Shunning External Memory Via Synopsis Data Structures

Y. Matias
Bell Labs and Tel Aviv University

Abstract

Traditional data structures are often inefficient for massive data sets, since they need to be stored on external memory. As a result, there has been recently considerable effort in designing algorithms and data structures that make effective use of external memory. One of the possible approaches is to compute and maintain in main memory data structures that are used to minimize, if not eliminate, access to external memory. Such data structures, which we call synopsis data structures, are significantly smaller than the data sets they represent; they need to represent the data set succinctly, in a manner that is suitable to the queries they are meant to support. In this talk we review recent results in the design and usage of synopsis data structures.