The PhD project proposals listed below will be considered for 2023/24 studentships available in the Department of Informatics to start 1 October 2023 or later during the 2023/24 academic year.

Please note that this list is not inclusive and potential applicants can alternatively identify and contact appropriate supervisors outlining their background and research interests or proposing their own project ideas.

The PhD projects are listed in two groups. In the first group are the projects with allocated studentships: each project in this group has one allocated studentship.

The remaining studentships will be considered for the projects listed in the second group. The number of those remaining studentships is smaller than the number of the projects in the second group. The allocation of studentships will be based on the merits of individual applications.

Applications for PhD studies in the Department of Informatics, for all listed projects as well as for other projects agreed with supervisors, are also welcome from students applying for other funding (within other studentship schemes) and from self-funded students.

Contents

Projects with allocated PhD studentships

Models and Algorithms for Resource Allocation Problems with Machine Learning Predictions

Supervisor: Dr Dimitrios Letsios

Description:

This project lies at the intersection of algorithms, computational optimization, and data science. The aim of the project is to design and analyze optimization models and algorithms for temporal resource allocation problems, e.g. electric power distribution, logistics, and production scheduling problems, arising in different application domains, including the energy sector, manufacturing and process engineering [Letsios et al. 2020]. The goal is to effectively assign resources, e.g. machine time and energy, to activities, so as to optimize performance. Solving instances of such problems may result in substantial economic benefits. Typically, future resource requirements and customer demand are not precisely known in advance, but can be predicted using data science and machine learning capabilities [Bertsimas et al. 2018]. However, these predictions are subject to errors. In this context, determining efficient algorithms for supporting and automating the resource allocation process is a challenge. To this end, prior work develops efficient algorithms and optimization models accounting for the time-varying nature and uncertainty of temporal resource allocation problems [Antoniadis et al. 2020, Letsios et al. 2021,Manish et al. 2018].

This project aims to (i) develop novel discrete optimization methods for temporal resource allocation methods and analyze their performance theoretically, (ii) suggest ways to mitigate the effect of prediction errors in the quality of the obtained solutions, and (iii) evaluate the performance of the proposed approaches numerically using real data. Prior experience on discrete optimization, approximation/online algorithms and/or integer programming will be useful.

References:

1. [Antoniadis et al. 2020] Antonios Antoniadis, Christian Coester, Marek Eliás, Adam Polak and Bertrand Simon. Online Metric Algorithms with Untrusted Predictions. International Conference on Machine Learning (ICML), 2020.
2. [Bertsimas et al. 2018] Dimitris Bertsimas, Vishal Gupta, Nathan Kallus. Data-Driven Robust Optimization. Mathematical Programming, p. 235-292, 2018.
3. [Letsios et al. 2020] Dimitrios Letsios, Radu Baltean-Lugojan, Francesco Ceccon, Miten Mistry, Johannes Wiebe, Ruth Misener. Approximation Algorithms for Process Systems Engineering. Computers and Chemical Engineering 132, 2020.
4. [Letsios et al. 2021] Dimitrios Letsios, Miten Mistry, Ruth Misener. Exact Lexicographic Scheduling and Approximate Rescheduling. European Journal of Operational Research, 2021.
5. [Manish et al. 2018] Manish Purohit and Zoya Svitkina and Ravi Kumar. Improving Online Algorithms via ML Predictions. Advances in Neural Information Processing Systems (NeurIPS), p. 9661--9670, 2018.

Data architectures with humans-in-the-loop

Supervisor: Professor Elena Simperl

Description:

Fundamental data-centric tasks such as conceptual modelling, content labelling, entity extraction and query processing are routinely realised as hybrid processes, which consist of human and algorithmic elements. Examples include any AI system that depends on large amounts of labelled data, interactive machine learning systems, but also knowledge graphs such as Yago, Wikidata, or DBpedia, which are created by people alongside a range of more or less sophisticated bots.

The projects in this category explore methodologies, computational methods and tools that go beyond the capabilities of existing AI and machine learning stacks in terms of tasks, performance and user experience. For example topics will include:

Novel methodologies and tools to create knowledge graphs, offering advanced user experiences, accessible to non-experts and using the latest tech (audio and video processing, intelligent assistants, AR and VR etc.)

Methodologies and techniques to acquire and encode common sense knowledge at scale Quality of knowledge graphs, including frameworks to define it, methods to assess and repair it, and the link between process, provenance and outcomes

New interfaces and experiences e.g. conversational agents to collect and curate knowledge and improve algorithmic performance.

Managing discussions, collaborative decision making and conflicts.

Data-driven modelling of value chains for efficient decision support using digital twins

Supervisor: Dr Partha Dutta

Keywords:

Data-Driven Modelling, Deep Learning, Digital Twins, Resilient Cyber Physical Systems, Automated Decision Making, Industry Value Chain.

Research background, scope and outcome:

Modern value chains are the critical backbone of the world economy. Value chains are complex networks of activities and interactions both within and across organizations of various types. Such activities are essential for creating the various goods, services and products necessary for the sustenance of our daily lives. Some examples are producing raw materials required for industrial manufacturing, designing of engineering products, operations and maintenance of high-value equipment, providing various end-user or customer services, among others. Although these are distinctly different activities, delivering these reliably require the effective collaboration between organizations. One of the main challenges in such collaborative work is the impact of unforeseen events that can disrupt any activity within a value chain. Against this background, digital twins (DT) offer a solution to help organizations make decisions under uncertainties to better manage value chains. DTs are models of real-world systems that can be used to simulate their behaviour for generating real-time or right-time insights about their operations. However, building effective DTs of real-world entities require replicating their behaviour reliably by capturing their parameters and constraints in detail. Doing so, however, is highly challenging because value chain entities are characterised by complex and unique properties that can vary subtly even across related entities within the same family (for example, power plants can use different technologies – steam, natural gas, or waste – for producing electricity and the DTs for the various types of power plants will be very different). Hence, building DTs by enumerating the physical system properties can be difficult to replicate (e.g., even across different but related entities as exemplified before) and scale (e.g., from simpler to larger/more complex entities).

Against this background, this PhD project aims to develop an alternative method to building digital twins to address the limitations of current DT methods. In this context, rapid instrumentation of industrial systems through IoT sensors has made operational data more readily available. Furthermore, recent progress in advanced artificial intelligence algorithms such as deep learning has created promising opportunities to build complex models of real-world systems using a data-driven approach.

More specifically, the PhD project will research current methods and limitations of developing empirical models of value-chain systems using sensor data. It will then design methods for developing behaviour models of multi-level value-chain systems by leveraging multi-class deep learning algorithms or other related frameworks, applied to sensor data. The methodology will be demonstrated by developing representative models of real-world value-chains (for example, those taken from the energy or manufacturing domains, for which adequate open-source data sets can be obtained). It is expected that the research output will contribute towards developing more resilient cyber-physical systems by enabling the design of more robust digital twins which in turn can improve value-chain decision making under uncertainty.

Imperfect Rationality and Computation

Supervisor: Professor Carmine Ventre

Description:

Algorithmic Game Theory is a research field that provides a set of tools to account for strategic reasoning in computer science. One assumption underlying much of the work in the area is, however, pretty limiting: agents need to be fully rational. This is unrealistic in many real-life scenarios; we, in fact, have empirical evidence that people often misunderstand the incentives and try to game the system even when misbehaving is against their own interest.

This project will look at novel approaches to deal with imperfect rationality, including the analysis of known systems and the design of novel ones. This will involve theoretical work (such as, mechanism design) as well as more applied approaches (such as, agent-based modelling) to get a better understanding of the strategic interactions within a population of agents with imperfect rationality.

Towards Fair/Explainable AI for Robotics

Supervisor: Dr Martim Brandao

Description:

This project will contribute to the area of "Responsible Robotics" – which focuses on a critical analysis of existing systems and practices in AI for robotics. The goal is to uncover social, ethical, and interaction issues of robot systems, and develop new methods that alleviate them. More concretely, the project will focus on one of the following topics:

1. Fairness in robot motion planning and robot vision
    a. identifying hidden values in existing motion planners, robot vision algorithms, and robotics datasets;
    b. identifying fairness concerns in robotics through user studies and critical literature/media analysis;
    c. questioning current practices in fair AI;
    d. proposing new system configurations or technical methods to alleviate issues of fairness.

2. Explainable and human-in-the-loop robot motion planning
    a. modeling human expectations and human understanding of robot motion;
    b. developing new algorithms and user interfaces for explainable and human-in-the-loop planning;
    c. conducting user studies to evaluate the effectiveness of explainable/human-in-the-loop algorithms, and to characterize issues such as automation bias.

Understanding the Complexity of Negotiations

Supervisors: Dr Alfie Abdul-Rahman & Dr Rita Borgo

Description:

A negotiated text is the product of a formal decision-making process where a text has been negotiated and drafted over a period of time. Many of the foundational texts of the modern world have not been written by individuals, by negotiated by groups of people in formal settings. For example, treaties between states such as the Universal Declaration of Human Rights or the Treaty of Versailles; or constitutions, such as the one negotiated by the American states in the Constitutional Convention of 1787.

During such negotiations, it is important for us to keep track of the delegations and their involvements to grasp their influence on the negotiation process either using techniques such as close reading, distance reading, or machine learning. Even relatively short historical documents written collectively in this way have been the product of thousands of specific proposals and decisions.

This project will apply a visual analytics approach towards the understanding of the complexity of a negotiation and the influence of the delegations during a negotiation process.

Possible research questions:

a. Developing new static and interactive visualization to assist with data discovery and insight generation in large datasets of events within interacting timelines.
b. Developing new approaches to show the evolution of complicated, technical documents over the period of months or years.
c. Developing new approaches for indexing the datasets related to the negotiation of documents, and more intuitive displays of the results.
d. Developing natural-language-based approaches to relate information captured in 'informal' archives (such as private diaries, letters, social media feeds etc.) to the formal records of a negotiation.

This project will work closely with the Quill Project, based at Oxford University:
https://www.quillproject.ne

Towards Multimodal Reasoning and Rationale Generation for Science QA

Supervisor: Professor Yulan He

Keywords: Natural Language Processing, machine reading comprehension, Question-Answering, multimodal learning, casual discovery

Description:

Recent years have seen a surge of interest in using large-scale pre-trained language models to solve the Question-answering (QA) tasks for reading comprehension. In a typical setup, a text paragraph paired with a question is fed to a generative language model such as GPT-3 [1] or UnifiedQA [2] to generate answers. However, this process is often viewed as a black box and it is difficult to understand why the pre-trained language models generate certain answers. In a recently developed ScienceQA dataset [3], each multiple-choice question is accompanied with annotations of the correct answer with corresponding lectures and explanations. This makes it possible to develop QA models which can also generate relationales explaining its model decisions. In addition, some of the questions contain image context in addition to text context. As such, it is desirable to build multimodal reasoning into the QA models.

This project aims to develop a QA model to address the challenging task of answering questions in the Science domain. Apart from developing advanced approaches for multimodal reasoning and rationale generation, it will also investigate the use of causal models [4] to understand the internal working mechanism of QA models built on large-scale pre-trained language models.

References:

1. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A. and Agarwal, S., 2020. Language models are few-shot learners. Advances in neural information processing systems, 33, pp.1877-1901.
2. Khashabi, D., Min, S., Khot, T., Sabharwal, A., Tafjord, O., Clark, P. and Hajishirzi, H., 2020. UnifiedQA: Crossing format boundaries with a single QA system. In Findings of the Association for Computational Linguistics (EMNLP), pages 1896–1907.
3. Lu, P., Mishra, S., Xia, T., Qiu, L., Chang, K.W., Zhu, S.C., Tafjord, O., Clark, P. and Kalyan, A., 2022. Learn to explain: Multimodal reasoning via thought chains for science question answering. arXiv preprint arXiv:2209.09513.
4. Stolfo, A., Jin, Z., Shridhar, K., Schölkopf, B. and Sachan, M., 2022. A Causal Framework to Quantify the Robustness of Mathematical Reasoning with Language Models. arXiv preprint arXiv:2210.12023.

Event-Centric Natural Language Understanding

Supervisor: Professor Yulan He

Keywords: Natural Language Processing, event extraction, machine reading comprehension

Description:

In human reading, successful reading comprehension depends on the construction of an event structure that represents what is happening in text, often referred to as the 'situation model' in cognitive psychology. The situation model also involves the integration of prior knowledge with information presented in text for reasoning and inference. Fine-tuning pre-trained language models for reading comprehension does not help in building such effective cognitive models of text and comprehension suffers as a result. Generally speaking, language understanding requires the combination of relevant evidence, whether they come from contextual knowledge, common-sense or world knowledge, to infer the meaning underneath. It also requires a constant update of a memory as reading progresses.

Possible projects under this topic are:

(1) event extraction and representation learning;
(2) event semantic relation detection and event-centric QA;
(3) event graph construction and learning;
(4) storyline generation from related events;
(5) scientific hypothesis generation from event graphs built on scientific literature.

References:

1. Liu, K., Chen, Y., Liu, J., Zuo, X. and Zhao, J., 2020. Extracting events and their relations from texts: A survey on recent research progress and challenges. AI Open, 1, pp.22-39.
2. Lyu, Q., Zhang, H., Sulem, E. and Roth, D., 2021, August. Zero-shot event extraction via transfer learning: Challenges and insights. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers) (pp. 322-332).
3. Zheng, J., Cai, F. and Chen, H., 2020, July. Incorporating scenario knowledge into a unified fine-tuning architecture for event representation. In Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval (pp. 249-258).
4. Wang, X., Wang, Z., Han, X., Jiang, W., Han, R., Liu, Z., Li, J., Li, P., Lin, Y. and Zhou, J., 2020. MAVEN: A massive general domain event detection dataset. arXiv preprint arXiv:2004.13590.
5. Zhao, X., Wang, C., Jin, P., Zhang, H., Yang, C. and Li, B., 2021, October. Post2Story: Automatically Generating Storylines from Microblogging Platforms. In Proceedings of the 29th ACM International Conference on Multimedia (pp. 2786-2788).
6. Wilson, S.J., Wilkins, A.D., Holt, M.V., Choi, B.K., Konecki, D., Lin, C.H., Koire, A., Chen, Y., Kim, S.Y., Wang, Y. and Wastuwidyaningtyas, B.D., 2018. Automated literature mining and hypothesis generation through a network of Medical Subject Headings. bioRxiv, p.403667.

