



**FAKULTA
INFORMAČNÍCH
TECHNOLOGIÍ
ČVUT V PRAZE**

**Výzkumné léto na FIT 2024 (VýLeT 2024):
Program podpory letního studentského výzkumu na
FIT ČVUT**

VYPSANÁ TÉMATA

Část 1: Základní informace k přihlášení do programu VýLeT 2024

Tento dokument obsahuje **soupis výzkumných témat vypsanych v rámci programu VýLeT pro rok 2024.**

Podrobné informace k programu VýLeT 2024 a pravidla programu jsou uvedeny v samostatném dokumentu „Výzkumné léto na FIT 2024 (VýLeT 2024): Program podpory letního studentského výzkumu na FIT ČVUT – PROPOZICE“ zveřejněném na stránce programu na webových stránkách FIT ČVUT.

Kdo se může přihlásit?

1. Přihlásit do programu se může student **bakalářského** nebo **navazujícího magisterského** programu na FIT ČVUT, a to v období stanoveném v harmonogramu.
2. Témata se dělí na volná a rezervovaná. Rezervovaná témata jsou určena pro konkrétní studenty vybrané mentorem. Zvolit si rezervované téma může pouze student, pro kterého je téma rezervováno.
3. Kterýkoliv student, který je studentem dle bodu (1) tohoto odstavce se může přihlásit na jakékoliv volné téma.
4. Jeden student se může přihlásit na více témat, vlastní přiřazení zájemců k tématům proběhne až ve výběrovém řízení.
5. Student se může přihlásit a následně řešit výzkumné téma, které je v překryvu s tématem jeho závěrečné práce.
6. Student se může přihlásit a následně řešit výzkumné téma, a posléze si toto téma zvolit jako téma závěrečné práce.
7. Během doby, kdy je program otevřen pro příjem přihlášek od studentů, se student může přihlásit s novým rezervovaným tématem, a to za následujících podmínek: student si najde mentora, dohodne se s ním a následně mentor zadá nové rezervované téma a student se na toto téma přihlásí.
8. Během doby, kdy je program otevřen pro příjem přihlášek od studentů, je možné přidávat témata pouze podle bodu (7) Není možné přidávat témata volná, neurčená pro žádného studenta.

Jak se přihlásit?

Přihlášení probíhá odesláním elektronického formuláře. Odkaz na formulář je uveden na webu FIT na stránce programu VýLeT 2024. Tento způsob přihlášení je jediný možný a jediný platný. Přihlášení prostřednictvím jiného komunikačního kanálu nebude považováno za platné.

Do kdy je třeba se přihlásit?

Je třeba se přihlásit ve lhůtě určené pro přihlašování studentů. **Tato lhůta končí 30. 4. 2024.**

Část 2: Vypsaná výzkumná témata

Testování Markovských modelů pomocí Monte Carlo simulace na grafické kartě / Testing of Markov models using Monte Carlo simulation on GPU

ID zadání / Topic ID: 4

Mentor: Ing. Martin Kohlík, Ph.D.

Mentor specialista / Mentor specialist: Ing. Jan Řezníček

Zadání: V rámci projektu bude vytvořen algoritmus pro testování Markovských modelů pomocí Monte Carlo simulace, který poběží v paralelním prostředí na grafické kartě. Výsledkem práce budou nasimulovaná data, která budou použita pro publikaci článku na konferenci. Tento výzkum navazuje na téma disertační práce mentora specialisty a také na výzkum, kterým se zabývá Laboratoř spolehlivosti.

Assignment: The project will create an algorithm for testing Markov models using Monte Carlo simulation, which will run in a parallel environment on a GPU. The result of the work will be simulated data, which will be used for the publication of the article at some conferences. This research follows up on the topic of the dissertation of a Mentor specialist and also on the research dealt with by the SafetyLab.

Výrazová síla rozšířených regulárních výrazů / Expressive power of extended/practical regular expressions

ID zadání / Topic ID: 9

Mentor: Ing. Ondřej Guth, Ph.D.

Zadání: Výrazová síla regulárních výrazů jako teoretického konceptu odpovídá regulárním jazykům. Vyhledávací nástroje a knihovny implementují syntaktická rozšíření poskytující jednoduchou možnost vyjádření toho, co je potřeba v praxi vyhledávat. Některé syntaxe rozšířených regulárních výrazů, jako UNIX BRE nebo PCRE, dokáží popsat i neregulární jazyky. Příklady takových rozšíření jsou zpětné reference, balanced construct nebo volání subrutiny.

