

# Phylogenetische Methoden in der historischen Linguistik

*Die IELex-Datenbank  
Maximum Likelihood und  
Bayessche Inferenz*

Gerhard Jäger  
Forum Scientiarum  
16. Dezember 2014

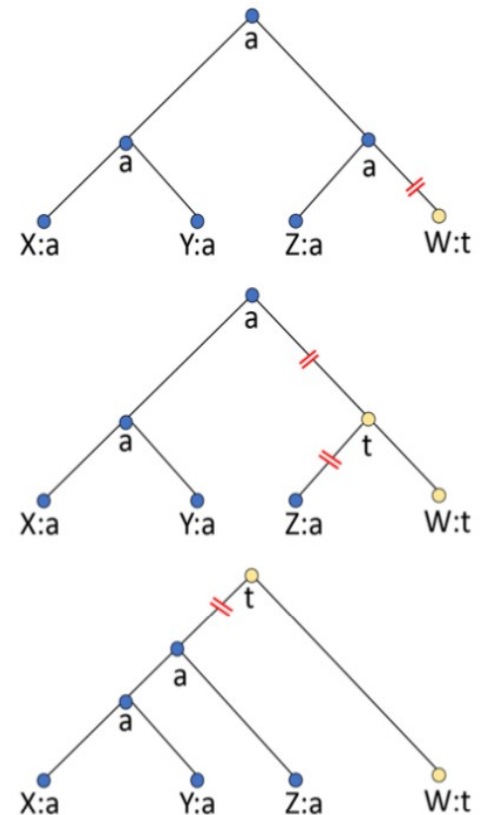
# Maximum Likelihood

- Maximum Parsimony basiert auf der Annahme, dass Mutationen extrem unwahrscheinlich sind
- Wenn Mutationen nicht ganz so selten sind, liefert MP falsche Ergebnisse
- Bessere (und computationell teurere) Alternative:

**Maximum Likelihood**

# Maximum Likelihood

- Angenommen, wir kennen den korrekten Baum  $T$  einschließlich
  - der Kantenlängen
  - der Charakterzustände an jedem Knoten
- Ein **evolutionäres Modell** sagt uns, wie wahrscheinlich die beobachteten Mutationen sind



# Maximum Likelihood

- Mutationswahrscheinlichkeiten sind vergleichbar zu den Mutationsgewichten im Sankoff-Algorithmus

## The Sankoff algorithm

1.  $S_i(j)$  is defined as the number of steps required at or above node  $i$  given that node  $i$  is in state  $j$ .
2. Set these quantities at the tips for the character (they are either 0 or  $\infty$ ).
3. move down the tree doing this at each node:

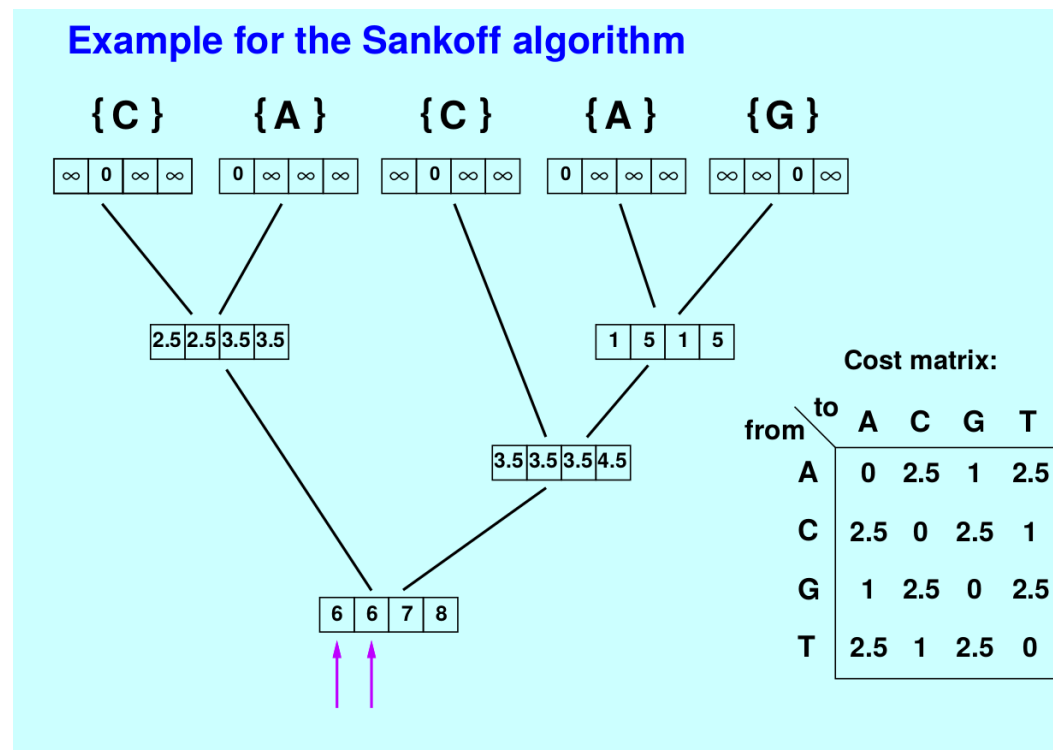
$$S_a(i) = \min_j [c_{ij} + S_l(j)] + \min_k [c_{ik} + S_r(k)]$$

4. At the bottom node of the tree:

$$S = \min_i S_0(i)$$

# Maximum Likelihood

- Unterschied: Kosten sind von Kantenlänge abhängig



Stattdessen:  $W_{x \rightarrow y} = -\log(1 - \exp(-2r_{x \rightarrow y} t))$

$r_{x \rightarrow y}$ : Mutationsrate von x nach y

t: Zeit (Kantenlänge) zwischen Mutterknoten und Tochterknoten

# Maximum-Likelihood

Gesamt-Likelihood des Baumes:

$$\sum_x \pi_x L(\text{root}, x)$$

$\pi_x$ : stationäre Wahrscheinlichkeit von Charakterzustand  $x$

$L(\text{root}, x)$ : Likelihood von Charakterzustand  $x$  an Wurzel des Baums

- Der Maximum-Likelihood-Baum wird gesucht durch
  - Optimierung der Kantenlängen
  - Durchsuchen des Baum-Raumes

# Praktische Fragen

- Gutes Programm auch für Maximum-Likelihood-Analyse: **paup\***
- Verfügbar von  
[http://people.sc.fsu.edu/~dswofford/paup\\_test/](http://people.sc.fsu.edu/~dswofford/paup_test/)

# Praktische Fragen

- Starte paup\* von dem Verzeichnis aus, in dem die Nexus-Datei liegt

```
> execute Ielex_binarizedFull.nex
```

```
> set criterion=likelihood
```

```
> set storebrlens
```

```
> Hsearch
```

```
> SaveTrees file='ielexFull_ML.tree'  
format=Newick brlens=yes
```

```
> q
```