A Space-Bounded Anytime Algorithm for the Multiple Longest Common Subsequence Problem

Abstract:

The multiple longest common subsequence (MLCS) problems, related to the identification of sequence similarity, are an important problem in many fields. As an NP-hard problem, its exact algorithms have difficulty in handling large-scale data and time- and space-efficient algorithms are required in real-world applications. To deal with time constraints, anytime algorithms have been proposed to generate good solutions with a reasonable time. However, there exists little work on space-efficient MLCS algorithms. In this paper, we formulate the MLCS problem into a graph search problem and present two space-efficient anytime MLCS algorithms, use an iterative beam widening search strategy to reduce space usage during the iterative process of finding better solutions. Based on, a space-bounded algorithm is developed to avoid space usage from exceeding available memory. A replacing strategy when SA-MLCS reaches a given space bound. Experimental results show SA-MLCS and use an order of magnitude less space and time than the state-of-the-art approximate algorithm while finding better solutions. Compared to the state-of-the-art anytime algorithm Pro- and can solve an order of magnitude larger size instances. Furthermore, can find much better solutions than on large size instances.