

Motif-finding by Gibbs Sampling

Problem. Given p strings and a length k , find the most “mutually similar” length- k substring from each string.

“Gibbs sampling” is the basis behind a general class of algorithms that is a type of local search.

It doesn’t guarantee good performance, but often works well in practice.

Assumes:

1. we know the length k of the motif we are looking for.
2. each input sequence contains exactly 1 real instance of the motif.