Node-Weighted Trees for RNA Secondary Structures: Representation and Comparison

Jing Jin[1], Biplab K. Sarker[1], Virendra Bhavsar[1], Harold Boley[2] and Lu Yang[1]
[1] Faculty of Computer Science, University of New Brunswick, Fredericton, New Brunswick, Canada
[2] National Research Council, Fredericton, New Brunswick, Canada

Abstract*— An improved representation of RNA secondary structures and a similarity algorithm for their comparison are presented. In this representation, the nucleotides and nucleotide-pairs of an RNA secondary structure are modeled by a normalized node-weighted tree. We show that our weighted-tree representation of RNA secondary structures is more informative and useful than earlier representations. Based on this new representation we propose a weighted-tree similarity algorithm for computing the similarity between RNA secondary structures. The performance of the algorithm is exemplified by computing similarities among various RNA secondary structures. The representation and time complexity as well as accuracy of our algorithm are contrasted with existing RNA secondary structure comparison approaches.

Index Terms-- **Computational biology, RNA secondary structure, node-weighted trees, tree comparison, tree similarity algorithm.**