Výrazová síla, tj. co je nebo není možné popsat, patří mezi otevřené problémy pro některé syntaxe.

Cílem tohoto projektu je nalezení výrazové síly vybrané kombinace syntaktických rozšíření použité v některé existující/praktické implementaci.

Assignment: Regular expressions as a theoretical concept have expressive power equivalent to regular languages. Regular expression engines employ syntactical extensions to provide simple, expressive support for real-world needs. Some extended regular expression flavours, such as UNIX BRE or PCRE, are capable of describing even non-regular languages. Examples of such syntactical extensions are backreference, balanced construct, or subroutine call.

Expressive power, i.e., what can or cannot be matched, is among open problems for some flavours.

The aim of this project is to find expressive power of selected combination of syntactical features of some real-world flavour.

Music Knowledge Graph

ID zadání / Topic ID: 10

Mentor: Ing. Milan Dojčinovski, Ph.D.

Assignment: The goal of this research is to develop a knowledge graph for the music domain. It is required to: identify data sources, collect/extract information, integrate the data, model the information as Knowledge Graph using technologies such as RDF, publish the Music Knowledge Graph, evaluate the create knowledge graph.

Generating Frontend Applications from Models

ID zadání / Topic ID: 12

Mentor: Ing. Marek Suchánek, Ph.D. et PhD

Assignment: This research project explores the potential of utilizing Elm, a functional programming language, to automatically generate intuitive and user-friendly frontend applications. By leveraging Elm's robust type system and declarative approach, the study aims to enhance frontend development with a focus on usability and maintainability. Through the integration of code templates, the project seeks to expedite development processes while ensuring consistency and adherence to best practices.

The project encompasses several key components: analysis and research to understand Elm's advantages and challenges in frontend application generation, design of a comprehensive framework for user-friendly application development, prototype implementation in Elm emphasizing its declarative nature, testing to ensure reliability and user-friendliness, evaluation of generated applications, and identification of future research directions such as extending support for additional interface elements and exploring other functional programming languages for frontend development.

Generating Haskell backend web applications / Generating Haskell backend web applications

ID zadání / Topic ID: 13

Mentor: Ing. Marek Suchánek, Ph.D. et PhD

Assignment: This research project explores the potential of Haskell, a functional programming language, for automatically generating RESTful APIs to meet the demand for scalable and maintainable software systems. Leveraging the principles of Normalized Systems, the study aims to ensure consistency and adaptability within the generated APIs, while integrating code templates to streamline development processes and enhance code reusability.

The project comprises several key components: an analysis and research phase to assess Haskell's suitability for REST API generation, a design phase to develop a comprehensive framework adhering to Normalized Systems principles, prototype implementation translating theoretical designs into functional code, rigorous testing to ensure reliability and scalability, evaluation of the generated APIs against established criteria, and identification of future research directions for potential enhancements and expansions in functional backend development.

Towards RDF-Based Model-Driven Development

ID zadání / Topic ID: 14

Mentor: Ing. Marek Suchánek, Ph.D. et PhD

Assignment: In the domain of software engineering, the fusion of model-driven development and RDF-based methodologies has sparked significant interest. This research project sets out to construct a pioneering RDF-Based method for model-driven development of software information systems. Taking cues from the fundamental principles of Normalized Systems, this approach aims to establish a sturdy framework where RDF graphs serve as the cornerstone for housing intricate details about models and configurations. By harnessing RDF's inherent flexibility and semantic richness, the method seeks to surpass traditional constraints in software development paradigms, offering a more adaptable and scalable solution.

Central to this research endeavor is the symbiotic interplay between RDF graphs and code templates. Instead of relying solely on conventional programming languages, the proposed method advocates for a more collaborative approach wherein code templates dynamically query RDF graphs to extract pertinent information, subsequently generating customized project structures. This collaboration between RDF and code templates not only streamlines the development process but also bolsters the system's resilience to evolving requirements. Through meticulous alignment with the Normalized Systems tooling and principles, this research project strives to pioneer a paradigm shift in software engineering methodologies, heralding a future where software systems are constructed upon a foundation of semantic richness and modular efficiency.

Efficient algorithms for calculating overlaps of harmonic oscillator basis states

ID zadání / Topic ID: 17

Mentor: Ing. Daniel Langr, Ph.D.