Equitable Privacy and Security

Supervisor: Dr Kovila P.L. Coopamootoo

Description:

Although a spectrum of privacy and security advice and protection mechanisms are available online, their accessibility and effectiveness vary across diverse user groups. The consequence is that particular user groups have poorer online privacy, security and safety outcomes, in addition to being specifically targeted. Projects listed below can employ a mix of interviews, focus groups, participatory methods within communities or large surveys.

Example project 1: Investigate interaction dynamics

This project involves deeply investigating the human aspect that determine differential access and outcome (choosing one aspect such as gender, ethnicity, socio-economic status, age), or one / a group of privacy and security technologies. This could include, among other things, looking into the interaction dynamics.

Example project 2: Co-design accessible solutions

This project involves (1) investigating why particular advice and privacy/security HCI do not work for particular user groups (e.g. what type of digital skills are required, what is the mismatch between user needs and protection information provided by technologies, language used, access routes) and (2) how to make them accessible and useful for these user groups.

Example project 3: Gender stereotypes in privacy / security designs

This project involves investigating the gender stereotypes that play out through privacy and security behaviours and evaluating whether / how current privacy and security human-computer interactions reinforce these stereotypes.

Causal machine learning for bioinformatics

Supervisor: Dr. David Watson

Description:

Biological systems are fundamentally causal, full of interdependencies that govern response to disease and treatment. Yet standard supervised learning algorithms – e.g., neural networks and random forests – are based on empirical risk minimisation (ERM), a strategy that explicitly prioritizes correlation over causation. The result is a black box model that may succeed at labelling new samples but fails to advance our understanding of the underlying process. This is unsatisfying in systems biology, where the goal is to map complex interactions between molecular phenomena. The implications for clinical practice are profound. When variables are interconnected – for instance, when one gene regulates another – only causal models can reliably predict the effects of interventions. Answering "what-if?" questions about potential outcomes under alternative treatments is crucial in clinical medicine, yet ERM algorithms are ill-suited to this task. We will extend recent work on machine learning for causal inference, combining biological knowledge with data-driven models to separate correlation from causation at varying levels of molecular detail.

Primary objectives include developing novel methods for:

(1) regulatory network inference within and across genomic platforms;
(2) causal representation learning, mapping high-dimensional data to low-dimensional manifolds;
(3) and causal imputation, predicting effects of unseen interventions to guide drug design.

Please note: Prior experience with bioinformatic data is preferred but not required.

The complexity of database query evaluation

Supervisor: Hubie Chen

Description:

Evaluating queries on databases is a primary means of extracting information from databases. This project aims to study facets of the complexity of and algorithms for database query evaluation. Our study will draw on tools, techniques, and ideas from areas such as decomposition methods, graph theory, logic, finite model theory, database theory, and parameterized complexity theory.

For this project, a **strong interest** in and **background** in *mathematical aspects of computing* is **expected**.

References:

1. Refer to these articles particularly those appearing in the venues PODS, ICDT, and LICS.

Efficiency Bias for AI Systems

Supervisor: Jie Zhang

Description:

Existing fairness metrics of AI systems focus on the differences in functional property performance (e.g., accuracy in classification and word error rate in speech recognition) among different groups of people. However, efficiency-critical AI systems such as machine translation and speech recognition systems may take a longer time to produce outputs for users in non-privileged groups, which affects the user experience for certain groups of people. This project studies and mitigates efficiency bias in efficiency-critical AI systems.

WP1: A large-scale study of efficiency fairness in efficiency-critical AI systems. This package aims to understand the fairness of existing efficiency-critical AI systems, such as speech recognition models, in terms of efficiency through a large-scale empirical study.

WP2: Efficiency fairness improvement using hierarchical learning. This package first classifies inputs in training data into different categories. The data in each category are trained respectively. The result can be ensembled with the general big model without hierarchical learning, to reduce the negative influence brought by possible overfitting.

Related Work:

1. Sarı, Leda, Mark Hasegawa-Johnson, and Chang D. Yoo. "Counterfactually fair automatic speech recognition." IEEE/ACM Transactions on Audio, Speech, and Language Processing 29 (2021): 3515-3525.
2. Chen, Zhenpeng, Jie M. Zhang, Max Hort, Federica Sarro, and Mark Harman. "Fairness Testing: A Comprehensive Survey and Analysis of Trends." arXiv preprint arXiv:2207.10223 (2022).
3. Hort, Max, Zhenpeng Chen, Jie M. Zhang, Federica Sarro, and Mark Harman. "Bia Mitigation for Machine Learning Classifiers: A Comprehensive Survey." arXiv preprint arXiv:2207.07068 (2022)

Support for programming languages

Supervisor: Dr. Stephen Kell

One studentship is allocated to the following four projects.

Project 1. Improved debugging support for high- and low-level languages

Description:

Interactive source-level debuggers must map between program states at the binary machine level and those at the source level. Commonly in compiler toolchains (e.g. for C, C++, Rust) this mapping is defined by metadata generated by compilers on a 'best-effort' basis: compilers include large numbersof complex optimisation passes transform code and should transform metadata to match, but in practice cut corners and leave the output metadata approximate. This causes debug-time problems, such as 'variable optimized out'. Language virtual machines sidestep the problem by offering debugging over only unoptimised code but perform feats of 'decompilation' to transition state once a debugger is attached at a known 'safepoint'. This is slowing down the program being debugged but can generally offer an accurate and complete debugging experience. This PhD project will complement an existing EPSRC-funded project on testing, synthesis and expressiveness gains in debugging information. The focus will be 'best of both' approaches to debugging, possibly by exploring how to 'residualise' state that a toolchainoptimised program would not normally maintain, allowing a debugger to 'fill in' missing state. Other aspects of interest including measuring the completeness and accuracy of the metadata (enabling compilers to compete on debuggability) and improving the stack coverage of debugging metadata (enabling its use in precise garbage collection).

Project 2. Runtime(s) for multi-language programming

Description:

It is usually difficult to combine code written in different languages, especially one or more higher-level languages involving a garbage collector. Typically, programmers must write onerous 'binding' or 'foreign function interfacing' (FFI) code, mapping between C and some higher-level language implementation, but the resulting maintenance burden is high, and the effort involved does not scale across many languages and many codebases. This project will exploit recent work adding run-time type information to native code, which changes the game between high-level language implementations and low-level code, making FFI logic plausibly unnecessary. A proof-ofconcept version of such an FFIless system has already been created as a CPython module offering seamless interfacing between Python and C, published at the VMIL workshop in 2019.

The PhD will address the challenges of a 'full-blown' approach, addressing one or more of the following issues: performance, garbage collection, support for tools such as debuggers and profilers, safety guarantees, and support for additional/multiple languages.

The project **requires** *strong systems programming skills*

Project 3. Visual live programming for scientists

Description:

Currently, working interactively with data means either using a pre-built application offering a fixed interface, which is visual but offers limited programmability, or using custom workflows built by programming/scripting or command-line wizardry, which are flexible but technically demanding and far less visual. Computational notebooks like Jupyter are in some senses a third way: they are somewhat visual and have proven approachable by those seeking to learn programming 'on the job'. However, they currently suffer many usability and reproducibility issues, and still present a 'walled garden' environment with poor integration into the surrounding system.

This PhD is about ways to combine the interactivity of applications and the flexibility of command lines, possibly by designing a notebook system that works differently than Jupyter et al. We observe that crude operating system (OS) interfaces are the bottleneck since they lack a rich data model on which to build visualisation as a system-wide service. This project will work relatively low in the stack to create demonstrate that is highly compatibleand interoperable, e.g. dealing in Unix-style files of existing formats, but can support working visually and programmatically via a palette of small, composable, user-tailorable graphical tools. Target audiences include computational scientists, data scientists, digital artists and the like; we have links to interested groups of computational scientists.

The project **requires** *systems programming skills* and **an interest** in *human–computer interaction topics.*

Project 4. Software performance and induced demand

Description:

Tradition has it that computing resources are scarce and therefore that efficiency is a prime concern of programs, especially in infrastructure programs such as compilers. However, in the modern world this is no longer true: computing resources are plentiful and powerful, and most software's functionality rarely challenges the limits of the hardware. Despite this, users continue to experience software that is slow, and continue to replace hardware with newer hardware to 'keep up'—especially in the era of continuously updated web-based software. It has long been observed that everyday software is getting slower (e.g., the famous Wirth's law). One theory to explain this is 'induced demand', where greater capacity changes habits of programmers and users in ways that effectively 'soak up' the extra capacity and, often, worsen the apparent infrastructure shortfall. (One classic text on induced demand is Hart & Spivak's 'The Elephant in the Bedroom', 1993.)

This project will study the phenomenon of induced demand in commodity software stacks. It will most likely consist of developing novel profiling tools to study the evolution of the performance of commodity software, and of case studies that pinpoint technical decisions or changes which explain the loss of performance. One possible angle is to study open-source desktop software over the period from the mid-1990s to the present; one tool-building tactic would be to exploit how a single Linux kernel can host user software environments spanning a large interval of time.

Model-based Security of Medical Cyber-Physical Systems

Supervisor: Dr Nicola Paoletti

Description:

Many medical conditions require therapy via implantable and wearable devices, such as cardiac devices to treat arrhythmia treatment and artificial pancreas systems for glucose regulation in diabetes. Such medical cyber-physical systems (medCPSs) have experienced dramatic technological advancements, and include control algorithms for automated therapy delivery, internet connectivity for remote patient monitoring, and machine learning (ML) to aid therapy decisions. This complexity introduces broad attack surfaces that can jeopardize patient safety. While prior work mainly focused on the practical feasibility of the attacks, we aim to investigate sophisticated sensor spoofing attacks that are both stealthy and tailored to the target patient.

Objectives:

The project aims to developing a model-based framework to provide verified defenses against stealthy attacks on medCPSs, which you will apply to the ICD (Implantable Cardioverter Defibrillator) and artificial pancreas case studies. In particular, you will carry out research on synthesis of Pareto-optimal attacks (and corresponding defenses), logic-based formal verification of the defenses, personalization of attacks and defenses from physiological signals, and adversarial robustness of ML vs non-ML device controllers.

Prior experience in any *of control, cyber-physical systems, healthcare applications, reinforcement learning, generative models, security of machine learning, multi-objective optimization* will be a **plus**.

References:

1. Paoletti, Nicola, et al. "Data-driven robust control for a closed-loop artificial pancreas." IEEE/ACM transactions on computational biology and bioinformatics 17.6 (2019): 1981-1993.
2. Paoletti, Nicola, et al. "Synthesizing stealthy reprogramming attacks on cardiac devices." Proceedings of the 10th ACM/IEEE International Conference on Cyber-Physical Systems. 2019.
3. Rushanan, Michael, et al. "Sok: Security and privacy in implantable medical devices and body area networks." 2014 IEEE symposium on security and privacy. IEEE, 2014.
4. Kune, Denis Foo, et al. "Ghost talk: Mitigating EMI signal injection attacks against analog sensors." 2013 IEEE Symposium on Security and Privacy. IEEE, 2013.

Data science methodologies and applications.

Supervisor: Dr. Atsushi Suzuki

One studentship is allocated to the following six projects.

Project 1. Trusted data science not relying on explainability:

Description:

Trust in data science, including machine learning, will increase data science's applicability in industries. Recent complex data science methods, including deep learning, have often been criticised for their non-explainability, which some people think leads to a lack of trust in data science. However, it would be physically impossible to achieve explainability without lowering machine learning's performance if the principle behind the task is too complex for human beings to explain. This phenomenon may happen in multimedia data science, such as computer vision and natural language processing. Hence, the project seeks other approaches to achieve trusted data science. Example approaches are to provide performance guarantees of complex data science methods using statistics and information theory.

For Consideration: https://ash-suzuki.github.io/opportunities.html

Project 2. Understanding the principle behind multimedia data processing by analysing data science methods

Description:

Processing complex multimedia data, such as images, videos, acoustic signals, and natural languages, is necessary for daily human life, and human brains successfully do it. However, it has yet to be clarified how human brains process multimedia data. Compared to human brains, whose algorithms are far from formulated, recent successful data science methods, such as deep learning, are far clearer in that their algorithms are mathematically formulated, though they are still complicated. This project aims to understand the principle behind multimedia data processing by analysing existing experimentally successful data science methods. For example, students could explore the trained parameters of successful deep-learning models.

For Consideration: https://ash-suzuki.github.io/opportunities.html

Project 3: Theory of fairness in data science

Description:

In the data science area, including machine learning, fairness refers to a method's being independent of sensitive variables, which may include but are not limited to, languages, ethnicity, disability, gender, and sexual orientation. Society requires data science methods to have fairness. Hence, we are interested in data science methods' behaviour on the condition where fairness is achieved. However, existing theories to analyse data science methods have not sufficiently considered the fairness factor.

This project aims to study the behaviours of data science methods with fairness theoretically. Students might use statistics and information theory to solve the problem.

For Consideration: https://ash-suzuki.github.io/opportunities.html

Project 4. Small data science for medicine

Description:

Data science, including machine learning, has successfully contributed to medical diagnosis. Still, medical diagnosis is challenging. One significant reason is that we cannot always collect big data in medical applications for, e.g., privacy reasons. Data science methods successful on big datasets do not always succeed in small data settings owing to overfitting. This project aims to enhance and analyse data science method performance on small datasets for applications in medicine. Example approaches are transfer learning and small machine learning models.

