
Cyril GAVOILLE
Tel: + 33 5 40 00 88 12
Fax: + 33 5 40 00 66 69
E-mail: gavoille@labri.fr

Talence, March 31, 2025

Subject: *Report on the doctoral dissertation of WOJCIECH JANCZEWSKI.*

The thesis of WOJCIECH JANCZEWSKI primarily focuses on informative labeling schemes for graphs and, to a lesser extent, dynamic algorithms for maintaining the longest increasing subsequence.

The aim of informative labeling schemes is to design data structures distributed among the vertices of a graph, represented as a label or bit-string, such that queries on vertices of the graph can be solved from the labels of the vertices involved in the query, and nothing else. The main goal is to minimize the size of the labels while supporting fast queries. In other words, the challenge is to pre-process a graph so that information about the vertices can be extracted from local information only and with as little information as possible. Such graph optimization problems are nowadays an important topic in algorithmics, with many applications including but not limited to queries in large databases, networking, and distributed computing. It is also a very challenging topic.

The longest increasing subsequence is a rather classical problem for which standard dynamic programming solutions exist. The dynamic variant involves maintaining (dynamically) the longest increasing subsequence (or an approximation of that length) in a sequence of comparable elements under insertions and deletions. This is a very challenging and non-trivial problem for which no efficient solutions are known. This variant has applications in data streaming and has connections with other combinatorial problems like Erdős-Szekeres partitioning.

The 160-page document, written in English, is structured into six chapters, which are detailed below.

Chapter 1 presents the main results of the thesis, including the complete state-of-the-art and related results. Most state-of-the-art selections are nicely presented in tables.

Chapter 2 introduces labeling schemes, as the next three chapters will focus on this topic. It is, however, centered on trees. Known schemes for trees are presented in a cleaner and simpler way so that new results can easily emerge. In particular, it is shown that rooted trees with n vertices have a multi-functional labeling scheme with $\log_2 n + O(\log \log n)$ -bit labels. The supported queries (or functions) include ancestry, adjacency, siblings, and parenthood. In fact,

all these queries can be solved easily since the labeling scheme assigns to each vertex v of the tree a two-part label $\ell(v) = (\text{ID}(v), \lambda(v))$, where $\text{ID}(v)$ is a unique $\log_2 n + O(\log \log n)$ -bit identifier for v , and $\lambda(v)$ is an extra $O(\log \log n)$ -bit label. The scheme allows decoding from $\ell(v)$ the ID of v 's parent and answering ancestry queries given two such labels. All these queries take constant time to decode.

Chapter 3 concerns routing in trees with very short labels. It is assumed that each incident edge of any node v of the tree receives a specific number, called a port number, carefully selected in $\{1, \dots, \deg(v)\}$ by the designer of the labeling scheme (a.k.a. the designer port model). It was known that $O(\log n)$ -bit labels suffice to answer every $s - t$ -routing query: Given only the two labels $\ell(s)$ and $\ell(t)$, what is the port number of the first edge in the tree on the path from s to t ? The current best label length is actually $\log_2 n + O(\log n \log \log \log n / \log \log n)$ bits. After reviewing the best solution, WOJCIECH JANCZEWSKI presents several schemes, each one more succinct than the previous ones, and eventually ends with a $O(\log^2 \log n)$ second-order term. The approach is interesting as, for each intermediate step, the weakness of the scheme is pointed out, and a solution is provided. A lower bound of $\log_2 n + \Omega(\log \log n)$ is presented, which is very close to the upper bound. The chapter also presents two other interesting label-based routing schemes, each one splitting differently the label information of v into two asymmetric pieces: the address of v and the local routing tables of v . The classic routing labeling scheme identifies both parts. Then, the routing process has to return the right port number in s given the address of t and the local routing table of s . I should say that the resulting schemes are highly non-trivial.

