DATA COMPRESSION AND BIOINFORMATICS SEARCHING FOR STRUCTURE IN ALL KINDS OF PLACES

Khalid Sayood

University of Nebraska -Lincoln

ABSTRACT

At the heart of data compression is the search for how information is organized in data. This search for organization can focus on the data itself, on the source of the information, or on the sink, or user of the information. Analogous approaches can be used to identify how information is organized in the DNA sequence. In this talk we describe how techniques which have proven useful in developing compression algorithms are also capable of helping uncover organizational structures in DNA sequences. We also show how this process of discovery can lead to the development of tools which are useful in investigating and perha ps understanding genomic sequences.