and a labelling $\lambda : V(T) \rightarrow \{0, 1\}^k$ of the vertices such that

AML, Reformulated

Previous observation implies

- **Input:** A set S of n binary sequences, each of length
- **Goal:** Find a tree T with n leaves, an assignment $p : E(T) \to [0,1]$ of edge probabilities, and a labelling $\lambda : V(T) \to \{0,1\}^k$ of the vertices such that
  1. The n labels of the leaves are exactly the sequences from S.

AML, Reformulated

Previous observation implies

- **Input:** A set S of n binary sequences, each of length
- **Goal:** Find a tree T with n leaves, an assignment $p : E(T) \to [0, 1]$ of edge probabilities, and a labelling $\lambda : V(T) \to \{0, 1\}^k$ of the vertices such that
  1. The n labels of the leaves are exactly the sequences from S.
  2. the sum of all “edge entropies” $\sum_{e \in E(T)} H\left(d_e/k\right)$ is minimized.

AML vs. MP

Optimization criteria

- **Input:** A set S of n binary sequences, each of length k.

AML vs. MP

Optimization criteria

- **Input:** A set S of n binary sequences, each of length k.
- AML: Minimize the sum of all "edge entropies" $\sum_{e \in E(T)} H\left(d_e/k\right)$.

AML vs. MP

Optimization criteria

- **Input:** A set S of n binary sequences, each of length k.
- AML: Minimize the sum of all "edge entropies" $\sum_{e \in E(T)} H\left(d_e/k\right)$.
- MP: Minimize the sum of all "edge differences" $\sum_{e \in E(T)} d_e/k$.

AML vs. MP

Optimization criteria

- **Input:** A set S of n binary sequences, each of length k.
- AML: Minimize the sum of all “edge entropies” $\sum_{e \in E(T)} H\left(d_e/k\right)$.
- MP: Minimize the sum of all “edge differences” $\sum_{e \in E(T)} d_e/k$.
- Can think of the two problems as attempting to minimize different edge weights (functions of d_e).

NP hardness of AML: Ideas

- MP was shown NP-hard by Day, Johnson, Sankoff using reduction from vertex cover (VC).

NP hardness of AML: Ideas

- MP was shown NP-hard by Day, Johnson, Sankoff using reduction from vertex cover (VC).
- Analogy of AML and MP optimization criteria suggests using similar approach.

NP hardness of AML: Ideas

- MP was shown NP-hard by Day, Johnson, Sankoff using reduction from vertex cover (VC).
- Analogy of AML and MP optimization criteria suggests using similar approach.
- Reduction from VC indeed identical.

NP hardness of AML: Ideas

- MP was shown NP-hard by Day, Johnson, Sankoff using reduction from vertex cover (VC).
- Analogy of AML and MP optimization criteria suggests using similar approach.
- Reduction from VC indeed identical.
- Proof substantially more involved as entropy $H\left(d_e/k\right)$ is not as “well behaved” as plain edge differences d_e/k.

Conclusion and Open Problems

- Analytic solutions to additional ML problems with few taxa may be feasible, and may reveal additional properties of likelihood surface (*e.g.* number of local maxima).

Conclusion and Open Problems

- Analytic solutions to additional ML problems with few taxa may be feasible, and may reveal additional properties of likelihood surface (*e.g.* number of local maxima).
- Multiple ML points for MC trees with more than 4 taxa?

Conclusion and Open Problems

- Analytic solutions to additional ML problems with few taxa may be feasible, and may reveal additional properties of likelihood surface (*e.g.* number of local maxima).
- Multiple ML points for MC trees with more than 4 taxa?
- Hardness proof for big AML as a stepping stone for big ML?

Conclusion and Open Problems

- Analytic solutions to additional ML problems with few taxa may be feasible, and may reveal additional properties of likelihood surface (*e.g.* number of local maxima).
- Multiple ML points for MC trees with more than 4 taxa?
- Hardness proof for big AML as a stepping stone for big ML?
- Is small ML in poly-time?

Conclusion and Open Problems

- Analytic solutions to additional ML problems with few taxa may be feasible, and may reveal additional properties of likelihood surface (*e.g.* number of local maxima).
- Multiple ML points for MC trees with more than 4 taxa?
- Hardness proof for big AML as a stepping stone for big ML?
- Is small ML in poly-time?