For Consideration: https://ash-suzuki.github.io/opportunities.html

Project 5. Enhancing financial technology (fintech) by data science

Description:

Data science, including machine learning, has successfully supported financial technology in many ways, e.g., by detecting cybersecurity issues and frauds, achieving smart trading, and supporting the industry to observe regulations. The project aims to enhance the support. Examples of research questions are the following:

- How accurately and how fast can we detect cyber-attacks and frauds?
- How much can we reduce the risk of trading and guarantee performance?
- What types of regulations can data science support us to observe and how?

For Consideration: https://ash-suzuki.github.io/opportunities.html

Project 6. Multimedia data generation by data science

Description:

Generative deep learning models, such as GANs and diffusion models, significantly improved the generation performance of multimedia data, such as images, videos, natural languages, acoustic signals, etc. The project aims to tackle novel data generation tasks using existing models. The project expects students to find novel problem settings and novel ways to use existing models rather than to improve models to achieve state-of-the-art performance in existing problem settings.

For Consideration: https://ash-suzuki.github.io/opportunities.html

Privacy and security

Supervisor: Dr. Ruba Abu-Salma https://rabu-salma.github.io/

One studentship is allocated to the following three projects.

Project 1: Designing with and for the vulnerable and marginalized

Description:

Recently, researchers have started to realize that designing digital technologies for one population in mind risks ignoring the security, privacy, and safety needs of (as well as creating concerns for) other populations. While research on understanding the needs and concerns of vulnerable and marginalized populations is evolving, with the US dominating the field, the research is still in its infancy without a clear understanding of the scale and impact of technologies on such populations. The objective of this project is to empirically study the security, privacy, and safety needs (based on the lived experiences) of vulnerable and marginalized populations like the poor, the young, and the disabled – using qualitative (e.g., interviews, focus groups, participatory design workshops) and quantitative (e.g., surveys) methods. The empirical evidence gleaned from these studies will inform the design of current and future technologies.

References:

1. Julia Bernd, Ruba Abu-Salma, Junghyun Choy, and Alisa Frik. Balancing Power Dynamics in Smart Homes: Nannies' Perspectives on How Cameras Reflect and Affect Relationships. In USENIX Symposium on Usable Privacy and Security (SOUPS), 2022.
2. Julia Bernd, Ruba Abu-Salma, and Alisa Frik. Bystanders' Privacy: The Perspectives of Nannies on Smart Home Surveillance. In USENIX Workshop on Free and Open Communications on the Internet (FOCI), 2020.
3. Julia Slupska, Selina Cho, Marissa Begonia, Ruba Abu-Salma, Nayanatara Prakash, and Mallika Balakrishnan. "They Look at Vulnerability and Use That to Abuse You": Participatory Threat Modelling with Migrant Domestic Workers. In USENIX Security Symposium (USENIX SEC), 2022.

Project 2. Exploring the needs and privacy concerns of bystanders with regard to smart home surveillance

Description:

The proposed project will focus on how the growth of smart home devices (e.g., CCTV cameras, smart speakers, smart TVs, smart toys, location trackers) as well as AR/VR devices affects the privacy of bystanders – not just the privacy of those who make the decision to own and deploy the devices – and how those effects can be mitigated in product development. Groups of people for whom bystander concerns are likely to be amplified will be interviewed/surveyed, such as domestic workers (e.g., in-home care attendants, babysitters), older adults who have safety monitors in their homes, and roommates of smart home device owners. Data analysis will focus on how smart home devices differentially affect the privacy of people with different levels of social and economic power. Based on

findings, a set of design recommendations and system controls for protecting the privacy of bystanders as well as balancing the interests of primary users and bystanders will be proposed.

References:

1. Julia Bernd, Ruba Abu-Salma, Junghyun Choy, and Alisa Frik. Balancing Power Dynamics in Smart Homes: Nannies' Perspectives on How Cameras Reflect and Affect Relationships. In USENIX Symposium on Usable Privacy and Security (SOUPS), 2022.
2. Julia Bernd, Ruba Abu-Salma, and Alisa Frik. Bystanders' Privacy: The Perspectives of Nannies on Smart Home Surveillance. In USENIX Workshop on Free and Open Communications on the Internet (FOCI), 2020.

Project 3. Designing user-centered secure communication tools supporting group chat

Description:

People around the world use group communication tools. These tools such as Facebook Messenger, Google Hangouts, Skype, and Telegram are very popular for group chatting, but they often do not include security features or hide those features beneath a complex user interface. Even when tools include secure communication features, existing research shows that users tend not to understand these features, do not turn them on when they are optional, and do not know how to use these features to protect themselves against attacks.

Despite the current lack of secure group communications (from the tools themselves or from user practices), people around the world, especially vulnerable populations, need secure multi-party communication systems. Group communications can cover a wide range of topics that may be the focus of censorship, such as religion, politics, and sexuality.

To this end, to design and build communication tools that effectively protect users, we need to understand how users perceive secure communications, and what influences their decision to adopt (or not adopt) secure tools. While the usable security community has made some progress in this area regarding one-to-one communications, there is relatively little work that explores user needs, practices, and mental models of secure communications in the context of group chat. This project aims to address this gap.

References:

1. Ruba Abu-Salma, M. Angela Sasse, Joseph Bonneau, Anastasia Danilova, Alena Naiakshina, and Matthew Smith. Obstacles to the Adoption of Secure Communication Tools. In IEEE Symposium on Security and Privacy (Oakland), San Jose, CA, USA, 2017.
2. Ruba Abu-Salma, Kat Krol, Simon Parkin, Victoria Koh, Kevin Kwan, Jazib Mahboob, Zahra Traboulsi, and M. Angela Sasse. The Security Blanket of the Chat World: An Analytic Evaluation and a User Study of Telegram. In IEEE EuroS&P Workshop on Usable Security (EuroUSEC), Paris, France, 2017.
3. Ruba Abu-Salma, Elissa M. Redmiles, Blase Ur, and Miranda Wei. Exploring User Mental Models of End-to-End Encrypted Communication Tools. In USENIX Workshop on Free and Open Communications on the Internet (FOCI), Baltimore, MD, USA, 2018.

4. Sean Oesch, Ruba Abu-Salma, Oumar Souleymane Diallo, Juliane Kramer, James Simmons, Justin Wu, and Scott Ruoti. Understanding User Perceptions of Security and Privacy for Group Chat: A Survey of Users in the US and UK. In Annual Computer Security Applications Conference (ACSAC), Austin, TX, USA, Online, 2020.
5. Omer Akgul, Ruba Abu-Salma, Wei Bai, Michelle Mazurek, Elissa M. Redmiles, and Blase Ur. From “Secure” to “Military-Grade”: Exploring the Effect of App Descriptions on User Perceptions of Secure Messaging. In ACM Workshop on Privacy in the Electronic Society (WPES), Seoul, South Korea, Online, 2021.
6. Sean Oesch, Ruba Abu-Salma, Oumar Souleymane Diallo, Juliane Kramer, James Simmons, Justin Wu, and Scott Ruoti. User Perceptions of Security and Privacy for Group Chat. In ACM Journal of Digital Threats: Research and Practice (DTRAP), 2022.

A Theory of Curriculum Learning through the lens of Information Theory

Supervisor: Dr. Matteo Leonetti

Keywords: reinforcement learning, curriculum learning, information theory.

Description:

Curriculum learning for Reinforcement Learning (RL) consists in training the RL agent on a sequence of increasingly complex tasks, to direct the learning towards gradually more complex tasks. It is common in human learning and has been shown to be beneficial to RL agents in terms of training time and quality of the learned behaviour.

Many heuristic and generative black-box algorithms have been proposed (including by my group) to create and optimise curricula. However, curriculum learning is lacking a theory of what makes a task appropriate at any stage of learning, allowing to construct such training tasks accordingly. This led to a proliferation of heuristics with variable effectiveness in practice, and generally strongly domain dependent results. A theory of curriculum learning will allow to evaluate existing methods more rigorously and direct the development of optimised curricula.

In this project, we will study, analyse, and model curriculum learning with the tools of information theory. Information theory has been used to characterise many aspects of intelligent decision making and is a promising lens through which we can study and interpret behaviour and learning. You will learn and experiment with state-of-the-art reinforcement learning, curriculum learning, and transfer learning algorithms. We will apply our findings to popular benchmarks, such as the MineRL Mindcraft challenge organised by Microsoft, and outside of simulation environments to our real-world autonomous service robot.

References:

1. Narvekar, S., Peng, B., Leonetti, M., Sinapov, J., Taylor, M.E., & Stone, P. (2020). Curriculum Learning for Reinforcement Learning Domains: A Framework and Survey. J. Mach. Learn. Res., 21, 181:1-181:50.

Safe and trusted machine learning

Supervisor: Dr Yali Du

One studentship is allocated to the following twoprojects.

Project 1. Safe Reinforcement Learning from Human Feedback

Description:

Reinforcement learning (RL) has become a new paradigm for solving complex decision making problems. However, it presents numerous safety concerns in real world decision making, such as unsafe exploration, unrealistic reward function, etc [1]. While humans understand the dangers, human involvement in the agent's learning process can be promising to boost AI safety [2,3].

Early research [2] adopted human preference as a replacement for reward signals, without considering safety and trustedness of agents. [3] uses human's guidance in a supervised learning manner; agents are asking for guidance randomly without adapting to its knowledge of the environment or task.

This project considers leveraging human feedback to build safe RL agents based on symbolised preference or abstracted states. On the one hand, symbolic feedback can be easily generated by humans and effectively applied to the agent learning phase, such as a human's binary preference on an agent's actions or policies. On the other hand, different approaches for state abstractions will be considered to build up the knowledge base of safe or dangerous behaviours, such as spatial-temporal abstractions [4]. Based on the symbolized preference and behaviour abstractions, there are several potential scenarios be explored:

a. Active parenting. Firstly, like a toddler learning to walk, Human guidance is when parents say "no" or redirect a toddler attempting something dangerous. With a parent agent that is knowledgeable of the dangerous states, it can provide guidance to an AI agent. When the AI agent attempts to go to a dangerous state, the parent agent with the knowledge of the dangerous set will forbid the AI agent to do so.
b. Active learning. Secondly, the parent agent does not proactively provide guidance to the AI agent but only helps when the AI agent asks for it. The AI agent will have two policies, one policy is for decision making, and the other policy is for generating decisions of whether it should ask parents for guidance.
c. Sharing autonomy. Explainable models can be employed to predict situations where the AI agent is not performing well. On such occasions we can take control from the agent and ask for expert/human advice. The key challenge is to achieve a balance between exhausting experts and reducing the false negative rate of prediction of unsafe situations.

References:

1. [1] Wirth C, Akrour R, Neumann G, Fürnkranz J. A survey of preference-based reinforcement learning methods. Journal of Machine Learning Research. 2017;18(136):1-46.
2. [2] Christiano PF, Leike J, Brown TB, Martic M, Legg S, Amodei D. Deep reinforcement learning from human preferences. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS), 2017, pp. 4302-4310.
3. [3] Frye C, Feige I. Parenting: Safe reinforcement learning from human input. arXiv preprint arXiv:1902.06766.

4. [4] Zahavy, T., Zrihem, N. Ben, & Mannor, S. (2016). Graying the black box: Understanding DQNs. 33rd International Conference on Machine Learning (ICML) 2016, 4, 2809–2822.

Project 2. Unifying Principles in Safe and Trusted Assistive AI

Description:

AI agents are often required to assist humans in many day-to-day tasks, such as in recommendation systems, restaurant reservation and self-driving cars [1]. As AI agents are frequently evaluated in terms of performance measures, such as human-stated rewards, many challenges are posed. Firstly, due to the involvement of multiple users, agents have to learn to strike a balance between the widely different human preferences [3]. Secondly, while it is usually assumed that humans are acting honestly in specifying their preference, such as by rewards or demonstrations, the consequence of humans mis-stating their objectives is commonly underestimated. Humans may maliciously or unintentionally mis-state their preference, leading the assistive AI agent to perform unexpected implementations. An example is the Tay chatbot from Microsoft; prankster users falsify their demonstrations and train Tay to mix the racist comments into its dialogue.

This project aims to unify many principals to achieve fairness and social welfare, towards building safe and trustworthy assistive AI agents that avoid bias and manipulation like Tay Chatbot. The human preference can be explicitly stated as 'like' or 'dislike' of the agent's performance, or implicitly stated through the demonstrations. Two popular learning paradigms can be considered, Reinforcement Learning (RL) from specified preference [1] and Apprenticeship Learning (AL) [2] with human's value implicitly expressed by their demonstrations. By reinforcement learning, agents learn to perform given tasks based on preference. By apprenticeship learning, agents observe human demonstrations (historical trajectories) that reveal human's interest and learn to perform tasks to align with human values. Example questions that can be explored:

a. Multi-objective learning: given the objectives specified either by reward or demonstrations, how can we balance the different and possibly conflicting objectives from users?
b. Manipulating the assistive learning: a famous result from social choice theory is that, a non-trivial collective decision is subject to manipulation [4], how easy is it for one or some users to change the behavior of an assistive agent? Or how can a human bias the system towards their own interest? By studying how to manipulate assistive learning, the ultimate goal is still to develop robots that can delegate multiple humans' interests fairly and correctly.

References:

1. [1] Chen, X., Du, Y., Xia, L., & Wang, J. (2021). Reinforcement recommendation with user multi-aspect preference. The Web Conference 2021 - Proceedings of the World Wide Web Conference (WWW) 2021, 425–435. https://doi.org/10.1145/3442381.3449846
2. [2] Fickinger, A., Zhuang, S., Critch, A., Hadfield-Menell, D., & Russell, S. (2020). Multi-Principal Assistance Games: Definition and Collegial Mechanisms. NeurIPS, 2020, 1–10.
3. [3] McAleer S, Lanier J, Dennis M, Baldi P, Fox R. Improving Social Welfare While Preserving Autonomy via a Pareto Mediator. arXiv preprint arXiv:2106.03927. 2021.
4. [4] Allan Gibbard. Straightforwardness of game forms with lotteries as outcomes. Econometrica: Journal of the Econometric Society, pages 595–614, 1978.

