208a DNA Sequencing by Hybridization with Errors Via Integer Programming
YoungJung Chang and Nick Sahinidis

DNA sequencing by hybridization (SBH) is a sequencing method that uses DNA chips to reconstruct a DNA sequence from fragmental subsequences. SBH has found many applications in custom re-sequencing and mutation detection, such as identification of SNPs, and remains an important experimental method. If the experimental data is free of error, the problem of reconstructing original DNA sequence reduces to the Eulerian path problem, for which linear-time algorithms are known. However, the same problem is NP-hard thus computationally challenging when there exist errors in the measurements.

Problems of DNA SBH with errors have been studied by many researchers, but no systematic works provide global optima, location of all alternative solutions, or analysis of the correctness of solutions. In particular, to guide correct selection and further experimental design, alternative solutions play a very important role in most experimental settings.

In this work, the DNA SBH problem is formulated as a mixed-integer linear program with an exponential number of constraints. A row generation solution algorithm is developed to solve this model. The proposed approach solved large SBH problems to global optimality efficiently and was able to locate all the alternative optimal solutions. These alternative solutions showed a wide range of quality in their correctness of reproducing the original sequence. This implies that obtaining a single solution could mislead the experimenter, thus justifying the development of algorithms for finding all alternative optimal solutions.