

MatInfo Research Center

– detailed research plan 2023–2025–

In what follows we propose a detailed research plan for the period 2023–2025:

- **Deep learning:** Many real-world application use data captured over the time constituting a time-series. The analysis of these time-series may require domain experts able to define hand-crafted features. Deep learning proposes new models for unsupervised learning of the features, however, each type of time series may require specific deep model architecture. The group are especially interested in proposing reliable deep model architectures for time series generated by human behavior and to develop biometrical systems based on such models. One specific research direction would be to apply transfer learning. In this case the time series should be transformed into images then a proper architecture of deep model should be used. Another research direction is to use other deep learning models such as autoencoders, Long Short-Term Memory models and Generative Adversarial Networks (GAN). The group is also interested in the interpretation of these models' predictions.
- **Applied Computational Intelligence:** The group focuses on the study and creation of new models in the fields of machine learning, soft computing and intelligent systems with a focus on real-world applications. The group is particularly active in the areas of combinatorial optimization, bio-inspired computing, metaheuristics for NP-hard problems, model-based search and optimization, classification, fuzzy clustering, multi-agent systems, data mining, feature selection and engineering, ensemble methods in machine learning and deep structured learning.

On a short and medium term, the group intends to develop methods applicable to segmentation and pattern recognition in medical images. Medical imaging is a fast-evolving industrial sector, where improved image-making technologies are iterated at a rapid pace. The number of operational medical imaging equipment is also growing rapidly. For extending life span and increasing life quality, health care programs in both developed and developing countries, are increasingly focusing on health preservation and on the pervasiveness of screening tests. In our modern world, the growth in availability and the sheer number of medical imaging records far outpaces the supply of human professionals needed for evaluation. Therefore, it is paramount to develop methods and algorithms for assisting human experts. If automated data processing algorithms perform the pre-screening of the images with sufficient accuracy, doctors will be able to diagnosis and begin treatment at an early stage all patients willing to participate in the screenings.

The medium-term output goal of the group is to develop sophisticated methods, that among other things facilitate the accurate segmentation of various medical images and the identification of pathological lesions from the image data. When processing the image data, we use the developed methods and models to classify pixels or superpixels. We wish to increase the classification accuracy by using multi-atlases. We use images from negative cases (healthy patients) to create atlases that describe what type of tissues are generally present in different parts of the organ being tested or what is the probability distribution of the normalized intensity values. A fusion procedure combines the information provided by the atlases and the result of the classification and based on this we recommend a diagnosis to the doctor.

- **Computer assisted learning (AlgoRythmics: promoting computational thinking for all)**

Computational Thinking (CT) is a fundamental skill for everyone in our digital era. Since algorithms are central to CT and they are everywhere in modern society, a developed CT may be beneficial for a wide range of human activities. Accordingly, developing students' CT should be included as an objective in all educational programmes.

The AlgoRythmics research group is a promoter of the “algorithms for all” initiative. We developed an e-learning environment that includes interactive animations and dance choreographies illustrating basic computer algorithms. In other words, we have proposed to enhance CT by inviting users on a “delicious algorithmics tasting tour” in the mini-world of computer algorithms. Previous research has revealed that the AlgoRythmics visualizations are effective tools: (1) within the framework of programming education and (2) for improving the CT of all students (including those that have no prior programming experience).

The research group (besides extending the environment with new algorithms) is interested in the following research topics: (i) Schematic versus realistic algorithm visualization; (ii) Human movement effect enhanced algorithm visualization; (iii) Social belonging phenomenon in dynamic visualization environments; (iv) Can the AlgoRythmics environment be effective for improving the CT of different category of learners? (v) Is there an optimal level of interactivity in algorithm visualization environment?

- **Game Theory:** Modern societies increasingly rely on the proper functioning of complex networks like the telecommunication and transportation networks, power grids, global trade networks, etc. The collective actions of many independent decision-makers influence the dynamics of these networks. Therefore, stable growth and development require a proper understanding and control and optimization of such networked systems. We wish to study and model networked systems by designing efficient algorithms for graph games with stochastic transitions, evolutionary games on graphs and learning on graphs.