Projects for the remaining studentship

Automatic Discovery of Software Vulnerabilities using AI

Supervisor: Dr. Maher Salem

Description:

In the Software development Lifecycle (SDLC), there are a lot of vulnerabilities can be predicted and mitigated or fixed before the software going to production release. This project aims to search for vulnerabilities in the SDLC and repairs it automatically. This PhD topic can investigate the use of AI in automatically discovering vulnerabilities. Therefore, background in *software security, program analysis (static or dynamic), and AI methods* are **required**.

References:

1. Ghaffarian, Seyed Mohammad, and Hamid Reza Shahriari. "Software vulnerability analysis and discovery using machine-learning and data-mining techniques: A survey." ACM Computing Surveys (CSUR) 50, no. 4 (2017): 1-36.
2. J. Antunes, N. Neves, M. Correia, P. Verissimo and R. Neves, "Vulnerability Discovery with Attack Injection," in IEEE Transactions on Software Engineering, vol. 36, no. 3, pp. 357-370, May-June 2010, doi: 10.1109/TSE.2009.91.

AI-based Model toward Course Improvement

Supervisor: Dr. Maher Salem

Most education systems are considering major items in course improvement. These items are feedback, assessments, student backgrounds, and course contents. These items become important especially after emerging the online education approach. Traditional analysis tools are no longer able to deliver significant qualitative results and the massive increasing of data is considered another hurdle in effectively analysing these items. This PhD aims to use the AI to analyse these items and suggest improvements to the degree program and the core courses. The student **requires** knowledge on *education system structure, degree programs, qualitative data analysis, AI models, and programming*.

References:

1. E. Muuli, M. Lepp, R. Palm and P. Luik, "Automation of assessment and feedback in IT teaching from the teaching staff perspective," 2021 IEEE Frontiers in Education Conference (FIE), 2021, pp. 1-9, doi: 10.1109/FIE49875.2021.9637290.
2. Hooda, Monika, Chhavi Rana, Omdev Dahiya, Ali Rizwan, and Md Shamim Hossain. "Artificial Intelligence for Assessment and Feedback to Enhance Student Success in Higher Education." Mathematical Problems in Engineering 2022 (2022). https://doi.org/10.1155/2022/5215722

Visual recognition with minimal supervision in deep learning context

Supervisors: Dr Miaojing Shi and Dr Michael Spratling

Description:

The goal of this PhD is to study object detection/segmentation in images or video with minimal supervision. This task will be placed into a setting where only image-level annotation is provided. To begin, additional supervision such as clicks, strokes, or bounding boxes may also be assumed. Towards the end of the PhD, the student is expected to work with datasets of mixed levels of supervision, including a harder, semi-supervised setting where there are only a few image-level labels as well as a large number of unlabeled images.

Several ideas can be investigated in the context of deep learning. For instance, generative adversarial learning can be employed to either augment the dataset or bridge the predicted detections with their ground truth. Recurrent neural networks can be applied to video segmentation to localize and segment semantic parts across nearby frames. On unstructured image datasets, ideas like deep metric learning and random-walk label propagation can be extended across pairs or groups of images. Cross-category transfer learning can be a further extension.

Few-shot learning is another challenging direction to explore. After learning on a set of base classes with abundant examples, new tasks are given with only few examples of novel (unseen) classes. For such cases, the learning strategy of multi-million parameters architectures in deep learning needs to be rethought to allow the networks to squeeze out the maximum amount of information from the few available samples.

Wearable, Discreet Augmentative and Alternative Communication

Supervisor: Dr Timothy Neate

Please note that applicants to this proposal are welcome to self-fund, or apply for the studentship or K-CSC Scholarship.

Description:

Approximately 2.2 million people in the UK experience a form of communication impairment [1], including a third of stroke survivors and 2/3 children in each classroom.

Communication impairments might mean that people find it hard to convey or understand

information when they need it most. This is different for everyone, but communication impairments can affect one's reading, writing, speaking and/or listening. People with communication impairments often use AAC (Augmentative and Alternative Communication) to support them in communication, generally, via a laptop, tablet or smartphone. These, assistive devices are not always quick to access and often carry with them a stigma [2].

Wearables, such as smartwatches and smart glasses, have the potential to provide a range of sensors and modes of input/output within an unobtrusive, commonplace form factor. Wearables are discreet.

Their always-available nature, coupled with instant access to the internet and processing (e.g., recognition models) on a companion device (e.g., a smartphone), have the potential to support people with communication impairments in accessing and expressing information in a subtle and less obtrusive manner.

Building upon work supporting access with wearables [3], this PhD project will conduct co-design of wearable applications and models which can support people with communication impairments in everyday life. Using established co-design approaches with users with communication impairments (e.g. [4]) this work will develop a range of input and output approaches with consumer and potentially custom form factor wearables, working closely with end-users and evaluate them in real-world contexts.

References:

1. UKGov, "Disability prevalence estimates," 2012.
2. Phil Parette and Marcia Scherer. Assistive Technology Use and Stigma. Education and Training in Developmental Disabilities
3. Vol. 39, No. 3
4. Dhruv Jain, Hung Ngo, Pratyush Patel, Steven Goodman, Leah Findlater, and Jon Froehlich.
5. 2020. SoundWatch: Exploring Smartwatch-based Deep Learning Approaches to Support Sound Awareness for Deaf and Hard of Hearing Users. In The 22nd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '20).
6. Timothy Neate, Aikaterini Bourazeri, Abi Roper, Simone Stumpf, and Stephanie Wilson. 2019.
7. Co-Created Personas: Engaging and Empowering Users with Diverse Needs Within the Design Process. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19).

Object-Based Access: Enhancing Accessibility with Data-Driven Media

Supervisor: Dr Timothy Neate

Please note: applicants to this proposal are welcome to self-fund, or apply for the studentship or K-CSC Scholarship.

Description:

The tools by which we use to create and consume media-rich digital content such as video streaming, podcasts, TV and radio, are undergoing substantial change. The core idea behind the future media ecosystem is Object-Based Media (OBM). OBM is the practice of linking media assets, such as individually recorded audio and video, to metadata. These metadata might include anything from time-stamped information about what happened in a video, who shot the video, or details about how the media should be played on different devices [1]. Digital media are sent as a collection of objects, arranged on a user's device according to their exact needs. Vitally, this means each user's experience of creation and consumption can be different.

While OBM principals might be used to automatically change a programme's duration to meet our time requirements [2] or make a story more relatable to us by including local information based on our location [3], there are also substantial implications for accessibility. As the content creation and consumption process differs for everyone, they have the potential to be accessible to everyone.

This PhD project will work with a range of stakeholders to develop accessible alternative media formats, workflows and interaction techniques for the creation of novel interactive media experiences. Building upon prior work accessible content workflows (e.g. [4, 5]) this project will consider how the future of content creation might be made more accessible through extreme customisation of content that is different to everyone and bespoke to their needs.

References :

1. D. Varghese, P. Olivier, and T. Bartindale, "Towards participatory video 2.0," in proc ACM CHI, 2020.
2. M. Armstrong, M. Brooks, A. Churnside, M. Evans, F. Melchior, and M. Shotton, "Object-based broadcasting-curation, responsiveness and user experience," 2014.
3. S. Concannon, N. Rajan, P. Shah, D. Smith, M. Ursu, and J. D. Hook, "Brooke leave home: Designing a personalized film to support public engagement with open data," in proc. CHI, 2020.
4. T. Neate, A. Roper, S. Wilson, and J. Marshall, "Empowering expression for users with aphasia through constrained creativity," in proc. ACM CHI, pp. 1–12, 2019.
5. T. Neate, A. Roper, S. Wilson, J. Marshall, and M. Cruice, "Creatable content and tangible interaction in aphasia," in proc. ACM CHI, 2020.

Socially Supported Privacy and Security (General Pool)

Supervisor: Dr Kovila P.L. Coopamootoo

A large percentage of internet users seek privacy and security advice from family, friends, or colleagues, who can provide diverse quality of advice that (1) do not ensure the user has learnt protective skills, (2) take (away) control of the user's privacy and security, or (3) even mock them for their poor skills or to have experienced a cybercrime or online harm. Projects listed below can employ a mix of interviews, focus groups, participatory-methods within communities or large surveys.

Example project 1: Advice and skill sharing framework

This project involves developing a practical and accessible framework for end users and advisors to better engage in sharing privacy and security skills. This may include, among other things, supporting end users in advisors and their advice.

Example project 2: Community supported privacy and security

This project involves designing collaborative human-computer interactions to enable communities, social connections, or peers to collaboratively engage with privacy and security technologies.

Efficient Mechanism Design for Markets and Reallocation of Goods

Supervisor: Dr Bart de Keijzer

Description:

This project focuses on the designing computationally and economically efficient mechanisms for market and exchange platforms. On such platforms, a number of agents are present who have the intention of selling or buying items from other agents. A mechanism interacts with these agents and determines, based on this interaction, how the agents should trade and against which payments. The agents are assumed to act rationally, in the sense that they have a certain utility function which they want to optimise: For example, a natural setting would be one where agents want to maximise the total value of the items received, plus the payment that they potentially receive in compensation for losing some of their goods. The agents will interact with the mechanism in such a way that their utility is optimised, and mechanisms for such scenarios need to be designed in such a way that trade happens in an optimal way, while agents are not able to "cheat" the mechanism for their own benefit. Moreover, these mechanisms should perform their computations reasonably (and provably) fast. How to design the trading mechanism in such a way that these requirements are satisfied?

As there are many details that need to be specified in the above sketch to yield a very concrete model, this gives rise to a wide range of interesting mechanism design challenges. Different properties of the market require different mechanisms, where one can think of e.g. a static "one-shot" trading

scenario versus a scenario where agents can dynamically enter and exit the market, or indivisible versus divisible goods, shareable vs unshareable goods, etc. In this project we will work on trying to solve various challenging variants of this design problem.

This is a project in algorithmic game theory, which means that it lies in the intersection of theoretical computer science and economics. This project relates strongly to computational complexity theory, approximation algorithms, matching theory, auction theory, and (clearly) mechanism design. This is a theoretical research field which is in the lucky position of also being relevant in practice: As

examples of where this field is applied, one may think of ad-auctions in search engines, and various automated market platforms where goods are exchanged, or where clients are assigned to service providers (think of various popular platforms for taxi drivers, finding holiday accommodation, food delivery, and transportation services for goods).

Multiscale agent-based modelling of collective cell behaviour during vascular network formation

Supervisor: Dr. Katie Bentley

Description:

Simulations are becoming more widely used in biomedicine to understand how low-level interactions between cells lead to emergent collective cell behaviour during the development of tissues and organs in the body. They also offer a testbed to predict new interventions and therapeutic strategies targeting aberrant cell behaviour during disease. For example, our lab has developed highly predictive agent-based models, validated by follow-up biological experiments, where cells are modelled as autonomous agents during the process of blood vessel growth in normal development, cancer, and retinopathy e.g. [1,2].

This project will aim to investigate establishment of an efficient multiscale agent-based-model simulation, calibrated, and validated directly to a specific in vitro tissue on a chip device, where vascular networks spontaneously form under different controllable conditions – with experiments performed by our collaborators at EMBL Barcelona. The aim is for the model to enable us to study how fine-grained cell shape changes at a subcellular scale impact emergent cell behaviours and formation of the entire vascular network across the tissue device. To achieve this the project will involve a software engineering element as well as biological modelling, investigating best practice in designing the multiscale ABM to be as efficient and robust as possible in order to simulate the hundreds of cells in the device at once across scales. The model will also enable exploration of several specific hypotheses we have about combinations of different tissue factors that could be added to fine tune cell behaviour and the resulting vascular network density leading to predictions of new therapeutics to improve vascularisation in diseased tissues.

Candidate expectations: the ideal candidate will be highly proficient in C++ and principled software engineering practices. An interest or experience in biological problem solving or modelling is desirable but not essential as training can be given in all biology required for the project. Excellent communication skills to handle discussion across disciplines are essential.

References:

1. Bentley, K et al. "Do endothelial cells dream of eclectic shape?." Developmental cell 29.2 (2014): 146-158.
2. Bentley, K. "The temporal basis of angiogenesis." Royal Society PTB (2017): 20150522.

Administrative Access Control Policies

Supervisors: Professor Maribel Fernandez & Dr Jose Such

Description:

Administrative access control policies specify the rights that security administrators have in the system (e.g., to add or remove users, or change users' rights). These policies are critical to ensure the overall security of the system, but not much work has been done on the development of general models for administrative access control. In this project we aim to define formal generic models of administrative access control, based on the Category-Based Meta Model of access control (CBAC), which can be used to analyse access control systems and help identify the impact of changes made by administrators (impact change) on the overall security of the system.

“Alexa, cover your ears!”: Privacy-Aware AI Personal Assistants

Supervisor: Dr Jose Such

Description:

AI personal assistants are becoming mainstream in practice, with the widespread introduction of desktop, phone and home assistants. For instance, over 70 million users utilise smartphone assistants like Siri, Google Assistant, and Cortana every day; and smarthome assistants have been sold in massive numbers, like the five million units of Amazon Echo with the Alexa personal assistant sold in less than two years. However, recent incidents involving AI personal assistants like Alexa recording a private conversation and sending it to a random contact, have increased users' privacy concerns, with some users trashing their assistants all together and companies like Mattel cancelling assistant projects.

It is therefore paramount to consider and respect users' privacy to realise the benefits of AI personal assistants and foster trust from users. In this project, you will formalise the social norms that govern information sharing, management, collection, processing and learning in AI personal assistants. Based on this, you will design novel methods to personalise privacy in AI assistants based on the social norms but also on the users' contextual, group, and individual preferences with an optimal accuracy-intervention trade-off.

