

1086-VG-1358      **Angela Angeleska\*** (aangeleska@ut.edu), **Sabrina Kleessen** and **Zoran Nikoloski**. *The Sequence Reconstruction Problem as an Integer Programming Problem*. Preliminary report.

Assembly of genomes from high-throughput data generated by the Next Generation Sequencing (NGS) technologies remains one of the most challenging tasks in modern biology. Here we address the Sequence Reconstruction (SR) problem whereby, for a given collection of subsequences or factors, one is to determine the set of sequences compliant with the collection. First, we give an overview of the SR problem from a language-theoretic perspective, and present the advantages and shortcomings of the existing algorithmic approaches. We then propose an optimization-based formulation, which casts the SR problem as a quadratic integer program whose solutions can be enumerated with appropriate transformations. This approach is relevant for the genome assembly problem and can be matched with the employed NGS technologies. (Received September 21, 2012)