LIRIMA Scientific Days
September 12 – 14, 2017, Tunis

Jean-Daniel Zücker
UMMISCO UMI 209, IRD-UPMC, France

Title: « Machine Learning Challenges in Metagenomics, Data Analysis for Personalized Medicine »

Abstract:
High-throughput technologies and today Next Generation Sequencing (NGS) have allowed the production of large genomic datasets: for instance, microarray data contain the simultaneous expression of tens of thousands of genes or NGS data may reach several millions of genes counts. Building Biomarkers from these data is a key problematic for tomorrows diagnostic and personalized medicine. In this talk we will address several questions raised when building classifiers based on such High-Dimensional data. Feature selection, Network Reconstruction, learning classifiers, Model Stability are several of the issues we will discuss in such settings. We will introduce the notion of ternary weight classifiers well present an algorithm adapted to deal with metagenomics data. Finally, we will discuss the feasibility to address such type of problems using Deep Learning.

Keywords: High dimension Data Mining, Metagenomics, Feature Selection, Network reconstruction.

Biopics: Jean-Daniel Zucker is a former Engineer (ESIAE, 1985) in Computer Science and Aeronautical Engineering. He then received in 1986 a Master in Artificial Intelligence applied to Life Science. He worked for the New England Medical Center (Boston, USA), IBM and Thomson in R&D for six years. After a Master in Artificial Intelligence (Major in Machine Learning) in 1992 he got is PhD. in 1996 in Machine Learning from Paris 6 University where he became an assistant and then associate professor focusing on representation changes and abstraction in machine learning. In 2002, he became Full Professor of Computer Science at Paris 13 University where he led a CNRS (the French NSCF) team on Machine Learning and Transcriptomics. In 2008, he became a Senior Researcher at the National Institute of Research for Development (IRD) working on Data Mining and Complex Systems. He is the director of the Laboratory UMMISCO UMI 209 (IRD and University Pierre and Marie Curie in Paris, France). He was posted in Hanoi Vietnam for five years and did participate to several projects in South-East Asia and Europe. He is now back in France where he pursues research in Machine Learning and Data Integration for Personalized Medicine. In the past twenty years his main research interest has been Machine Learning, Data Mining and Multi-Agent Modeling with applications ranging from Post-genomics to Environmental Decision Support systems. He has a Google Scholar h-index of 35, he has written more than 240 publications. He co-authored recently a book published by Springer “Abstraction in Artificial Intelligence and Complex Systems”.