Automated Signature Generation for Network Intrusion Detection Systems (NIDS)

Supervisor: Dr Fabio Pierazzi

Description:

A Network Intrusion Detection System (NIDS) is a probe that passively monitors network traffic and triggers a “security alert” whenever a signature matching a particular pattern is found. However, signatures are still mostly generated manually, a process which is error-prone and time consuming.

This project will explore how AI and ML can support automated generations of signatures which are effective, efficient and interpretable. Evaluations will include adaptivity to different network environments, and performance speedup in terms of DR and FPR with respect to manually-defined signatures and existing statistical approaches.

Backbone Guided Local Search Methods for MAX-SAT

Supervisors: Dr Kathleen Steinhofel & Dr Dimitrios Letsios

Description:

Satisfiability (SAT) is a key problem in combinatorial optimisation and has a huge range of real-life applications. It seeks for a given Boolean Formula (conjunctive normal form) an assignment of variables such that the formula returns True. In case such an assignment does not exist, we seek an assignment that satisfies a maximum number of clauses (MAXSAT). As backbone structure, we denote the set of variables that have the same assignment in all optimal solutions.

Knowledge about the backbone structure can be used to guide heuristic methods which aim to find near optimal solutions. For instance, the size of the backbone can give indications of how many optimal solutions exists and consequently how hard it is for the search method to converge to an optimal solution. The guidance can be provided in two different types:

1. Deriving instance dependent methods by using pre-processing to approximate the backbone structure and to derive parameter settings for local search.
2. Estimating the backbone structure based on configurations visited by the local search method.

The findings will lead to faster convergence to optimum solutions and more importantly can produce methods which adapt to instance dependent properties. At the same time, methods to derive and analyse the backbone structure can be used to classify candidate solutions and to model additional, sought-after properties such as robustness of candidate solutions.

Behavioral Modeling of Process Memory for Real-Time Detection of Attacks

Supervisor: Dr Fabio Pierazzi

Description:

Memory vulnerabilities such as buffer overflow, heap spraying, heap vulnerabilities, are still one of the major threats in all modern systems. Most modern approaches to detect memory attacks are based on heavyweight monitoring and analysis which causes significant overhead and prevents real-time application. This project will explore how AI and ML can be used to create a behavioural model of process memory for real-time anomaly detection of attacks occurring in memory.

Better Error Help Using Large Scale Programmer Data

Supervisors: Professor Michael Kolling & Dr Neil Brown

Could large scale beginning programmer data be used to give useful hints and help to beginners stuck on an error? For example, if a novice had problems with a task, could perhaps useful hints be automatically generated by analysing previous users who had similar problems, what they did, and whether their actions led to solving the problem?

Making use of the Blackbox data set [1] is one option to automatically generate helpful hints and tips for novice programmers.

References:

a. https://bluej.org/blackbox/

Big Data in Programming Education

Supervisors: Professor Michael Rolling & Dr Neil Brown

Description:

The Blackbox project has collected a large amount of data about the behaviour of novice

programmers. We have data about hundreds of millions of programming sessions. So far, this data has been analysed only very superficially. An interesting project would be to use a big data approach for deeper analysis of this data set, and to work out what we could learn from this.

Characterization of Immunoglobulins

Supervisors: Professor Costas Iliopoulos, Dr Sophia Karagiannis & Dr Grigorios Loukides

Description:

Antibodies, or immunoglobulins, belong to the ‘gamma globulin' protein group and can be found mainly in the blood of vertebrates [1]. Antibodies constitute the major serological line of defense of the vertebrates with jaws (gnathostomata) by which the immune system identifies and neutralizes threatening invaders (viruses, fungi, parasites, bacteria). The contrivance underlying the reaction efficiency of our immune system to specifically recognize and fight invading organisms or to trigger an autoimmune response and disease remains to be elucidated. The efficient reaction of our immune system against all kinds of intruders is highly dependent on the number, condition and availability of antibodies, as reaction times are ‘key' to the successful elimination of the foreign pathogen. The importance of antibodies in health care and the biotechnology industry demands knowledge of their structures at high resolution. This information can be used for antibody engineering, modification of the antigens binding affinity and epitope identification of a given antibody.

Computational approaches provide a cheaper and faster alternative to the commonly used, albeit laborious and time consuming, X-ray crystallography. Available immunogenetics data can be utilized for computational modelling of antibody variable domains. Standardized amino acid positions and properties can assist in optimizing the relative orientation of light and heavy chains as well as in designing homology models that predict successful docking of antibodies with their unique antigen. As a result, it comes down to identifying conserved motifs or patterns that are implicated and mediate antibody-antigen interactions. Detection of such motifs by simple sequence comparison is impossible. Our research is fixated on the investigation of alternative approaches to efficiently study antibodies, mainly by the multimodal fusion of information from genetic, structural, and physicochemical analysis.

We propose a holistic approach in the realm of immunoinformatics that will focus on elucidating the mechanism of antibody-antigen recognition. The results and the final tool (in the form of either an online service or a downloadable tool) will be made freely available to the scientific community. We are confident that many fellow researchers from all walks of immunology, bioinformatics and antibody-related sciences will benefit from such a tool, both in terms of applied research and basic understanding of the function of CDRs.

Nowadays, it is certain that such specialized and specific recognition properties cannot be based on random and hypervariable sequences. It is just that using the 20 amino acid code is not a suitable approach to explain the phenomenon. Therefore, herein, we will calculate more than 430 different physicochemical properties to represent each residue of all antibodies, to identify what is the right dictionary (or indeed the right combination of dictionaries) required to decrypt the antibody-antigen interaction puzzle. The calculation of the physicochemical properties will be done using the QSAR module as it is implemented into the Molecular Operating Suite (CCG).

A brief overview of our main objectives includes the following:

a. Collecting and building the working dataset
b. Collect and curate antibody structural data from numerous databases
c. Deep learning for feature extraction and prediction • Predict reliable classification markers through the use of convolutional neural networks (CNNs)

Computational Analysis of Ageing Brains

Supervisors: Dr Kathleen Steinhofel & Professor Zoran Cvetkovic

Description:

The ability to acquire and store information is a key function of the brain. This ability is affected by ageing and in various age-related disease, including dementia. In old age the acquisition of new information is more difficult than in young age. Moreover, updating of acquired information is also affected by ageing. The mechanistic basis of the age-related decline is not well understood. It is known that changes at synapses, the connections between nerve cells, are the basis of information storage. But it remains unknown how the synaptic basis of information storage changes with age.

Recently, ultrastructural changes at synapses were discovered and analysed after training in a memory task in young and aged mice. In the research programme, we want to investigate the impact of these changes by using computational approaches based on models of these biological observations. In addition to the modelling of ultrastructural changes, the regulatory function and expression level of microRNA in the neuro cell will be analysed towards the impact to the ability to store information. The findings will not only advance insights into mechanisms of information storage, but also support the analysis of age-related diseases, such as dementia, that affect cognition.

The supervisor team will include Prof Peter Giese (IoPPN) and Anna Zampetaki (Cardiovascular Division).

Contextualising Big Programming Data

Supervisors: Professor Michael Kolling & Dr Neil Brown

Description:

The Blackbox project has been running for over five years. It collects data from noviceprogrammers: source code that is written, compilation errors displayed, and various other data about a programmer's interaction with the BlueJ IDE. The data is collected without any further context: we do not know the age or experience of the programmer, whether they are on a course or not, whether they are doing well on their assignments, and so on. This allows for a large data set, but one that is stripped of useful context. This project would investigate collecting useful data (e.g. experience, course grades) for a subset of Blackbox participants, to provide a richer subset for other researchers, and to be used in the project to investigate associations between programming activity and success on a course.

Data Science Strategies for Cancer Immunotherapy Application

Supervisors: Dr Sophia Tsoka & Dr Grigorios Loukides

Description:

Computational analysis of biomedical datasets can lead to understanding of disease systems and therapeutic interventions. We propose a project that will target the computational analysis of experimental data on immune activation against cancer using antibodies. Integration of experiments with publicly available data on known cellular interactions will establish a resource for data mining. Such a resource will be used to implement machine learning algorithms to link gene features to cancer response, network analyses to represent molecular interactions and logical modelling to explore regulatory effects from proteomic experiments.

The combination of these Data Science frameworks will elucidate signalling networks related to the control of tumor growth by antibody- enhanced human immune cells and identify key altered pathways and their regulation state. The long-term prospect is to improve understanding of disease mechanisms and cell signalling, to improve the design of novel drugs and therapies.

Development of Scalable General Artificial Intelligence (AI) Problem Solving Systems

Supervisor: Dr Amanda Coles

Description:

This project aims to develop scalable general Artificial Intelligence (AI) problem solving systems, capable of reasoning with the large combinatorial problems that arise in effectively managing the oversubscribed infrastructure of densely populated cities. This project builds on a study, supervised by Dr Amanda Coles (KCL Informatics) an expert in AI Planning and Professor Christopher Beck (University of Toronto) an expert in Constraint Programming (CP), exploring the application of CP and AI planning to disruption recovery in the UK rail network.

The PhD project aims to significantly increase the solution quality and scalability of AI problem solving technologies, based on our new understanding of the strengths these approaches, by automatically decomposing problems so CP solvers and AI Planners solve the parts best suited to their strengths. The successful candidate will extend the state-of-the-art in AI research and can apply this to real-world UK rail network problems.

Distributed Computing by Population Protocols

Supervisor: Professor Tomasz Radzik

Description:

Population protocols are a simple model of distributed computing, with applications extending to other areas, including processes in chemical network and online social networks. This model assumes that the computing system consists of many identical devices, called agents or nodes, which communicate with each other in pairwise interactions. The pattern of interactions depends on external factors and interacting nodes follow a simple protocol, which should ensure that all nodes gradually learn some global property of the system. This project is a study of the computational potential and limitations of this model and an investigation of applications.

Fairness in Automatic Assessment

Supervisor: Dr Zheng Yuan

Research areas: fairness in artificial intelligence, bias in artificial intelligence, artificial intelligence in education, machine learning, natural language processing, automatic assessment

Description:

Automated assessment (AA), the task of employing machine learning models to automatically score written/spoken text, is one of the most important educational applications of natural language processing. Emerged as a means to overcome issues arising with standardised assessment, AA supports a faster assessment and provides instant feedback, not only facilitating self-assessment and self-tutoring, but also addressing educational shortfalls promptly. Moreover, the potential of a reduced workload is becoming more attractive, especially in large-scale assessments. As a lot of teaching has moved online and the number of students keeps rising, AA is crucial to the scalability of teaching and marking. Further advantages become more pronounced when it comes to marking constructed responses, a task prone to an element of subjectivity. AA systems guarantee the application of constant marking criteria, thus reducing inconsistency, which may arise when more than one human examiner is employed.

Over the last few years, there has been a significant amount of work done on ensuring fairness, accountability, and transparency for machine learning models. With the deployment of AA in both summative and formative scenarios (e.g. high-stakes testing and classroom instruction, respectively), it is important to ensure fairness in these AA systems and all test-takers are treated fairly, especially for making high-stakes decisions like college admissions, employment, or visa applications. Recently, there has been increasing interest in AA fairness/bias, and research in this area has mainly focused on detecting bias in a post-hoc setting. For example, studies have documented differing performance of existing AA systems for test-takers with different gender, race, native language, socioeconomic status, or disabilities.

Research Objectives:

a. This project will study fairness and ethics in artificial intelligence (AI), with a special focus on AA. Studies in machine learning have highlighted that algorithms often introduce their own biases either due to an existing bias in the data or due to a minority group being inadequately represented. The aim of this project is to develop machine learning models with fairness and ethical considerations. As a result, the decisions made by the new systems will be unbiased and the decision-making processes will be transparent, which will eventually build up trust in AI and benefit all.
b. This project is expected to detect, understand and mitigate both algorithmic bias and data bias in machine learning models, as well as to define and measure fairness in AI systems. In particular, the project will focus on developing accountable and responsible machine learning models for AA, so as to ensure fairness in AA. However, the models and techniques produced as well as lessons learnt will be sufficiently generic such that they can be applied to other AI applications and the diverse range of contexts for AI.

References:

1. Andersen et al. Benefits of alternative evaluation methods for Automated Essay Scoring. EDM 2021.
2. Litman et al. A Fairness Evaluation of Automated Methods for Scoring Text Evidence Usage in Writing. AIED 2021.
3. Ke and Ng. Automated Essay Scoring: A Survey of the State of the Art. IJCAI 2019.
4. Madnani et. al. Building Better Open-Source Tools to Support Fairness in Automated Scoring. EthNLP@EACL 2017.
5. Romei and Ruggieri. A multidisciplinary survey on discrimination analysis. The Knowledge Engineering Review 2014.

Formal verification of smart contracts

Supervisor: Dr Hana Chockler

Description:

Formal verification of software is gaining popularity for verifying increasingly complex and safety-critical software. While the full verification task is unsolvable (the problem is easily reduced to the halting problem, which is undecidable), numerous existing solutions to subproblems are general enough to provide thorough verification and correctness assurance for real-life systems. There is a number of teams currently working on tools for software verification, with the Formal Verification Team at the University of Lugano (USI), led by Prof. Sharygina, being one of the most established ones. Prof. Sharygina recently received Swiss government funding for a large project titled "Beyond Symbolic Model Checking through Deep Modelling", in which Dr. Chockler (the first supervisor) is a named collaborator. The proposed Ph.D. project will be done in collaboration with the team at USI.

The student will be able to travel to work face-to-face with the team in Lugano, and close collaboration via skype and emails is expected when the student is in London.

Current model-checkers (automated formal verification tools) are mostly suitable to verify programs in C and C++. In this project, we will research the direction of formally verifying smart contracts. The student will research different options of extending the verification platform to smart contracts written in Solidity (or other languages) and will analyse whether the verification should be done on the source code level or on the bytecode level.