Assignment: Study the following journal article: <https://doi.org/10.1016/j.physleta.2019.126162>. Focus on the mathematical methods and formulas described there. Get familiar with the following C++ code: https://gitlab.com/tdytrych/lisu3shell/-/blob/master/programs/tools/def2sph_HO.cpp. Match the relevant parts of the code with the mathematical formulas within the article. Rewrite the code into the form of a generic-purpose C++ library (header-only if applicable), possibly with additional interfaces for the C/Fortran and Python programming languages. Write the paper that will introduce this library to the scientific community. The expected target journal for this paper is Computer Physics Communications (ScienceDirect/Elsevier).

Paralelní násobení řídkých matic / Parallel multiplication of sparse matrices

ID zadání / Topic ID: 19

Mentor: doc. Ing. Ivan Šimeček, Ph.D.

Zadání: Práce se zabývá paralelním násobením řídkých matic a obsahuje analýzu a porovnání výkonu vlastního implementovaného algoritmu v OpenMP s existujícími algoritmy. Byly popsány již existující algoritmy a jejich postupy a optimalizace, díky kterým dokáží dosáhnout svých časů. Tyto myšlenky byly použity ve vlastní implementaci. Pro porovnání byla provedena sada měření, při kterých byly testovány různé matice ze souboru matic SuiteSparse. Naměřené výsledky byly poté analyzovány a porovnány, aby se zjistila relativní efektivita implementovaného algoritmu v porovnání s existujícími algoritmy.

Assignment: The work focuses on parallel sparse matrix multiplication and includes analysis and comparison of performance between a custom implemented algorithm in OpenMP and existing algorithms. The existing algorithms and their approaches and optimizations, which enable them to achieve their respective times, have been described. These ideas were utilized in the custom implementation. For comparison, a set of measurements was performed, testing various matrices from the SuiteSparse matrix collection. The obtained results were then analyzed and compared to determine the relative efficiency of the implemented algorithm compared to the existing algorithms.

The Erdős-Szekeres Maker-Breaker game

ID zadání / Topic ID: 20

Mentor: Arun Kumar Das, Ph.D.

Assignment: The Erdős-Szekeres Maker-Breaker game is a two-player competitive game where both players alternately place points in the plane such that no three points are colinear.

The first player (Maker) wants to obtain an empty convex polygon of a given size k such that the vertices of the polygon are chosen from these points and the second player (Breaker) wants to prevent it. The assignment is to investigate the strategies of both players. The precise question is whether Maker can win the game. Or, the Breaker can prevent Maker from winning till some finite time.

Detekce anomálií ve vysokorychlostním síťovém provozu pomocí neuronových sítí / Anomaly detection in high-speed network traffic using neural networks

ID zadání / Topic ID: 21

Mentor: Ing. Josef Koumar

Zadání: Hlavním předmětem tohoto výzkumu je návrh a vyhodnocení inovativních metod detekce anomálií využívajících neuronové sítě, přizpůsobených pro datové sady odvozené z vysokorychlostního síťového provozu. Výzkum je zaměřen na řešení složitosti a nuancí detekce a analýzy anomálií v prostředích charakterizovaných obrovským množstvím dat a rychlými přenosovými rychlostmi. Cíle a očekávané výstupy tohoto výzkumu jsou podrobně uvedeny níže.

Cílem výzkumu je návrh modelů detekce anomálií založených na neuronové síti. Což zahrnuje otestování existující či navržení nové metody detekce anomálií, která využívá hluboké učení, například konvoluční neuronové sítě (CNN) nebo rekurentní neuronové sítě (RNN). Přizpůsobit tuto existující nebo nově vzniklý model speciálně tak, aby zvládl jedinečné výzvy, které představuje vysokorychlostní síťový provoz, včetně velkého objemu, rychlosti a rozmanitosti dat. Přičemž datová sada pro experimenty bude poskytnuta.

Assignment: The main subject of this research is the design and evaluation of innovative anomaly detection methods using neural networks adapted for datasets derived from high-speed network traffic. The research is focused on solving the complexities and nuances of anomaly detection and analysis in environments characterized by huge amounts of data and fast transmission speeds. The objectives and expected outputs of this research are detailed below.

