

<IS-1>: On an enhancement of RNA probing data using information theory

Thomas J. X. Li^{1,*}, Christian M. Reidys^{1,2}

¹*Biocomplexity Institute & Initiative, University of Virginia, 995 Research Park Blvd, Charlottesville, VA, 22911*

²*Department of Mathematics, University of Virginia, 141 Cabell Dr, Charlottesville, VA 22903*

jl9gx@virginia.edu

Identifying the secondary structure of an RNA is crucial for understanding its diverse regulatory functions. This presentation focuses on how to enhance target identification in a Boltzmann ensemble of structures via chemical probing data. We employ an information-theoretic approach to solve the problem, via considering a variant of the Rényi-Ulam game. Our framework is centered around the ensemble tree, a hierarchical bi-partition of the input ensemble, that is constructed by recursively querying about whether or not a base pair of maximum information entropy is contained in the target. These queries are answered via relating local with global probing data, employing the modularity in RNA secondary structures. We present that leaves of the tree are comprised of sub-samples exhibiting a distinguished structure with high probability. In particular, for a Boltzmann ensemble incorporating probing data, which is well established in the literature, the probability of our framework correctly identifying the target in the leaf is greater than 90%.