- **Applied Algebraic Geometry:**

- *Infinite Discriminant.* For a general data point the number of complex critical points is constant, while the number of real critical points is typically not constant for all general data. There are four ways in which we can have different number of critical points than expected. First, a data point can have a critical point with multiplicity. This locus is the ED discriminant. The second reason is because a critical point may wander of into the singular locus of the variety. In a similar fashion, the third case is when a critical point becomes isotropic with respect to the Euclidean inner product; this locus is called ED data isotropic locus. In the latter two cases the number of critical points is always smaller than the ED degree. Finally, a data point can have infinitely many critical points. One of the concrete goals of this research is to determine the locus of data points which have infinitely many critical points. A classical example would be that there are infinitely many critical rank 2 approximations of a matrix with two identical singular values. Preliminary computations show that the locus of data points with infinitely many critical points is part of the ED discriminant.
- *Euclidean distance degree and curvature.* Till this moment the most effective way to determine the Euclidean Distance degree of a general variety was by computing *Chern classes*. The promising article of Amelunxen and Bürgisser from 2012, shows a connection between the curvature measure of the cone of positive semidefinite matrices and the complexity of the corresponding semidefinite programming problem. Since in the case of algebraic varieties the algebraic complexity is expressed by the Euclidean Distance degree (at least for distance optimization) it would be interesting to find connections to between Euclidean Distance degree and the curvature of the underlying variety, in the fashion of the article of Amelunxen and Bürgisser. Let us remember as well that for real curves, the function, that counts the real critical points, is a function of *the curvature of the given curve*. So our goal is *to find connection between the curvature measure of the underlying variety X and its Euclidean Distance Degree*. In case of positive answer we would have new method

of computing Euclidean Distance degree, which for some varieties could be more effective than computing Chern classes.

- **Geometric Analysis:**

- *Functional inequalities: theory.* are often manifestation of natural physical phenomena as they often express very general laws of nature formulated in physics, biology economics and various aspects of engineering. They are also form the basis of fundamental mathematical structures such as calculus of variations, geometrical analysis etc. The optimal Gagliardo-Nirenberg (GN) inequality generates the optimal logarithmic inequality (called also as the entropy-energy inequality involving the Shannon entropy). Conversely the logarithmic inequality can be viewed as an endpoint of a family of optimal GN inequalities that also connects to Sobolev's inequality. Using mass transportation argument, Cordero-Erausquin, Nazaret and Villani extended the results to prove optimal Gagliardo-Nirenberg inequalities on arbitrary normed spaces. The family of the GN inequalities on curved space was studied by A. Kristály. Del Pino, Dolbeault, Filippas and Tertikas studied the connection between the classical Hardy inequality and the optimal Sobolev inequality, obtained a weighted logarithmic Sobolev inequality. Considering the previous studies, it is a challenging problem to study some first or higher order Sobolev-type inequalities and their limit cases (as log-Sobolev-type inequalities, interpolation inequalities) on curved spaces.
- *Functional inequalities: applications.* Most of the physical and engineering phenomena can be modeled by means of certain elliptic equations which can be studied using different variational methods. Functional inequalities are indispensable to investigate PDEs where the solutions are being sought on Sobolev spaces; see e.g. the analytic form of the Yamabe problem or Perelman's approach to the famous Poincaré conjecture. To be more precise, consider the following model problem

$$\begin{cases} -\mathcal{L}u(x) = f(x, u(x)), & x \in \Omega, \\ u \text{ subject to some boundary condition,} \end{cases} \quad (\mathcal{P})$$

where $\mathcal{L}$ is an elliptic Laplace-type operator (e.g., the classical Laplace operator on Euclidean setting, the Laplace-Beltrami operator on Riemannian manifolds, the Finsler-Laplace operator on Finsler manifolds etc.), Ω is an open subset in any of the above spaces, and $f : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ is a Carathéodory function. Assuming f has suitable growth at the origin and infinity, the question is: "how is it possible to find solutions for $(\mathcal{P})$ in the Sobolev space $W^{1,2}(\Omega)$?"

A possible answer is the following: we associate the *energy functional* $E : W^{1,2}(\Omega) \rightarrow \mathbb{R}$ to the problem $(\mathcal{P})$ and assuming it is well defined and enough smooth (say, of class C^1), it turns out that its critical points are precisely the weak solutions of $(\mathcal{P})$. To be more precise, if $\mathcal{L} = -\Delta_g$ is the Laplace-Beltrami operator, the energy functional associated to $(\mathcal{P})$ reads as

$$E(u) = \frac{1}{2} \int_{\Omega} |\nabla_g u(x)|^2 dv_g - \int_{\Omega} \int_0^{u(x)} f(x, s) ds dv_g, \quad u \in W^{1,2}(\Omega),$$

where ∇_g and dv_g denote the gradient and canonical volume form on the Riemannian manifold (M, g) , respectively. In order to guarantee existence of such critical points we can explore variational principles as minimization techniques, Ekeland, Borwein-Preiss and Ricceri variational principles, or mountain pass and linking arguments; for multiplicity of critical points, see e.g. Pucci and Serrin. The most demanding part of the investigation is the validity of the so called Palais-Smale compactness condition where the compactness of Sobolev embeddings plays a crucial role.

The main goal of this issue to study some elliptic equations on manifolds. The main question is "How can we guarantee existence of solutions for elliptic PDEs on compact or non-compact Riemann-Finsler manifolds via variational methods?"

Prof. dr. Farkas Csaba,
Director of the research center

A handwritten signature in blue ink, consisting of a large, stylized 'F' followed by a cursive 'arkas' and a horizontal line extending to the right.