

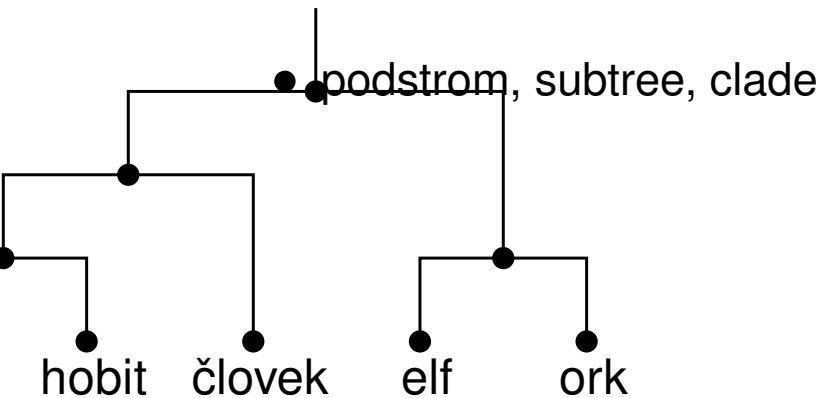
Fylogenetické stromy

Broňa Brejová

7.11.2022

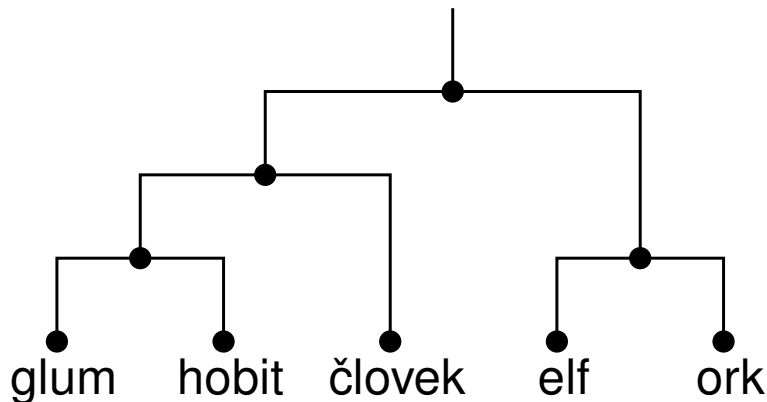
Terminológia

- zakorenený strom, rooted tree
- nezakorenený strom, unrooted tree
- hrana, vetva, edge, branch
- vrchol, uzol, vertex, node
- list, leaf, leaf node, tip, terminal node
- vnútorný vrchol, internal node
- koreň, root



Zopár faktov o stromoch

- Majme zakorenený strom s n listami, v ktorom má každý vnútorný vrchol 2 deti. Takýto strom vždy má $n - 1$ vnútorných vrcholov a $2n - 2$ vetiev (prečo?)
- Majme nezakorenený strom s n listami, v ktorom má každý vnútorný vrchol 3 susedov. Takýto strom vždy má $n - 2$ vnútorných vrcholov a $2n - 3$ vetiev.
- Koľkými spôsobmi môžeme zakoreniť nezakorenený strom s n listami?

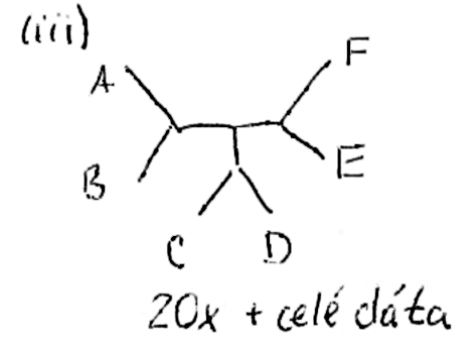
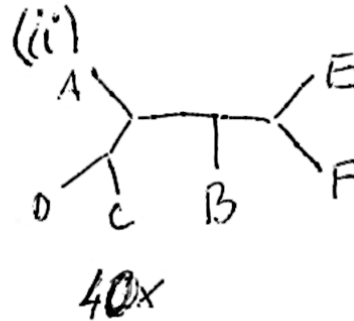
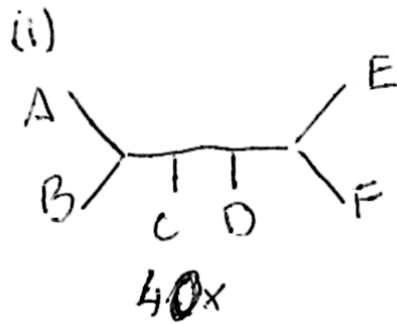


Bootstrap

- Náhodne vyberieme niektoré stĺpce zarovnaní, zostrojíme strom
- Celé to opakujeme veľa krát
- Značíme si, koľkokrát sa ktorá hrana opakuje v stromoch
(v nezakorenenom strome je hrana rozdelenie listov na dve skupiny)
- Nakoniec zostavíme strom z celých dát a pozrieme sa ako často sa ktorá jeho hrana vyskytovala
- Môžeme zostaviť aj strom z často sa vyskytujúcich hrán
- Bootstrap hodnoty sú odhadom spoľahlivosti, hlavne ak máme celkovo málo dát (krátke zarovnanie)
- Ak však dáta nezodpovedajú vybranej metóde/modelu, tak aj pre zlý strom môžeme dostať vysoký bootstrap

Bootstrap

Robili sme $100\times$ bootstrap, dostali sme tieto výsledky:



Doplňte bootstrap hodnoty hranám výsledného stromu (iii)

Ktoré ďalšie vetvy majú podporu aspoň 20%?

Aký strom by sme dostali, ak by sme chceli nechať iba vetvy s podporou aspoň 80%?