The goal of the research is to design anomaly detection models based on neural networks. This includes testing an existing or designing a new anomaly detection method that uses deep learning, such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs). Adapt this existing or emerging model specifically to handle the unique challenges presented by high-speed network traffic, including the large volume, speed, and diversity of data. The dataset captured on the high-speed ISP network for the experiments will be provided.

Neural Operators for Transfer Learning in Weather Predictions

ID zadání / Topic ID: 27

Mentor: Mgr. Petr Šimánek

Assignment:

Assignment: Neural Operators for Transfer Learning in Weather Predictions

This assignment aims to explore the potential of neural operators in enhancing transfer learning methodologies for weather prediction models. Students are expected to investigate the integration of neural operators into existing weather prediction frameworks, focusing on improving accuracy, computational efficiency, and generalizability across different geographic regions and climatic conditions.

Research Goals:

Literature Review: Conduct a comprehensive review of current models and techniques in weather prediction, focusing on the application of transfer learning.

Understanding Neural Operators: Delve into the concept of neural operators, their mechanisms, and their advantages over traditional neural network approaches in processing spatial-temporal data.

Model Integration: Propose and develop a model that incorporates neural operators into a transfer learning framework tailored for weather prediction. This model should aim to leverage learned representations from one region or set of conditions to predict weather in a different context.

Comparative Analysis: Evaluate the performance of the proposed model against existing weather prediction models. Criteria for comparison will include MSE, MAE and runtime.

Machine Learning Techniques for Neuromodulation

ID zadání / Topic ID: 28

Mentor: Mgr. Alexander Kovalenko, Ph.D.

Assignment: In collaboration with the CEITEC Institute, we apply machine learning techniques to data utilized in neuromodulation, encompassing fields such as biophysics and electrophysiology. Our focus includes both signal processing and image processing.

Interlingual Continuous Learning in Large Language Models

ID zadání / Topic ID: 29

Mentor: Mgr. Alexander Kovalenko, Ph.D.

Assignment: This research topic delves into the strategies and methodologies for implementing continuous learning mechanisms in LLMs that enable them to learn from new linguistic data across multiple languages. It explores the challenges of maintaining and enhancing the proficiency of language models in understanding, and generating text in a diverse array of languages, including low-resource languages that have traditionally been underrepresented in NLP research.

Decorrelated Neural Network Training for Parameter Efficient Learning

ID zadání / Topic ID: 30

Mentor: Mgr. Alexander Kovalenko, Ph.D.

Assignment: This research topic investigates methods for decorrelating weights in neural networks during both the training phase and initialization, aimed at enhancing the efficiency of training smaller models. By examining strategies that reduce the interdependence of model parameters, it seeks to improve learning efficiency and model performance without the need for extensive computational resources. The focus is on developing techniques that allow for more parameter-efficient learning, potentially leading to faster training times and reduced memory requirements for deploying neural networks in resource-constrained environments.

Parametrizované algoritmy pro hru Geografie / Parameterized algorithms for the Geography game

ID zadání / Topic ID: 31

Mentor: RNDr. Ondřej Suchý, Ph.D.

Zadání: Ve hře Geografie je hrací plocha dána orientovaným nebo neorientovaným grafem a počátečním vrcholem. Hráči Alice a Bob si střídavě vybírají vrchol, který předtím nebyl vybrán, a sousedí s posledním vybraným vrcholem. Souvisejícím výpočetním problémem je určit, který z hráčů má pro danou hrací plochu vítěznou strategii. Hra má mnoho verzí, z nichž významná je například zkrácená verze, kde je kromě šachovnice dáno celé číslo k a otázka zní, zda má Alice vítěznou strategii, která umožní vyhrát během k tahů.

Klasická složitost hry se značně liší i pro neorientované grafy, kde je známo, že neomezená varianta je řešitelná v polynomiálním čase, zatímco krátká varianta je PSPACE-úplná. Cílem výzkumu bude určit parametrizovanou složitost hry s ohledem na některé strukturální parametry vstupního grafu nebo jeho symetrizace v případě orientovaných grafů. To znamená buď naleznout efektivní parametrizované algoritmy, nebo ukázat obtížnost, případně včetně dolních na čas běhu na základě ETH. Mezi parametry, které budeme uvažovat, jsou např. maximální počet listů v kostře grafu, feedback edge set number, feedback vertex set number, velikost vrcholového pokrytí nebo stromová šířka či hloubka.

