

**Poster I-10**  
**Negative Result on Maximum**  
**Common Subgraph of Two Graphs**



**Authors:**

Xiuzhen Huang (*Department of Computer Science, Arkansas State University*)

**Short Abstract:** The problem of finding maximum common subgraph has important applications in biological structure matching. We investigate the maximum common subgraph of two graphs. We derive the hardness result for the exact algorithms of this problem. Our lower bound result is asymptotically tight.

**Long Abstract:**

One of the many challenging biological problems related to DNA, RNA, protein structures, and metabolic and signaling pathways, is the structure matching. The biological structures could be abstracted as graphs. The biological structure matching involves the comparison of a set of abstracted graphs. In bioinformatics, the biological structure matching problems can be formulated as the problem of finding the Maximum Common Sub-graph. The Maximum Common Sub-graph problem has found very important practical applications in bioinformatics as well as in other areas, such as pattern recognition and image processing. Approaches for the Maximum Common Sub-graph problem and different variants of this problem are intensively studied in the literature. We study the Maximum Common Sub-graph of two graphs. We first give a review on the parameterized computation and recent progress on parameterized intractability. Recent investigation in parameterized complexity has derived stronger computational lower bounds for well-known NP-hard parameterized problems in [1, 2]. For example, for the Clique problem, which asks if a given graph of  $n$  vertices has a clique of size  $k$ , it is proved that unless an unlikely collapse occurs in parameterized complexity theory, the problem is not solvable in time  $f(k)n^{o(k)}$  for any function  $f$ . Note that this lower bound is asymptotically tight in the sense that the trivial algorithm that enumerates all subsets of  $k$  vertices in a given graph to test the existence of a clique of size  $k$  runs in time  $O(n^k)$ . Based on the hardness of the Clique problem, lower bound results for a number of computational biology problems have been derived [3]. For example, our results for the problems longest common subsequence and shortest common super-sequence have significantly strengthened the former results on them and advanced our understanding on the complexity of the problems. Based on the current research progress in the area of parameterized computation, we first formally defined the parameterized version of the Maximum Common Sub-graph problem, where we choose the number of vertices in the common sub-graph as the parameter. Then we derive the hardness result for the exact algorithms of this problem. We prove that unless some unlikely collapse ( $\text{Mini}[1] = \text{FPT}$ ) occurs in parameterized computation, there is no algorithm of time  $f(k)n^{o(k)}$  for the Maximum Common Sub-graph problem, where  $k$  is the number of vertices in the common sub-graph and  $f$  is any recursive function. Our proof is based on a linear fpt-reduction from the Clique problem to the parameterized version of the Maximum Common Sub-graph problem as follows. We also discuss the approaches for addressing this problem. Basically, this problem can be reduced to find a maximum clique in the product graph of the two given graphs [4, 5]. By a close observation of the product graph  $G$  of two given graphs  $G_1$  and  $G_2$ ,

where  $G_1$  and  $G_2$  both have  $n$  vertices, we can see that  $G$  is indeed a  $n$ -partite graph, where the vertices are partitioned into  $n$  disjoint partitions with each partition  $n$  vertices. There is no edge between any two vertices within the same partition. Edges only appear between two vertices that are in two different partitions. So, at most one vertex from each partition (of the  $n$  vertices) could be in a clique of the graph. Therefore, to find a clique of size  $k$ , there will be  $n^k$  possible ways for choosing the clique vertices. For each possible way, it needs time  $O(k^2)$  to check if it constructs a clique of size  $k$ . Therefore, it takes time  $O(n^k * k^2)$  to find a clique of size  $k$  in the graph  $G$ . That is, the Maximum Common Sub-graph problem is solved in time  $O(n^k * k^2)$ , where  $k$  is the number of vertices in the common sub-graph. In consideration of the upper bound result, we point out that our lower bound result for the Maximum Common Sub-graph problem presented here is asymptotically tight. With the hardness result in mind, for future research, it will be very interesting research to further explore efficient approaches for different variants of this problem and also on this problem with constraints from real-world applications. [1] J. Chen, B. Chor, M. Fellows, X. Huang, D. Juedes, I. Kanj, and G. Xia, "Tight lower bounds for parameterized NP-hard problems," in Proc. Of the 19th Annual IEEE Conference on Computational Complexity, pp. 150-160, 2004. [2] J. Chen, X. Huang, I. Kanj, and G. Xia, "Linear FPT reductions and computational lower bounds," in Proc. of the 36th ACM Symposium on Theory of Computing, pp. 212-221, 2004. [3] J. Chen, X. Huang, I. Kanj and G. Xia, "W-hardness linear FPT reductions: structural properties and further applications," in proceedings of the Eleventh International Computing and Combinatorics Conference (COCOON 2005), Lecture Notes in Computer Science, vol. 3595, pp. 975-984, 2005. [4] V. Kann, On the approximability of the maximum common subgraph problem, Proc. 9th Annual Symposium on Theoretical Aspects of Computer Science, Lecture Notes in Computer Science 577, pp. 377-388, 1992. [5] J. W. Raymond and P. Willett, Maximum common subgraph isomorphism algorithms for the matching of chemical structures, Journal of Computer-aided Molecular Design 16, pp. 521-533, 2002.