Stochastic Diffusion Search in Bioinformatics

Andrew Moynihan

BSc in Artificial Intelligence

ABSTRACT

This paper explores the implementation of a Stochastic Diffusion Search within a Bioinformatic search space. A solution was designed to search DNA strings for known genes that may have been altered by any combination of common mutations: Additions, Deletions or Substitutions. The nature of SDS allowed the search to locate even heavily mutated genes and produce a list of errors for the user. The solution is also flexible enough to handle introns – non-coding sequences of DNA that may be present within the gene – provided that the pure gene is known.
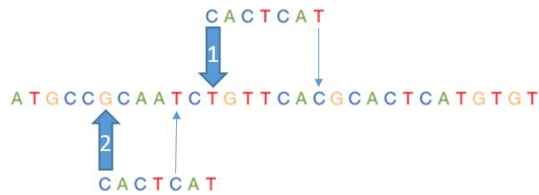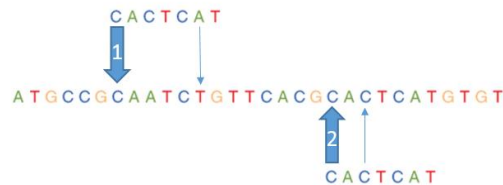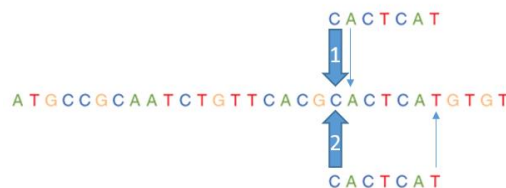
Figure 1

Figure 2

Figure 3

Figure 1 shows the initial state of two agents on a small section of gene. After both failing their tests, they stochastically diffuse to the positions in Figure 2. After testing again, one agent has a candidate location, and so remains stationary, where the other agent joins it via communication. Figure 3 repeats with random tests until the end condition is satisfied.