Chapter 4 concerns adjacency labeling in n -vertex planar graphs. Here, the goal is to decide whether u and v are adjacent, given $\ell(u)$ and $\ell(v)$. The problem has a long history with a recent breakthrough with a $\log_2 n + O(\sqrt{\log n \log \log n})$ -bit label solution. I know this solution very well, and I must admit that it is very involved. The new scheme is again elegantly presented in this chapter. It is simplified enough to save a $\sqrt{\log \log n}$ factor in the second-order term. This is done by a reduction to a data-structure problem in trees and with the use of B -trees (which are more efficient here than bulk-trees as used in the original scheme). This better scheme also allows slightly improving the complexity to decode adjacency from the labels, from $o(\sqrt{\log n})$ to constant. It follows a slightly better bound on the size (vertices and edges) in the induced-universal graph for n -vertex planar graphs.

Chapter 5 concerns another type of labeling scheme, distance labeling schemes, where the goal is to compute the distance between two vertices given their labels. It focuses on permutation graphs for which it was known how to achieve such schemes with $9 \log_2 n + O(1)$ -bit labels. Again, the state-of-the-art previous solution is presented and analyzed, and elegant hints are provided to achieve a new scheme with $3 \log_2 n + O(\log \log n)$ -bit labels and constant time decoding. This is quite technical and matches the lower bound, up to the second-order term.

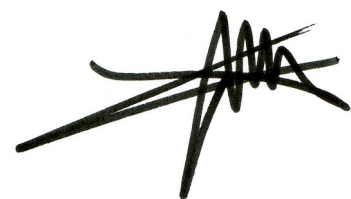
The thesis ends with Chapter 6 about a different problem: Maintaining an approximation of the longest increasing subsequence (LIS) under insertions and deletions of any element in the sequence of length n . It is only very recently that sublinear update solutions in the fully dynamic setting are known, namely an $(1/\varepsilon)^{O(1/\varepsilon)}$ -approximation with $\tilde{O}(n^\varepsilon)$ updates, and an exact algo-

rithm with $\tilde{O}(n^{2/3})$ updates. WOJCIECH JANCZEWSKI presents the first $(1-\varepsilon)$ -approximation algorithm with poly-logarithmic updates, namely $O((1/\varepsilon)^5 \log^{11} n)$ update time. As intermediate steps, a connection with the Erdős-Szekeres partitioning is made (making a partition of any sequence of integers into $O(\sqrt{n})$ monotone subsequences), providing a better solution for this latter problem. The new partitioning algorithm runs in time $\tilde{O}(n)$, a significant improvement upon the $O(n^{3/2})$ time of the previous best solution. The $\tilde{O}(n)$ time complexity is almost tight according to an $\Omega(n \log n)$ lower bound of any comparison-based algorithm presented also in this chapter. Although the solution is built step by step, the resulting solution is very intricate.

I am impressed by the quality of the drafting of the document and the significance of the solutions. Most of the presented results have already been validated and published in well-established international conferences: SODA '21 (Chapter 2 about routing in trees), SOSA '22 (Chapter 3 about adjacency in planar graphs), and STOC '21 (Chapter 6 about dynamic approximated LIS algorithm); other chapters correspond to arXiv preprints. The significance of these works is shown by the selectivity of these venues; in particular, STOC and SODA, which are the top annual conferences on algorithmics research and accept only the best submissions of the year in this area.

Overall, the thesis is very readable and well-written, presenting difficult and technical material in an accessible way, without any sacrifice in precision and rigor. Doing so is a difficult but important skill, which WOJCIECH JANCZEWSKI has mastered well. The contributions presented in this dissertation are strong, deep, and important in the challenging context of labeling schemes and fully dynamic algorithms. The high-quality thesis clearly demonstrates WOJCIECH JANCZEWSKI's mastery of these subjects, his ability to formulate and solve important problems in this area, and his ability to communicate those results clearly. It is clearly well above the standard for acceptance for a doctoral thesis in this area.

I would be very happy to have a student who produced such a thesis with such a collection of results, and I highly recommend its acceptance as a doctoral dissertation in its present form. In conclusion, I consider the dissertation as excellent, and for a University in France, I would recommend it for a scientific prize of the university.



Cyril GAVOILLE, Professor
University of Bordeaux.