

Inference frameworks

Wright-Fisher process at the core

Summary statistics
and simple calculations
(small data, simple population)

Π : average pairwise divergence

S : # segregating sites

T_D : deviation from neutrality

Probabilistic models
(big data, more complex populations)

Forward time:

- PDEs / diffusion

- selective sweeps



Reverse time:

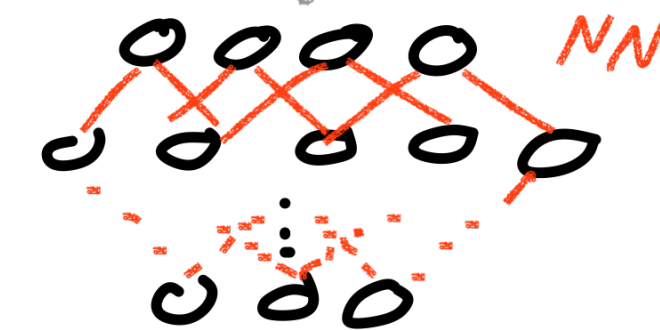
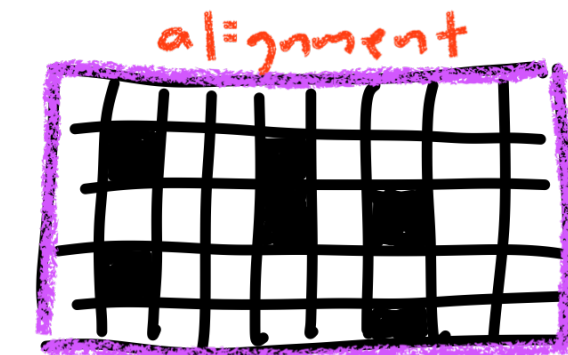
- coalescent genealogy

- coalescent HMM



Machine learning
(big data, arbitrarily complex population)

Supervised learning



Problem

The Wright-Fisher process is wrong!

Question: what's missing w/ WF model?