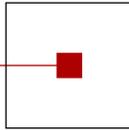


s c c h

software competence center
hagenberg



Advances in Knowledge-Based Technologies

Proceedings of the
Master and PhD Seminar
Winter term 2016/2017, part 1

Softwarepark Hagenberg
SCCH, Room 0/2
18 November 2016

Software Competence Center Hagenberg
Softwarepark 21
A-4232 Hagenberg
Tel. +43 7236 3343 800
Fax +43 7236 3343 888
www.scch.at

Fuzzy Logic Laboratorium Linz
Softwarepark 21
A-4232 Hagenberg
Tel. +43 7236 3343 431
www.flll.jku.at

Program

Session 1. Chair: Susanne Saminger-Platz

- 9:00 B. Moser:
Similarity Recovery from Threshold-Based Sampling beyond Bandlimitedness
- 9:30 J. De Jesus Arias-Garcia:
Supermodular quasi-copulas

Session 2. Chair: Bernhard Moser

- 10:15 W. Zellinger:
Central Moment Discrepancy for Domain-Invariant Representation Learning
- 10:45 L. Paoletti:
Top-k Matching Queries for Filter-Based Profile Matching in Knowledge Bases

Similarity Recovery from Threshold-Based Sampling beyond Bandlimitedness

B. Moser, *AdvKBT Seminar, Hagenberg, Nov. 2016.*

This talk reports from ongoing research and gives a sketch of a mathematical framework for similarity recovery from threshold-based sampling based on a generalized concept of signed measures. By this approach a signal f is mapped to a signed measure μ_f that satisfies a generalized pseudo-additivity condition $\mu_f(I) \oplus \mu_f(J) = \mu_f(I \cup J)$ for disjoint intervals $I \cap J = \emptyset$. This model comprises standard threshold-based sampling schemes such as “Send-on-Delta” (SOD) or “Leaky Integrate-and-Fire” (LIF). Furthermore, it also contains the psychophysical Weber-Fechner law as examples and gives rise to novel sampling schemes such as “Leaky SOD”. We motivate this approach by previous findings in the context of Weyl’s discrepancy measure and point out that this model induces a pseudo-metric $d(\cdot, \cdot)$ in the signal space given by $d(f, g) = \sup_{s,t} |(\mu_f - \mu_g)([s, t])|$ and a corresponding pseudo-metric in the resulting sampling space. With these pseudo-metrics the input and output spaces become quasi-isometric and the threshold-based sampling scheme turns out to be a quasi-isometric mapping and an asymptotic isometry for arbitrary small thresholds. The corresponding inequalities are valid under general conditions which go beyond bandlimitedness or similar restrictions as typically encountered in the signal reconstruction context.

Keywords: Threshold-based sampling, Weber-Fechner law, quasi-isometry, discrepancy norm, Alexiewicz norm, \oplus -measures

Supermodular quasi-copulas

José De Jesús Arias-García¹

KERMIT, Department of Mathematical Modelling, Statistics and Bioinformatics
Ghent University, Ghent, Belgium
`josedejesus.ariasgarcia@ugent.be`

In this talk we recall the concept of an n -quasi-copula, its basic properties and their relationship with n -copulas and highlight some properties of 2-copulas that cannot be extended to higher dimensions.

Then, we notice that as dimensionality increases there are more intermediate classes between the class of n -quasi-copulas and the class of n -copulas. One of such classes is the class of supermodular n -quasi-copulas which have similar properties to the ones of 2-copulas that cannot be generalized to higher-dimensional copulas.

Central Moment Discrepancy for Domain-Invariant Representation Learning

Werner Zellinger

Department of Knowledge-Based Mathematical Systems,

Johannes Kepler University Linz, Austria

E-mail: werner.zellinger@jku.at

Abstract: The learning of domain-invariant representations in the context of domain adaptation with neural networks is considered. In particular a new regularization method (CMD) is proposed that is based on differences of higher order central moments. CMD is used to minimize the domain discrepancy of the latent feature representations explicitly in the hidden activation space. In contrast to standard approaches, e.g. "Maximum Mean Discrepancy" (MMD), computationally expensive distance- and kernel matrix computations are unnecessary. We define CMD to be an empirical estimate of a new metric introduced in this paper. We prove that convergence of bounded random variables w.r.t. to the new metric implies convergence in distribution of the random variables. We test our approach on two different benchmark data sets for object recognition (Office) and sentiment analysis of product reviews (Amazon reviews). CMD achieves state-of-the-art performance on most domain adaptation tasks of Office and Amazon reviews. In addition, a post-hoc parameter sensitivity analysis shows that the new approach is stable w.r.t. parameter changes in a certain interval.

Online Transfer Learning for Climate Control in Residential Buildings

Lorena Paoletti

Department of Data Analysis Systems,
Software Competence Center Hagenberg GmbH,
E-mail: lorena.paoletti@scch.at

Abstract: Finding the best matching job offers for a candidate profile or, the best candidates profiles for a particular job offer, respectively constitutes the most common and most relevant type of queries in the Human Resources (HR) sector. This technically requires to investigate top-k queries on top of knowledge bases and relational databases. We propose in this paper a top-k query algorithm on relational databases able to produce effective and efficient results. The approach is to consider the partial order of matching relations between jobs and candidates profiles together with an efficient design of the data involved. In particular, the focus on a single relation, the matching relation, is crucial to achieve the expectations.