Smart contracts are typically small. However, they interact with other contracts and are being called in a loop or recursively, thus leading to a number of subtle bugs (see, for example, the exploit of the DAO bug, leading to loss of $50 Millions). It is then reasonable to expect that the best way to formally verify smart contracts is by using modular reasoning: for each smart contract, the other contracts with which it interacts can be considered an environment.

This environment can be overapproximated using learning techniques in combination with sampling and traditional model checking approaches. After verification of a single contract passes successfully, some symbolic representations of the contracts with respect to the correctness properties will be combined to prove correctness of the overall system.

The project will include a significant implementation component. The implementation is done using the software verification platform developed at USI. The main development task is the new front-end, so that the verification platform is able to analyse programs in Solidity (or EVM bytecode). As model-checkers require writing a large and complex software, the advantage of having such a software available already and being in contact with the team that develops and maintains it is hard to overestimate. In addition, being a part of a very active and experienced research team guarantees discussions and collaborations that further aid the research, especially in the initial stages.

From Requirements to Models Using Natural Language Processing

Supervisors: Dr Kevin Lano

Description:

The construction of UML models such as class diagrams can be a complex and time-consuming activity, even with tool support, and the modification and evolution of these diagrams is also challenging.

This PhD project will investigate the automated production of models from natural language requirements statements, using rule-based or neural net approaches to identify model elements such as classes and operations from the statements.

The project should involve a comparison of the relative effectiveness of rule-based versus neural net approaches and investigate how these could be combined.

The project is part of a large research programme for ``User-centered model-driven engineering" carried out within the Software Systems research group, which aims to make MDE techniques usable by mainstream software practitioners.

Human data interaction

Supervisor: Professor Elena Simperl

Description:

Technology can play an important role in improving people's experiences with data, whether in a professional context, or in everyday life. Projects in this space look at human factors that affect our ability to find, make sense and communicate with data, including topics such as:

a. Dataset search and discovery, including Google's dataset search engine
b. Data portals: how are they used and how can they be improved
c. Communicating and presenting data, metadata and data-related activities
d. User experience in data science and data engagement
e. Tools and experiences to increase accessibility of data and data science work
f. Collaboration in data science
g. Data storytelling tools with narrative support
h. Data science communities: where are they, how do they work, how can we make them better?

Example project: New interfaces and experiences to data engagement

The project will explore novel ways to present a dataset, for example a CSV file, using speech, audio or video technologies. The aim is to propose an algorithm that given a dataset produces a media summary of the content and context of the data and evaluate the results in a user study. The algorithm could use a range of techniques, including machine learning, computer vision and speech generation.

This will also require capabilities to generate text from data, as text is more accessible that metadata to convey what a dataset is about and how it should be used. In previous studies we used a manual approach to create summaries, which does not scale. The aim here would be to use natural language generation to automatically create short text summaries for a given tabular dataset, formatted, for example, in CSV. The project could use machine learning or rule-based techniques.

Example project: Personalising dataset search

In previous studies we explored different ways to present datasets in the context of search, including structured metadata, text descriptions, data previews and visualisations. In this project, the aim is to develop an information retrieval algorithm that tests the impact of these different result presentations and personalizes them based on user preferences and feedback.

This could include, among other things, personalised analogies for numerical data. Research has shown that using familiar concepts to describe numbers and numerical datasets can improve engagement. The aim of this project is to explore the same approach for a wider range of datasets (beyond spatial data such as distances and areas) and to develop an approach that for a given numerical dataset learns to recommend relevant analogies.

Example project: Communicating data quality

Most work in data visualisation has focused on choosing and customising charts and stories to communicate data. The aim of this project is to look into contextual aspects of data use, including sources, uncertainty, missing or incorrect values, timeliness and the way this additional information

could be embedded into visual design. The project will first undertake a survey of existing approaches for numerical datasets and then propose and test ways to communicate less explored quality aspects.

Human and social factors in information systems

Supervisor: Professor Elena Simperl

Description:

Some of most remarkable online platforms and tools we are using today bring together human and social intelligence with data and algorithms in ingenious ways. Underlying them, there is a huge, interdisciplinary research space concerned with the design principles, methods and tools that allow us to build such systems and understand and predict their evolution. The most successful of them seem to share a core set of principles:

a. They are decentralised and self-organizing, and can mobilize a critical mass of resources effectively, whether that's people, data or computational devices.
b. They make extensive use of mobile, sensor and web platforms, alongside openly available data and software to enable communication, knowledge exchange and coordinated action.
c. They know how to bring crowd and machine capabilities together to achieve their aims in a sustainable way.
d. They empower individuals to self-organise and commit to being fair, transparent and accountable about the data and resources these contribute.

Relevant topics include applications of crowdsourcing and social computing to AI systems, as well as fundamental crowdsourcing research around task and workflow design, crowd learning, quality assurance, and ethical crowdsourcing. The research would potentially focus on a class of social machines, including peer-production systems, human computation platforms and participatory sensing networks.

Example project: Improving task design

There is a large body of literature exploring how to achieve a particular goal via crowdsourcing and proposing workflows and improvements. The aim of this project is to derive such task design guidelines from a new source: discussion forums used by the crowd, for example on Mechanical Turk or in citizen science projects on the Zooniverse platform. The project will collect a sample of relevant discussions and extract comments pertinent to design guidelines, using, for instance, quantitative (NLP) or qualitative techniques.

Example project: Crowd self-assessment

In crowdsourcing, asking participants to self-assess their skills and performance helps designers understand the feasibility of the task and identify areas of improvement. Previous research has looked at the ability of crowd participants to self-assess. The aim here would be to carry out a follow-up study to understand whether the initial conclusions apply to other tasks, domains, and workflows.

Learning of Software Design Patterns

Supervisor: Dr Hana Chockler

Description:

A software engineer joining a development team typically does not start writing software immediately; first, she needs to understand the large existing body of code and recognise the key components and how they interact with each other. Documentation is typically sparse and not updated regularly. The software, on the other hand, is large and difficult to understand. “Software is like entropy: It is difficult to grasp, weighs nothing, and always increases.” (Norman Augustine). Project development tools aid understanding the software by identifying the participating classes, and the static dependencies between them.

The next step is to identify a set of design patterns common to this project, e.g.: when are resources allocated and freed, in what manner are certain components of a class visited? In this area, the static analysis tools are insufficient. The proposed project is to learn the design patterns in the given code automatically by applying grammatical inference (learning) algorithms.

The benefits for automatically learning design patterns go beyond helping the software engineer getting a clear representation of the design patterns. An automatic analysis can discover areas where the same goal is achieved by utilising different patterns: one of the patterns can be erroneous or

obsolete, or the multitude of patterns can point to the lack of precise development guidelines for a certain task, indicating a need for a guiding design pattern. These challenges require developing learning algorithms that can learn several automata simultaneously, such that the resulting automata correctly capture the main abstractions in the given corpus of code.

The project will build upon previous research results of Dr Hana Chockler in collaboration with the University of Oxford.

Machine Learning Augmented Algorithms

Supervisor: Dr Frederik Mallmann-Trenn

Description:

Machine learning and in particular deep learning has gained much attention over the past several years, yet theoretical understanding is still very limited. As a remedy, a recent line of research emerged in which neural networks are used as a black box in many online problems, where the data arrives over time. The idea is to use a neural network to give predictions of the data that will arrive in the stream. The goal is to design algorithms that perform much better than previously if the prediction is good and on the other side, to show that even if the prediction is bad, the solution found by the algorithms is still reasonably good.

A toy example is the ski rental problem, where each day a skier on vacation has to make a decision: either rent skis for 10$ or buy skis for 100$. We assume that the ski trip can end abruptly (chosen adversarially). See https://en.wikipedia.org/wiki/Ski_rental_problem for classical algorithms. Now if we assume that a neural network makes a prediction on when the ski trip will end, how much better can we do? The field is very young and great problems of practical and theoretical importance await!

Related literature:

(1) http://www14.in.tum.de/personen/albers/papers/inter.pdf for a technical survey on online algorithms.

Model Driven Engineering in Finance

Supervisors: Dr Kevin Lano

Description:

In the finance industry there is a strong emphasis on the rapid time-to-market of new financial software products and financial models, which can conflict with the achievement of software quality and correctness. The proposed research will investigate how these conflicting aspects can be managed and partly resolved through, for example, the reuse of trusted components, and the use of model-based rapid application development and iterative (agile) development.

Modelling Predictive Space-Time Cube for Urban Informatics

Supervisors: Dr Yijing Li, Prof. Nicolas Holliman

Description:

This project will build up the space-time cube(s) predictive model for urban information on multiple dimensions, e.g., greenspace accessibility and values, land-use simulated mobility, residents' happiness and geodemographic profiles, and the development of local crimes, in the expectation to enlighten policy makers with data-driven evidence. The Predictive Space-Time Cube model will be trained and tested with multi-sourced trajectory open data (for example, remote sensing images, census data, google mobility data, detailed crime incidents data, statistics on socio-economic, etc.) in selected metropolitan cities like London, New York, Sydney, and Hong Kong (https://comparecitycrime.com/, preliminary exploration). Besides of the widely applied spatial data analytical skills and machine learning techniques, student will develop a 3D understanding of the urban crimes in a dynamic and forecasting way and contribute to the tradition literature on spatial analysis from an innovated angle by adding the dynamic temporal and layers' dimensions.

Modular and Hierarchical Learning and Representation of Large Software

Supervisor: Dr Hana Chockler

Description:

A software engineer joining a development team typically does not start writing software immediately; first, she needs to understand the large existing body of code and recognise the key components and how they interact with each other. Documentation is typically sparse and not updated regularly. The software, on the other hand, is large and difficult to understand. “Software is like entropy: It is difficult to grasp, weighs nothing, and always increases.” (Norman Augustine).

Being able to represent large software in a graphical way with the ability to zoom in and out of components would help tremendously towards understanding of the software structure and its functionality. The proposed project is to learn a hierarchical compositional structure that will be used for such a graphical representation. The most likely candidate for such a hierarchical structure is state charts which have been used for software design for many years.

There are no existing learning algorithms for learning state charts. There is, however, a number of algorithms for learning similar structures, such as finite automata, transducers, etc. The first part of the proposed project consists of constructing a new learning algorithm for learning state charts. The second part of the proposed project is using this algorithm to learn complex software in a hierarchical way, allowing the user to zoom in and out of components (composite states).

The project will build upon previous research results of Dr Hana Chockler in collaboration with the University of Oxford.

Multiple Robots Performing Random Walks

Supervisor: Dr Frederik Mallmann-Trenn

Description:

The goal of the project is to study distributed algorithms for dynamic and noisy settings robot swarms, and biological systems. We will seek new algorithms to solve fundamental problems of communication, construction, reaching agreement, estimation, data processing, searching, shape formation, task allocation, and more.

One example is the setting of [1], where robots must estimate the fraction of black tiles in a grid.

Each of the robots is very simple and performs a random walk. Whenever, two or more robots are close to each other, they can communicate with each other. In the end, the robots must agree on a joint estimate of the fraction of black tiles.

The arising questions here are:

a. How much can multiple random walks speed up the process?
b. How many samples have to be taken?
c. What happens if the communication is noisy?

The goal is also to collaborate with researchers in the robotics community by modelling and analyzing systems theoretically. In addition to a solid understanding of Markov chains, students should be interested in collaborating with researchers across different disciplines.

References:

(1) [1] https://dl.acm.org/citation.cfm?id=3237953

Natural language explanations for artificial intelligence

Supervisor: Dr Zheng Yuan

Research areas: artificial intelligence, deep learning, explainable artificial intelligence, natural language processing

Description:

In recent years, artificial intelligence (AI) has been successfully applied to various applications with the breakthrough of deep learning (DL). Despite the impressive performance, the decision-making processes of DL models are still generally not transparent or interpretable to humans due to their 'black-box' nature.

Explainability is becoming an inevitable part of machine learning systems. This is especially important in domains like healthcare, education, finance and law where it is crucial to understand the decisions made by AI systems and build up trust in AI. Several directions for explainable artificial intelligence (XAI) have been explored and the majority of explainability methods focus on providing explanations at the input feature level, which consist of assessing the importance or contribution of each input feature, after the models have been trained and fixed. However, these methods may 1) fail to provide human-readable explanations as the underlying features used by AI models can be hard to

comprehend even by expert users (e.g. tokens for text and pixels for images); 2) only detect the incorrect learned behaviour, without providing any general solution for improvement.

As an appealing new research direction, this project will focus on generating human-friendly and comprehensive natural language explanations (NLEs) for AI, where NLEs normally consist of natural language sentences that provide human-like arguments supporting a decision or prediction. In particular, the aim of this project is to develop AI models that can make use of NLEs to provide better performance, counteract existing biases in the data, and provide human-readable explanations for the decisions made by the models. The AI models produced will have the advantage of making use of explanations and providing human-level explanations, just like how humans both learn from explanations and explain their decisions.

The project will focus on natural language processing as the primary application area and start from publicly available NLEs datasets. However, the AI models and techniques produced will be sufficiently generic such that they can be applied to other areas, such as computer vision, speech processing, policy learning, and planning.

References:

1. Arrieta et. al. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion 2020.
2. Camburu et. al. e-SNLI: Natural Language Inference with Natural Language Explanations. NeurIPS 2018.
3. Liu et al. Towards Explainable NLP: A Generative Explanation Framework for Text Classification. ACL 2019.

Network Optimisation Algorithms

Supervisors: Professor Tomasz Radzik & Dr Kathleen Steinhofel

Description:

Network Optimisation problems are computational problems with input data referring to a network structure. Such problems occur in computer science, operations research, engineering, and applied mathematics. From the computer science point of view, the general objective of studying network optimisation problems is to develop efficient algorithms, which provide strict performance guarantees.

This project will focus on algorithms for network optimisation problems with the dynamic network structure, which changes over time. One of the applications is to provide efficient routing in networks where individual node-to-node links are not always available.

