

## Project 2: Sequence Assembly

### Objectives:

1. Understand the sequence assembly problem.
2. Be familiar with different sequence assembly algorithms

### Task:

1. Implement a sequence assembly algorithm to assemble fragmented reads

### Instructions:

The input to this homework is a set of DNA reads (you can download it from Blackboard as [HW2\\_reads.fasta](#)). You can assume that if two reads have a suffix-prefix overlap longer than 50bp, they can be considered as being sampled from the same region of the genome and be reliably connected. Also note that minor sequencing errors (e.g. mismatches or insertion/deletion) may present in the overlapped region. [Submit your reconstructed genomic sequence in FASTA format.](#)

### Hint:

To ensure simplicity, the original genomic sequence contains no repeat.

### Submission:

Submit your reconstructed genomic sequence (in FASTA format) via Blackboard by Friday Mar 26<sup>th</sup>, 2021 11:59PM.