



Prague, August 20, 2007

The Board of Doctoral Study
Faculty of Mathematics and Physics


Report on the PhD Thesis by Irena Mlýnková

The submitted PhD thesis addresses an important issue in XML technology, namely XML-to-relational storage strategies. The problem is how to store XML data in a relational database, such that it can be approached effectively and efficiently. This theme is highly topical and relevant w.r.t its significant practical consequences.

The work focuses particularly on adaptive techniques that seem to have the most significant impact on efficient XML processing. Chapter 2 covers a good spectrum of the background and current research literature in this area. The author also formulates open issues concerning the problem. Her original approach – the Hybrid User-Driven Adaptive Method is contained in Chapter 3, including results gained from an experimental implementation. The method transforms fragments of a XML schema into a set of relations. Key notions used here are fragment annotations and similarity of fragments. Chapter 5 is devoted to the latter. The resulted measure is a special aggregation function with parameters tuned with a help of statistical properties of real XML data collections and machine learning methods. The statistical analysis of XML data uses results from a large-scale research proceeded in Department of Software Engineering in 2006. Details of this significant research done by I. Mlynkova a K. Toman are described in Chapter 6. Chapter 7 is devoted to experiments with query evaluation, in other words, to the proof of concept. A prototype system called UserMap is described there and first experiences its application.

The thesis uses an appropriate and sound formal apparatus as well as clear scientific methods, which indicates that the candidate has gained both theoretical and practical experience during her studies. The thesis also offers a lot of possibilities for other research.

The results presented in the thesis have been published in proceedings of representative international conferences as well as in the International Journal of Computer Science and Applications. Some results are also cited internationally in a number of conference papers. Two papers of Irena Mlynkova (with co-authors) were selected as the Best student papers of the conference COMAD (in 2006) and the Best paper of the conference RCIS (in 2007), respectively. Consequently, I recommend that the candidate be awarded the Doctor degree.



supervisor