

## Zhrnutie semestra

**Broňa Brejová**

**19.12.2024**

## Pravdepodobnostné modely

- Skryté Markovove modely (hľadanie génov, konzervovaných oblastí, fylogenetické HMM, profilové HMM)
- Fylogenetické stromy a substitučné modely
- Stochastické bezkontextové gramatiky
- Gibbsovo vzorkovanie
- Metóda maximálnej vierohodnosti
- Expectation maximization (EM)
- Wrightov-Fisherov model náhodného genetického driftu

## Štatistické metódy

- Pojem štatistickej významnosti, E-hodnota a P-hodnota
- Test na pozitívny výber
- Väzbová nerovnováha, mapovanie asociácií

## Precvičenie dynamického programovania

- Zarovnávanie sekvencií  
(globálne, lokálne, afínne medzery)
- Skryté Markovove modely (Viterbiho a dopredný algoritmus)
- Výpočty na stromoch  
(úspornosť, vierohodnosť - Felsensteinov algoritmus)
- Hmotnostná spektrometria (MS/MS)
- Sekundárna štruktúra RNA

## Iné

- Celočíselné lineárne programovanie
- de Bruijnove grafy
- Zhlukovanie a klasifikácia

## Ako modelovať problémy reálneho sveta

- Rozmyslieť si, aké máme dáta, čo by sme chceli ako výsledok
- Sformulovať ako informatický problém (napr. optimalizácia nejakého skóre)
- Pravdepodobnostné modely nám často dovoľia zvoliť skórovaciu schému systematickým spôsobom
- Výsledný problém často NP ťažký
  - Heuristiky, aproximačne algoritmy
  - ILP a iné techniky na presné riešenie
  - Nedá sa problém trochu preformulovať?
- Testovanie: sú výpočtové výsledky relevantné v danej doméne? (bola formulácia dostatočne realistická?)

## Ďalšie predmety

- **Strojové učenie** 2-INF-150, Vinař/Boža (ZS, 4P, 6kr)
- **Vybrané partie z dátových štruktúr** 2-INF-237, Kováč (ZS, 4P, 6kr)
- **Seminár z bioinformatiky (1)-(4)** 2-AIN-50[56],25[12] (oba semestre, 2S, 2kr)
- **Manažment dát** 2-INF-185 Brejová, Vinař, Boža (LS, 1P/2C, 4kr)
- **Všeobecná biológia** N-bCXX-085/1-BIN-113, Tomáška (LS, 2P, 2kr)
- **Genomika** 2-INF-269, Nosek a kol. (LS, 2P/1C, 4kr)
- **Výzvy súčasnej bioinformatiky** 1-BIN-105, Vinař (LS, 2S, 2kr)