

# The Spaghetti Model: DNA Sequencing using Graph Theory

## Abstract

Graph theory is used to model a specific application of the Physical Mapping problem, which is to reconstruct the relative position of DNA fragments along the genome from pair wise overlap information. We call the model that we address in this talk the spaghetti model. A DNA strand is copied many times and cut into pieces of varying length using either an enzyme or a cleaving process, much like cracking handfuls of identical hard pieces of spaghetti. The pieces are then separated by length, and similar sized pieces are collected. Graph theory is used to order this collection of fragments. This problem can be mathematically modeled using a graph called a unit interval graph. An interval graph  $G$  is a graph whose vertices can be assigned to intervals on the real line such that two vertices are adjacent in  $G$  whenever their corresponding intervals overlap. A unit interval graph is an interval graph in which the intervals are of equal length. We prove that given a unit interval graph that models the overlap data, one can successfully order the pieces of the DNA sequence, and we give an algorithm to that effect. We also provide an algorithm which produces a minimum set of intervals necessary to order the entire DNA sequence.

1056-Q5-240.