



New York University

A private university in the public service

Dennis Shasha
Courant Institute of Mathematical Sciences
Department of Computer Science
New York University
251 Mercer Street
New York, N.Y. 10012, USA
Telephone: (212) 998-3086
Electronic mail: shasha@cs.nyu.edu
Web: <http://cs.nyu.edu/cs/faculty/shasha/index.html>

January 1, 2006

Professor Alfredo Ferro
Professor Vincenzo Cutello
Dip. Matematica e Informatica,
Citta Universitaria,
viale A. Doria no 6
95125 CATANIA
Italy
Fax number: 011-39 095 733 7032

To Whom It May Concern,

Rosalba Giugno worked with me at New York University for two years of her PhD program on subgraph searching. Since her graduation, we have continued to collaborate. Her work resulted in the graphgrep software. That software has been available on the web and has been downloaded by more than 200 groups from all over the world. The work has also been published as two papers in the prestigious conferences ACM Principles of Database Systems 2002 and the IEEE International Conference in Pattern Recognition 2002.

The work has to do with searching for small query graphs in databases of graphs. This is an important problem in many applications from computer vision to chemistry. Several searching systems exist for chemical databases but not much research has been done for application-independent searching in database of graphs. Rosalba's work entailed both a language for searching that generalizes regular expressions for strings as well as algorithms and software to execute that language. Compared with the other best software in this area, the commercial product Daylight (a searching system for a database of molecules), the software is of similar speed (sometimes faster and sometimes slower) and is much more flexible. Dr. Giugno's research will therefore permit inexpensive data analysis across many domains. The language she proposes *Glide* is a useful combination of XPath with Smile. XPath is an XML standard and Smile is a biochemical standard. By combining the two, Glide allows the expression of a wide variety of queries including queries having fixed and variable length don't cares (the XPath contribution) while generalizing from

paths to graphs (the Smile contribution). The algorithms involve a relational and string searching sublanguage on an algebra. This was first implemented in K and then in C and C++. The primitives involve a representation of the database graphs as a collection of small subgraphs (typically paths that may contain cycles), joins between paths and selections of nodes. The net result has a time complexity that is exponential in the size of the data graph, but polynomial in the size of the data graph, and linear in the size of the number of graphs in the database.

Rosalba structured the problem very well and developed the necessary algorithmic components. Her intelligence is obvious; what sets her apart is her ability to think independently, her eagerness for rigor, and her systematic approach to problem solving and reading the literature. These traits, combined with her work ethic, strongly suggest that she will continue to make contributions of the highest scientific standards. Besides her work on GraphGrep and Glide, Rosalba advised several undergraduate and Master's level students at NYU, and prepared lecture and conference presentation material. From the student feedback and my own experience, Rosalba has excellent communication skills, and is a dedicated teacher. On a personal note, Rosalba is a wonderful collaborator, is very responsive, and writes well. She is collegial, willing to share and disseminate her work. My collaboration with Rosalba has continued. GraphGrep's algorithms have been extended, and its data structures have been made more efficient, resulting in new software named GraphBlast. This software is part of the Cytoscape software package, an internationally used package for computational biology.

I enthusiastically support her application for a position at the University of Catania. With her talents Rosalba would be an asset and an outstanding faculty member at any top-rated university. Please do not hesitate to contact me if you need any further information.

Warm Regards,

Dennis Shasha