

An Oracle for the Gapped Consecutive Ones Property, applied to Ancestral Genome Reconstruction

Ryan Coghlan¹, Cedric Chauve¹, and Tamon Stephen¹

Department of Mathematics, Simon Fraser University

The Consecutive Ones Property (C1P) of binary matrices has been shown to be an important theoretical concept in ancestral genome reconstruction [1]. Roughly speaking, a binary matrix is C1P if its columns can be reordered in such a way that the 1s on each row form a single block. In Ancestral Genome Reconstruction, columns represent orthologous markers and rows sets of markers that were believed to have been consecutive in the ancestral genome (Ancestral Contiguous Sets, or ACS).

However, the C1P is a relatively rigid combinatorial concept; it is very sensitive to small errors in the binary matrix, especially missing markers in ACS, or events such as convergent events of local rearrangements. This leads to investigating relaxed notions of C1P, such as the gapped-C1P, studied among others in [3]: in this problem, gaps of bounded size (a given parameter) are allowed to occur in the matrix after reordering the columns. It was shown in [2, 3] that this problem is NP-complete as are several parameterized variants.

In the present work, we follow the relationships between the gapped-C1P concept and graph bandwidth and which was described in [2]. We present an implementation of an exponential time algorithm for deciding the gapped-C1P of binary matrices, outlined in [2] and inspired of the algorithm described in [4] to decide the graph bandwidth property. We apply our program on data for the reconstruction of vertebrate ancestral genomes.

References

1. C. Chauve and E. Tannier (2008), *PLoS Comput. Biol.*, 4(11):e1000234.
2. C. Chauve, J. Manuch and M. Patterson (2009), in *EUROCOMB 2009*.
3. J. Manuch and M. Patterson (2010), in *RECOMB-CG 2010*.
4. J. Saxe (1980), *SIAM J. Algebraic Discrete Methods*, 1:363-369.