



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Scientific stay

Université Blaise Pascal, Clermont-Ferrand, France

Jan Outrata

Aug 22, 2011 - Sep 16, 2011

Place

- Laboratory of computer science, modeling and systems optimization (Laboratoire d'Informatique, de Modélisation et d'Optimisation des Systèmes, LIMOS), Université Blaise Pascal (UBP), Clermont-Ferrand, France



Figure: Building of LIMOS

Place

- Laboratory of computer science, modeling and systems optimization (Laboratoire d'Informatique, de Modélisation et d'Optimisation des Systèmes, LIMOS), Université Blaise Pascal (UBP), Clermont-Ferrand, France
- research areas: graph theory, algorithm optimization, image processing
- formal concept analysis (FCA) a relative areas: several basic FCA algorithms (known by author names Nourine, Berry, Sigayret who work there) and applying FCA in machine learning
- at present cca 80 professors and doctors of the university and cca 70 PhD students and posdocs (form that one half abroad)
- research is funded by CNRS national research organization, Auvergne region and private industry

prof. Engelbert Mephu Nguifo



- + co-workers and PhD students in LIMOS
- research: applying FCA methods in data mining (association rules generalizations and computing frequent closed itemsets, algorithms for computing formal concepts), machine learning (classification with use of concept lattice and neural network) and bioinformatics (classification of biological sequences) – significant results in using FCA
- grants and projects are funded from european, national and regional sources, mostly, however, from private industry
- close collaboration with people at El-Manar University in Tunis

Stay run

- discussion and preparing a paper on using Boolean factor analysis (BFA) of input data for FCA as a clustering method
 - the paper was started during my previous visit together with R. Bělohlávek
 - the paper was considerably improved and extended for presentation on international conference CLA 2011 and inclusion in the conference proceedings
 - collaboration with paper coauthors dr. Sylvie Guillaume (coworker) and Dhouha Grissa (PhD student)
 - results on researched data appear as very promising for applying BFA in other areas

Stay run

- discussion of possibilities of further research directions in this area, e.g. the influence of input data duplicity or noise removal on resulting factors, early observations:
 - duplicity removal has almost no effect on the first, most significant factors, while on the less significant the effect is large – this is in accordance with expectations that the most significant factors will be in this sense stable
 - in noise removal similar behavior is expected – noise in input data should be covered by the least significant factors
- discussion with dr. Raoul Medina (co-worker) on relations between functional dependencies in databases and decision trees in machine learning, in connection with FCA

Publication

R. Belohlavek, D. Grissa, S. Guillaume, E. Mephu Nguifo, J. Outrata: *Boolean factors as means of clustering of interestingness measures of association rules*. In: Napoli A., Vychodil V. (Eds.): CLA 2011: Proceedings of the 8th International Conference on Concept Lattices and Their Applications, 2011, pp. 207–222, Nancy, France, October 2011.

[INRIA Nancy – Grand Est and LORIA, France, Nancy 2011, ISBN 978-2-905267-78-8]

- using Boolean factor analysis as a clustering method over data describing measures of interestingness of association rules and comparing resulting factors with clusters produced by traditional clustering methods (agglomerative hierarchical clustering and K-means)
- factors form overlapping clusters while traditional clusterings produce nonoverlapping ones

Contacts

- strengthening of scientific collaboration with the group of prof. Engelbert Mephu Nguifo
- developing of current and setting up new topics for further collaboration

Conclusion

- the stay fulfilled its purpose
- strengthening of scientific collaboration with one of the most significant centers in the area of formal concept analysis (FCA) and relative areas (graph theory)
- a paper presented on international conference



Figure: A foto from discussion on the paper. Depicted prof. Mephu Nguifo (on the right), dr. Guillaume (in the middle) and dr. Outrata (on the left).