Assignment: In the Geography game the board is given by a directed or undirected graph and a starting vertex. Players, Alice and Bob, alternate in choosing a vertex not chosen previously and neighboring the last chosen vertex. The corresponding computational problem is to determine, given a board, which of the players has a winning strategy. The game has many versions, a prominent one being the short version, where apart from a board one is given an integer k and the question is whether Alice has a winning strategy to win within k moves.

The classical complexity of the game even for undirected graphs, where the unbounded variant is known to be solvable in polynomial time, whereas the short variant is known to be PSPACE-complete. The goal of the research will be to determine the parameterized complexity of the game with respect to some structural parameters of the input graph or its underlying undirected graph in case of directed graphs. That is, either to develop efficient parameterized algorithms or to show hardness, possibly including ETH based running time lower bounds. Some parameters to be considered are, e.g., maximum leaf number, feedback edge set number, feedback vertex set number, vertex cover number, or treewidth or treedepth.

Grafické rozhraní k balíčku PyRigi / Graphic user interface to PyRigi package

ID zadání / Topic ID: 32

Mentor: Dr. techn. Ing. Jan Legerský

Zadání: Realizace grafu je umístění jeho vrcholů v d -rozměrném prostoru. Jednou z hlavních otázek strukturální teorie pevnosti je rozhodnout, jestli je daná realizace grafu pohyblivá nebo pevná, neboli jestli ji lze spojitě deformovat při zachování vzdáleností mezi sousedními vrcholy.

Na workshopu Code of Rigidity (11.-15. března, RICAM, Linz, Rakousko) několik vědců včetně mentora začalo pracovat na balíčku v Pythonu, který implementuje poznatky ohledně realizací grafů. Balíček nazvaný PyRigi umožní určit (infinitesimální) pevnost či pohyblivost, globální pevnost, řídkost a další vlastnosti relevantní pro výzkum v oblasti strukturální teorie pevnosti.

Cílem tohoto VýLeTu je naprogramovat k balíčku PyRigi grafické rozhraní, které umožní zadávání realizací grafů, jejichž vlastnosti poté můžou být testovány pomocí PyRigi, a také vizualizaci výstupů jednotlivých algoritmů. Motivace pro tvorbu GUI je následující: mnoho vědců zabývajících se strukturální teorií pevnosti by ocenilo, kdyby mohli využívat PyRigi skrze grafické rozhraní místo terminálu či notebooku. Také při použití PyRigi v Jupyter notebooku je užitečné moci grafy kreslit stejně tak mít vizualizaci výstupů (infinitesimální pohyby, maximální pevné podgrafy...) nebo animovat pohyblivé realizace.

Očekává se, že vědecké komunita zabývajících se strukturální teorií pevnosti bude používat vytvořené rozhraní pro svůj výzkum. Mentor předvede vytvořené GUI na některém workshopu strukturální teorie pevnosti pro dosažení tohoto záměru. Rozhraní bude také popsáno v krátkém článku zahrnujícím souhrn vlastností, které bude umožňovat testovat.

Assignment: A bar-joint framework is a graph together with a realization of its vertices in the d -dimensional space. One of the main questions considered in rigidity theory is whether a given framework is flexible or rigid, namely, whether it admits a continuous deformation preserving the distances between adjacent vertices, or not.

Recently, a few researchers including the mentor have started to work on Python package (called PyRigi) implementing the knowledge of rigidity theory on bar-joint frameworks at the workshop Code of Rigidity (March 11-15, 2024, RICAM, Linz, Austria). The package PyRigi will allow one to check for (infinitesimal) rigidity and flexibility, global rigidity, sparsity and other properties that are of interest for research in rigidity theory.

The goal of this VýLeT project is to create a graphic user interface that would allow inputting frameworks whose properties can be then tested using the PyRigi package as well as to visualize outputs of the implemented algorithms. The motivation to have such GUI is the following: several people in the rigidity community would feel more comfortable to use PyRigi via GUI rather than in console/Jupyter notebook. Also for using PyRigi programmatically it is useful to be able to input frameworks by drawing as well as to visualize the outputs of algorithms (infinitesimal flexes, maximal rigid subgraphs...) or to animate motions of flexible frameworks.

It is expected that people in the rigidity community will use the created GUI for their research. In order to achieve this, the mentor will present the GUI on one of the rigidity theory workshops. Moreover, the GUI shall be described in a short paper including the summary of the properties that could be checked using it.