

## Abstract

In this work, we first present some basic concepts and definitions normally used in molecular biology. They are used, in order to describe some of most important combinatorial problems posed by the biologists.

We will give a special attention to the Closest String Problem (CSP), which consists in the determination of a sequence  $s_H$  of characters (belong to some alphabet  $\Sigma$ ) that is closer to a given set  $S = \{s_1, s_2, \dots, s_m\}$  of strings of  $\Sigma$  each of length  $n$ . The objective in this case is to find a sequence  $s_H$  such that the maximum distance (according to a given metric) between  $s_H$  and  $s_i$  for  $i=1 \dots m$  is minimized.

We study the proof of completeness of the CSP and concentrate our attention in the deterministic and randomized approximation algorithms listed in the literature. In fact, the major part of these techniques are based on probabilistic methods. For this reason, we present the most recent results, and give a detailed description of these strategies such as the randomized rounding procedure and the derandomization techniques, in particular, the method of conditional probability.

Finally, we develop a derandomization strategy using pessimistic estimators as proposed by Raghavan in 1988.

**Keywords:** closest string problem, computational biology, derandomization, randomized round.