The Nexus between Crime, Mental Wellbeing and the Built Environment in Urban Areas

Supervisors: Dr Rita Borgo & Dr Andrea Mechelli

Description:

This project will explore the nexus between crime, mental wellbeing, and the built environment in urban areas, using London as a case study. Using a data-driven approach, the student will develop a holistic understanding of the spatial and temporal dynamics of crime. For instance, objective notions of crime such as real-time crime reports from Metropolitan Police can be compared with more subjective notions of how safe a place "feels", measured using crowdsourcing using the UrbanMind App. Spatial variation in crime levels and Temporal Dynamics (night vs. day or weekday vs. weekend) will be mapped. Machine learning on images from Google Street View, Flickr etc can shed light on how the built environment affects perceived notions of safety and whether it has an effect on actual crime, as hypothesised by the "Broken Window" theory. The results will be used to inform future work on urban wellbeing, as well as urban planning.

A Novel Model-driven AI Paradigm for Intrusion Detection

Supervisor: Dr Fabio Pierazzi

Description:

Intrusion Detection Systems (IDSs) are commonly deployed in networks and hosts to identify malicious activities representing misuse of computer systems. The numbers and types of attacks have been constantly increasing, and detection based on manually-defined signature is no longer a viable option. Hence, AI-powered IDS solutions have been explored to keep up the arms race and scale to new threats, but they are not yet deployed at scale in companies; this is mostly because such AI- powered systems cannot be trusted and are not interpretable [1], and they suffer from a lot of false positives preventing their applicability in real-world scenarios. In particular, a major limitation is that most existing solutions for AI-powered IDSs are data-driven, where the relationships learned from the data are often artifacts or domain-agnostic, and thus harder to trust and interpret even for network administrators.

This project aims to explore the design of a novel model-driven AI paradigm for intrusion detection, where expert knowledge is embedded in a model to characterize user behaviors (e.g., through formal logic [2]), with the purpose of identifying malicious activities with trust, interpretability and verifiability of the IDS decisions, in particular when deployed to real-world contexts. In other words, this project aims to advance the state-of-the-art in AI-powered IDSs by integrating expert knowledge in the models to achieve trust, interpretability and verifiability of decisions. This will increase the overall safety of protected users by making IDS systems more effective and reliable, and progress towards industry-wide deployment of AI-based solutions for intrusion detection.

References:

1. R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection." IEEE Symp. Security and Privacy, 2010.
2. S. Jajodia, N. Park, F. Pierazzi, A. Pugliese, E.Serra, G.I. Simari, V.S. Subrahmanian. A probabilistic logic of cyber deception. IEEE Transactions on Information Forensics and Security, 2017.

Participatory Agent-Based Modelling of Emergency Department Patient Flow

Supervisors: Dr Steffen Zschaler

Description:

Emergency health care is in crisis; the core "4-hour" KPI has not been met since 2015. Emergency Departments (EDs) are socio-technical systems with complex interactions between a wide range of actors and with their urban environment. To help predict how changes in practice will affect the 4-hour KPI while ensuring patient safety and quality of care, we have been developing agent-based models (ABMs) of EDs, which can provide explainable analyses of behaviour of complex systems emerging from the lower-level interaction of large numbers of agents. However, ABMs currently are implemented in Java or C++, making them too technical to be understood and manipulated by clinical decision makers. Hence, findings from ABM-based analyses are often not translated into interventions.

In this PhD project, you will explore how using domain-specific languages (DSLs) closely aligned with clinical staffs' conceptualisation of the ED environment will affect acceptance of ABM. We collaborate with King's College Hospital ED and Westminster City Council.

Personalised Medicine

Supervisors: Professor Costas Iliopoulos & Dr Sophia Tsoka

Description:

The explosion of human genomic data is a key driver of the current transition in healthcare to an era of personalised medicine. The correct assembly and subsequent analysis of this data is, therefore, crucial.

Algorithms can be designed to provide answers for the various and vast number of specific questions that collectively elucidate the (dis)functioning of a biological entity, at the most fundamental level.

These algorithms can be categorised based on their purposes. For example, pattern matching/discovery algorithms can find biologically significant motifs in sequences; alignment algorithms can identify the similarity between sequences; and compression algorithms can allow the latter two problems to be solved in a more time- and space-efficient manner.

Predictive Visual Analytics for Urban Contingency Planning

Supervisors: Dr Rita Borgo & Dr Grigorios Loukides

Description:

The aim of this project is to investigate the power of integrating predictive analytics and data visualization to address the challenge of generation, validation, and deployment of contingency plans in the context of urban related scenarios.

Based on initial work already conducted in this area by both supervisors, the project will investigate development and evaluation of:

a. algorithms for mining city rich data, both static (stored by the London Councils) and realtime (retrieved from mobile platforms, remote sensors, and social media), in order to predict emergencies efficiently and accurately;
b. novel visual encodings to enhance the diagnostic and predictive capabilities of mining algorithms, through their integration within a flexible visual analytics system capable of supporting and leveraging domain expert knowledge.
c. Application domain: contingency plans are employed across different organizations, from government to businesses, to minimise risk of catastrophic impacts of unexpected events. Research will have impact beyond the city remit.

Privacy in the Internet of Things

Supervisor: Professor Maribel Fernandez

Description:

Data Collection policies are used to restrict the kind of data transmitted by devices in the Internet of Things (e.g., health trackers, smart electricity meters, etc.) according to the privacy preferences of the user. The goal of this project is to develop cloud/IoT architectures with integrated data collection and data sharing models, to allow users to specify their own policies and trade data for services. For this, new data collection and data sharing models will have to be developed, with appropriate user interfaces, policy languages, and policy enforcement mechanisms. An important aspect of the project is the development of policy recommendation systems that can suggest/create policies based on user profiles, making privacy an integral part of the system (according to the privacy-by-design" IoT paradigm).

Programming as an HCI Challenge - IDE Interaction Design

Supervisors: Professor Michael Kolling & Dr Neil Brown

Description:

Frame-based editing with Stride [1] was a first attempt to revisit the design of program editing from an HCI perspective, in the context of novice programmers. What would it look like to approach professional IDEs from this perspective?

This project would take a Stride-like approach to professional tools and design, build and evaluate a new, better editor.

References:

1. [1] https://www.greenfoot.org/frames/

Program editor design for accessibility

Supervisors: Professor Michael Kolling & Dr Neil Brown

Description:

Most program editors use text for editing. Screen readers can be used with text-based editing by visually impaired programmers, but the syntax can often be confusing. Whitespace and punctuation are highly significant in program text but often omitted or have poor interaction with screen readers. Block-based editors rely less on syntax, so are potentially more suitable for accessible programming - but blocks are often manipulated only through drag-and-drop interactions which are ill-suited to visually impaired users. Our existing Stride editor combines keyboard interactions with structural programming but does not yet have support for accessibility tools. This project would look at improving the Stride editor to work well with accessibility, especially for vision-impaired users, including the design, implementation, and evaluation of the editor with actual users.

Programming history for learning and reflection

Supervisors: Professor Michael Kolling & Dr Neil Brown

Description:

Version control provides a way to store and view the history of program code. This is generally considered an advanced tool, used for collaborating or once a programmer is working on a large code base. This project would investigate the implications of using built-in automatic version control. Can this help during novice program development, can it help students in reflecting on their learning progress, and could it be used to provide more accurate programming assessment. This would involve the design, development, and multiple evaluations of automatic version-control in a beginner's IDE.

Security and Safety of Cyber-Physical Systems

Supervisor: Professor Luca Vigano

Description:

Cyber-Physical Systems (CPSs) are integrations of networking and distributed computing systems with physical processes and associated instrumentation that monitor and control entities in the physical environment, with feedback loops where physical processes affect computations and vice versa. Emerging applications of CPS include all the essential pieces of our social infrastructure: telecoms, banking, manufacturing, health energy, transportation, government smart cities. CPSs have effectively become one of the driving factors of the so-called fourth industrial revolution (Industry 4.0), but all the new opportunities opened by CPSs will only materialize if we can ensure their security and safety.

However, this need is often not addressed in current practice because of the major challenges that are posed by the heterogeneous and distributed nature of the systems and their interaction with the physical world and with the human users. Consequently, there has been a dramatic increase in the number of attacks, e.g., influencing physical processes to bring the system into an undesired state.

System failure can be extremely costly and threaten not only the system's environment but also human life.

The main aim of this PhD project will be to develop model-based AI techniques for representing, analysing, and reasoning about the security and safety of both the technical components of a CPS (control, computation, communication) and its social components (e.g., user interaction processes and user behavior) together and at the same time. The goal will be to overcome the limits of the state-of- the-art to devise methodologies and technologies for the formal validation of properties of CPSs to include the human element together with the technical in a holistic, socio-technical approach for security and safety, and to rebound the findings over the users through behavior change techniques. This will greatly simplify the design, development, deployment, and management of socio-technically secure CPSs, and thus have a disruptive and lasting impact.

Smart Metering Voice Controlled Devices

Supervisors: Dr Rita Borgo and Dr Alfie Abdul-Rahman

Description:

Communication is an integral part of our daily lives, and no communication mean is more significant than the human voice.

The advent of the Internet of Things (IoT) and advances in computing technologies and natural language processing have made possible to exploit voice recognition in the context of voice-controlled devices. Such devices capture user's spoken words and employ sophisticated AI frameworks to analyse, interpret and act out their inner intention. Such devices share the same challenges as any IoT device, that is Privacy of the information flow and Security of the system. Oftentimes vulnerability of voice-controlled devices resides not as much in potential cyber-attacks as rather in the technology itself and its maker's interpretation of user Privacy, key element to Trust in Autonomous Systems, and as a consequence Safety of their data and persona.

In this project we are interested to move the attention from the system itself to the user side. We will focus our attention to devices such as Alexa and Google Nest where the human-AI interaction pervades multiple levels of a user life context.

The volume of information exchange is large, varied, may touch critical aspects of one's life which may, or may not, be explicitly interconnected. In this scenario the nature and type of information fed to the AI through the user-AI communication channel matters greatly, yet there is no mean to return to the user control of such flow.

In this project we will leverage Visualization, Natural Language Processing and Human-Computer Interaction as means to model the human-AI dialogue, its domain and parameter space. Core to this will be the ability to:

a. characterise the nature of information flows (human-AI and vice versa);
b. develop metrics to estimate the level of privacy of information exchanged, volunteered by the user and pried by the system, in each communication flow;
c. develop metrics to estimate the level of privacy when cross referencing information exchanged within more than a single flow.
d. Where a flow can be seen as either a temporal instance or a thematic instance.

Starting from these core elements we will aim to define a novel framework capable of supporting human understanding and agency within the human-AI dialogue dynamic and its characteristics, with special focus on the specific context of home voice-activated devices. We aim to explore users' abilities to gauge the level of threat versus gain incurred through the use of voice-controlled devices. How to empower the user to make informed decisions with respect to which elements of a dialogue could potentially be released to the AI system and for which deletion should be required, appraise the level threats and risks associated with each choice.

Outcomes of the proposed research will be grounded in theoretical foundations, validated and verified through empirical evaluation. Methods to achieve the project goals will include and not limited to: symbolic modelling to map user vs system knowledge, sentiment analysis and topic modelling, information visualization, grounded theory.

Software Verification and Nominal Dependent Type Theory

Supervisor: Professor Maribel Fernandez

Description:

Dependent Type Theory is a mathematical tool to write formal specifications and prove the correctness of software implementations. The proof assistants used to certify the correctness of programs (such as Coq), are based on dependently typed higher-order abstract syntax. The goal of this project is to explore alternative foundations for proof assistants using nominal techniques. The nominal approach has roots in set theory and has been successfully used to specify programming languages. This project will focus on the combination of dependent types and nominal syntax and explore the connections between the nominal approach and the higher-order syntax approach used in current proof assistants.

String Sanitisation with Applications to Internet of Things Data

Supervisors: Dr Grigorios Loukides, Professor Costas Iliopoulos, Professor Luca Vigano

Description:

The overall aim of the project is to develop and evaluate a robust and efficient approach that allows organisations and businesses to protect the privacy of data represented as strings. The project will consider the protection of aggregated data (event sequences), as well as string databases, and it will also address the interrelated issues of usefulness, security, and scalability. It aims to develop a methodology (model, algorithms, protocols) for sanitising (i.e., transforming) data that is: (I) privacy-preserving, by designing and applying a privacy model along with algorithms for sanitising string data. (II) Utility-preserving, by designing measures and tools for quantifying the level of usefulness of data that must be traded-off for achieving privacy. (III) Secure and scalable, by designing efficient protocols that allow multiple parties to protect their data securely and jointly. The methodology will be evaluated on data from the Internet of Things (IoT) domain.

Temporal and Resource Controllability of Workflows of Autonomous Systems

Supervisor: Professor Luca Vigano

Workflow technology has long been employed for the modeling, validation and execution of business processes, and will play a crucial role in the design, development and maintenance of future autonomous systems. A workflow is a formal description of a business process in which single atomic work units (tasks), organized in a partial order, are assigned to processing entities (agents) to achieve some business goal(s). Workflows can also employ workflow paths in order (not) to execute a subset of tasks. A workflow management system coordinates the execution of tasks that are part of workflow instances such that all relevant constraints are eventually satisfied.

Temporal workflows specify business processes subject to temporal constraints such as controllable or uncontrollable durations, delays and deadlines. The choice of a workflow path may be controllable or not, considered either in isolation or in combination with uncontrollable durations. Access controlled workflows specify workflows in which users are authorized for task executions and authorization constraints say which users remain authorized to execute which tasks depending on who did what. Access controlled workflows may consider workflow paths too other than the uncertain availability of resources. When either a task duration or the choice of the workflow path to take or the availability of a user is out of control, we need to verify that the workflow can be executed by verifying all constraints for any possible combination of behaviors arising from the uncontrollable parts. Indeed, users might be absent before starting the execution (static resiliency), they can also become so during execution (decremental resiliency) or they can come and go throughout the execution (dynamic resiliency).

Temporal access-controlled workflows merge the two previous formalisms by considering several kinds of uncontrollable parts simultaneously. Authorization constraints may be extended to support conditional and temporal features.

This PhD project will aim to ensure the safety and trust of autonomous systems by reasoning about the temporal and resource controllability under uncertainty of the workflows that govern them.

Tracing Trust - Visual Frameworks for Explainable AI

Supervisors: Dr Rita Borgo, Dr Alfie Abdul-Rahman

Description:

Explainable Artificial Intelligence (XAI) is a topic receiving close review and increasing interest across different fields. Crucial to explainability is understanding of cause-effect relationships which in complex intelligent systems are anything but clear. Lack of ability to present the rationale behind a decision-making process inevitably mines trust and introduces uncertainty with respect to accountability of consequences.

The proposed research program will focus on the creation of a theoretical and applied framework to support the creation of systems to help people interpret the reasoning behind decisions made by AI systems. The project will entail design, implementation, and testing of visualization interfaces connecting to and integrating with explainable intelligent systems designed by partners.

This project places itself across three different fields: visual analytics, human-computer interaction, and artificial intelligence.

Hub relevance: Autonomous Systems

Unstructured Big Data

Supervisors: Professor Costas Iliopoulos & Dr Grigorios Loukides

Description:

A major challenge in today's society is the explosive growth of unstructured data such as text, images, videos and speech data. These forms of data exhibit the three characteristics of velocity, volume and variety that make processing and comprehending them a challenging task.

The initial processing of this data is invariably done using automated methods, as manual processing would be prohibitively expensive. The output of this automated processing is uncertain, either due to inaccuracies or inconsistencies in the raw data, or due to the automated processing. The database community has recognised this phenomenon in recent years, and several probabilistic formulations of uncertain data have been proposed, with a focus on processing SQL-like or ranking queries on such data. However, the science of mining, pattern analysis and pattern discovery on uncertain data expressed in probabilistic terms is very much in its infancy.

Mining probabilistic uncertain data to obtain reliable and actionable information is a critical challenge. Since the proliferation of "data science pipelines" uncertainties in one stage can propagate and magnify in later stages. It is essential both that uncertainty is processed appropriately by the system and that the data is not artificially made certain by, for example, choosing the most probable outcome at each stage.

The central hypothesis of this proposal is that the new field of algorithms on uncertain sequences that we propose is an important and broad foundation for representing and mining uncertain data arising in a wide variety of contexts. In addition, novel algorithmic techniques and ideas will be needed and could be useful for other high throughput data processing. The breadth of the proposed area investigation can be illustrated by the three abstract models given below:

a) **Probabilistic sequences:** they model a number of real-world data, as - DNA sequences, either to represent single nucleotide polymorphisms, or errors introduced by wet-lab sequencing platforms during the process of DNA sequencing.
    a. Converting sensor readings into meaningful human actions (e.g., accelerometer readings into kinds of human activity, using blood pressure/voice pitch to infer emotions) due to the process' intrinsically uncertainty.
    b. Software behaviour is often characterised in terms of sequences of events, such as the order of user interactions with a GUI or a webpage, the order of function invocations within a program, or the order in which network packets are sent to a server. A common assumption is that system behaviour is deterministic. It is however easy to envisage situations in which this assumption is violated. Network packets might arrive in a different order, depending on their route through a network, a unit of code might include stochastic behaviour to arise from random number generators, or different interleaving of concurrent processes.
b) **Uncertain Event Sequences:** These arise from several sources including measurement error, randomness in the underlying phenomenon, and due to distributed and asynchronous data gathering. They are used in several real-world scenarios to model and analyse spatial or temporal data, which is of interest in diverse disciplines as computational neuroscience, earth

science and telecommunications. Marked event sequences are even more general and can be applied to computer and economic systems for examples.

c) **Uncertain Time Series**: are most naturally associated with measurement errors, but can directly represent a range of variation (e.g. high/low stock prices in a day's trading, confidence intervals for predictions) or deliberate obfuscation for reasons of privacy preservation. They can be seen as special cases of event sequences, but while in uncertain time series the uncertainty lies in the value, in uncertain event sequences, the uncertainty is in the time that the event occurred.

Research Programme and Methodology:

We will focus on highly scalable methods discovering repetitive structures in uncertain sequences. Given the uncertainty in the underlying data, these repetitions will of necessity be approximate, rather than exact.

There are two major technical obstacles to overcome: firstly, classical measures of approximation (edit or Hamming distance) are inadequate to measure similarity between uncertain sequences.

a) One objective of this project is to define new, alternative, well-founded and powerful approaches for measuring similarity between sequences.
b) Secondly, we need to develop novel algorithmic techniques for solving problems in the context of uncertain sequences.

These problems are directly motivated by bioinformatics applications, such as studying genetic mutations; DNA sequence analysis of antibodies and identification of "hairpins" that occur in DNA sequences in Tuberculosis and HIV virus strains, respectively. However, they are also closely related to pattern discovery tasks that arise in other problem domains. Furthermore, they are also the most intensively studied problems in mining time series data and, to the best of our knowledge, these problems have not been considered in the uncertain time series framework, and it is not at all clear how to extend the known methods to this case. A solution is to build upon the experience we have in musical and biological computation pattern analysis, which share some characteristics with uncertain sequence processing, to suggest lines of attack.

Objectives:

A. Devise appropriate uncertain / probabilistic sequence formulations for modelling large- scale complex heterogeneous data.
B. Develop highly-scalable algorithms for pattern / motif discovery and sequential pattern mining in uncertain sequence data.
C. Build a theoretical framework for pattern discovery in dynamic, streaming and high-throughput uncertain sequence data.
D. Develop robust and well-founded methods for inferring actionable models of uncertain sequence data.
E. Devise appropriate and tractable formal frameworks for modelling stochastic dependencies in uncertain sequence data.

Predictive Profiling from Biometrics Data in Educational Environment.

Supervisor: Dr Tasmina Islam

Description:

Mental health and well-being of students is very important in achieving their full potential during academic studies in university [1]. Predicting their mental and emotional status can be very useful in monitoring student's well-being and providing the appropriate support at the time when needed. Although the principle focus of biometrics is identification/verification of individuals, biometric data can be used to predict some lower level (age, gender, ethnicity, etc.) and higher level (mental state, emotion etc.) individual characteristics [2]. Different biometric modalities (e.g., face, voice, EEG signals, keystroke, handwriting etc.) can be explored utilising this predictive capability to predict students' mental and emotional status that may have impact on their academic performance. As well as monitoring well-being, both physiological and behavioural biometrics can play a big role in facilitating education, for example, tracking attendance, monitoring engagement, and learning behaviour (especially when learning remotely). These could be beneficial for both students and educators.

Due to the wider use of biometrics, the analysis of biometric data poses some challenges if the biometric data is captured under unconstrained environment, for example, voice recognition in a crowd or with noise/echo, full or partly covered mouth (e.g., wearing a mask), face recognition in limited/unevenly distributed light, pose variations of individuals, noise like other people in the background, where some parts of the face is occluded (e.g., wearing a mask or a sunglass) and many more.

This project aims to explore different factors that affects the biometric recognition performance and investigate how to manage and improve the performance in facilitating education. The project will also explore the predictive capabilities of biometric data under both constrained and unconstrained environment. Prospective students can discuss about different modalities and options with the supervisor.

References:

1. Smith, A.P., 2019, November. Student workload, well-being, and academic attainment. In International Symposium on Human Mental Workload: Models and Applications (pp. 35-47). Springer, Cham.
2. Fairhurst, M., Li, C. and Da Costa-Abreu, M., 2017. Predictive biometrics: a review and analysis of predicting personal characteristics from biometric data. IET Biometrics, 6(6), pp.369-378

Dense subgraph detection and breaking

Supervisors: Dr. Grigorios Loukides and Dr. Dimitrios Letsios

Description:

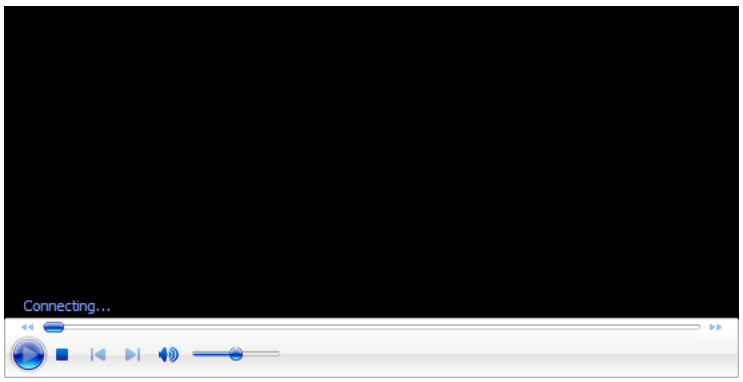
Graphs naturally model relationships between entities in domains ranging from social networks to communication networks and the web. In all these domains, a fundamental analysis task is dense subgraph discovery. This task aims at identifying parts of the graph that are cohesive. A clique, for example, is such a cohesive subgraph, while there are several other notions that relax the notion of clique, allowing for more efficient discovery. Dense subgraph discovery is important in a multitude of applications, such as detecting communities in social networks and preventing money laundering. Beyond discovery, studying how dense subgraphs may be broken with minimal graph distortion is also relevant.

It is practically useful in maintaining communities in social networks, assessing resilience to attacks or errors in communication networks, and enabling social network users to prevent discrimination. The main goal of this project is to develop algorithms for detecting and for breaking dense subgraphs. There is also possibility to propose new notions of dense subgraphs, or to develop algorithms employing existing notions for complex types of graphs.

Candidates with **strong background** in *algorithm design and optimization and with excellent programming skills* (preferably in C++) are welcome.

References:

1. Huiping Chen, Alessio Conte, Roberto Grossi, Grigorios Loukides, Solon P. Pissis, and Michelle Sweering. 2021. On Breaking Truss-Based Communities. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD '21). Association for Computing Machinery, New York, NY, USA, 117–126. https://dl.acm.org/doi/pdf/10.1145/3447548.3467365



Persuasive Natural Language Generation

Supervisor: Professor Yulan He

Keywords: Natural language processing, controllable language generation, user attitude detection

Persuasive natural language generation (NLG) aims to generate convincing arguments for attitude change. It has been found in [1] that persuasive NLG needs to build on psychological, linguistic, and technical concepts where multiple facets need to be considered. This project will build on our previous work on first learning latent discourse structure in argumentative text [2] in which the dynamic patterns of how the participants voice their opinions which affect the persuasion results can be automatically detected. It will then extend our approach for long-form text generation [3] by controlling the discourse structure. Being able to control structural discourse features is particularly important as it helps shape the argumentative structure for persuasive argumentation.

References:

1. Duerr, S. and Gloor, P.A., 2021. Persuasive Natural Language Generation--A Literature Review. arXiv preprint arXiv:2101.05786.
2. Zeng, J., Li, J., He, Y., Gao, C., Lyu, M. and King, I., 2020, April. What changed your mind: The roles of dynamic topics and discourse in argumentation process. In Proceedings of The Web Conference 2020 (pp. 1502-1513).
3. Adewoyin, R.A., Dutta, R. and He, Y., 2022. RSTGen: Imbuing Fine-Grained Interpretable Control into Long-FormText Generators. NAACL

Robust Explanations in Sequential Decision Making

Supervisor: Dr Nicola Paoletti

Description:

Explainable AI has become increasingly relevant, because in many domains, especially safety-critical ones, it is desirable to complement black-box deep learning (DL) models with comprehensible explanations of the models' predictions. The focus of this project is to investigate methods to derive robust explanations. An explanation E for some input x and prediction $F(x)$ is robust if E is sufficient to imply $F(x)$, that is, any input x' that shares the same explanation of x must produce the same prediction $F(x')=F(x)$. Robustness is clearly a desirable property, and yet it is neglected by most popular explainability methods.

Objectives:

Building on existing frameworks that solve this problem as a one of constraint satisfaction, this PhD project will contribute to developing approaches for robust explanations in sequential decision making processes. Such processes are found in AI planning, reinforcement learning, and control/cyber-physical systems, and they nowadays make use of DL models to e.g., represent the policy or the environment's dynamics. Unlike most explainability techniques that deal with input-output, i.e., one-step, predictions, the challenge here is to deal with sequence data that arise from multiple, inter-dependent steps taken over time. We envision two (non-exclusive) research directions in the context of explainability of sequential decision making:

(1) Temporal logic explanations: We will investigate methods for deriving temporal-logic (TL) explanations. TL is a popular language to specify correctness properties in verification and agent tasks in planning, and it allows to capture temporal relations in an expressive, human-interpretrable, and unambiguous manner. In this project you will develop the first methods to synthesise TL explanations that are both robust and, crucially, aware of model uncertainty.

(2) Causal explanations: In this project, you will develop methods for deriving causally-inspired explanations by investigating the duality between robust and counterfactual (CF) explanations. The latter, in the RL setting, can be defined as modifications to the agent's policy that are minor but can lead to a substantially different agent's reward. A robust explanation can be seen as a necessary set of the agent actions that do not admit CF explanations.

References:

1. Adadi, Amina, and Mohammed Berrada. "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)." IEEE access 6 (2018): 52138-52160.
2. La Malfa, Emanuele, et al. "On guaranteed optimal robust explanations for NLP models." Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2021), pages 2658-2665, 2021.
3. Bartocci, Ezio, et al. "Specification-based monitoring of cyber-physical systems: a survey on theory, tools and applications." Lectures on Runtime Verification. Springer, Cham, 2018. 135-175.
4. Tsirtsis, Stratis, Abir De, and Manuel Rodriguez. "Counterfactual explanations in sequential decision making under uncertainty." Advances in Neural Information Processing Systems 34 (2021): 30127-30139.